USE OF MICROVESICLES IN DIAGNOSIS, PROGNOSIS AND TREATMENT OF MEDICAL DISEASES AND CONDITIONS

CROSS REFERENCE TO RELATED APPLICATIONS

[0001] This application claims priority to US provisional applications 61/025,536, filed February 1, 2008 and 61/100,293, filed September 26, 2008, each of which is incorporated herein by reference in its entirety.

GOVERNMENTAL SUPPORT

[0002] This invention was made with Government support under grants NCI CA86355 and NCI CA69246 awarded by the National Cancer Institute. The Government has certain rights in the invention.

FIELD OF THE INVENTION

[0003] The present invention relates to the fields of medical diagnosis, patient monitoring, treatment efficacy evaluation, nucleic acid and protein delivery, and blood transfusion.

BACKGROUND OF THE INVENTION

[0004] Glioblastomas are highly malignant brain tumors with a poor prognosis despite intensive research and clinical efforts (Louis et al., 2007). The invasive nature of this tumor makes complete surgical resection impossible and the median survival time is only about 15 months (Stupp et al., 2005). Glioblastoma cells as well as many other tumor cells have a remarkable ability to mold their stromal environment to their own advantage. Tumor cells directly alter surrounding normal cells to facilitate tumor cell growth, invasion, chemo-resistance, immune-evasion and metastasis (Mazzocca et al., 2005; Muerkoster et al., 2004; Singer et al., 2007). The tumor cells also hijack the normal vasculature and stimulate rapid formation of new blood vessels to supply the tumor with nutrition (Carmeliet and Jain, 2000). Although the immune system can initially suppress tumor growth, it is often progressively blunted by tumor activation of immunosuppressive pathways (Gabrilovich, 2007).

[0005] Small microvesicles shed by cells are known as exosomes (Thery et al., 2002). Exosomes are reported as having a diameter of approximately 30-100 nm and are shed from many different cell types under both normal and pathological conditions (Thery et al., 2002). These microvesicles were first described as a mechanism to discard transferrin-receptors

from the cell surface of maturing reticulocytes (Pan and Johnstone, 1983). Exosomes are formed through inward budding of endosomal membranes giving rise to intracellular multivesicular bodies (MVB) that later fuse with the plasma membrane, releasing the exosomes to the exterior (Thery et al., 2002). However, there is now evidence for a more direct release of exosomes. Certain cells, such as Jurkat T-cells, are said to shed exosomes directly by outward budding of the plasma membrane (Booth et al., 2006). All membrane vesicles shed by cells are referred to herein collectively as microvesicles.

[0006] Microvesicles in *Drosophila melanogaster*, so called argosomes, are said to contain morphogens such as Wingless protein and to move over large distances through the imaginal disc epithelium in developing *Drosophila melanogaster* embryos (Greco et al., 2001). Microvesicles found in semen, known as prostasomes, are stated to have a wide range of functions including the promotion of sperm motility, the stabilization of the acrosome reaction, the facilitation of immunosuppression and the inhibition of angiogenesis (Delves et al., 2007). On the other hand, prostasomes released by malignant prostate cells are said to promote angiogenesis. Microvesicles are said to transfer proteins (Mack et al., 2000) and recent studies state that microvesicles isolated from different cell lines can also contain messenger RNA (mRNA) and microRNA (miRNA) and can transfer mRNA to other cell types (Baj-Krzyworzeka et al., 2006; Valadi et al., 2007).

[0007] Microvesicles derived from B-cells and dendritic cells are stated to have potent immuno-stimulatory and antitumor effects *in vivo* and have been used as antitumor vaccines (Chaput et al., 2005). Dendritic cell-derived microvesicles are stated to contain the co-stimulatory proteins necessary for T-cell activation, whereas most tumor cell-derived microvesicles do not (Wieckowski and Whiteside, 2006). Microvesicles isolated from tumor cells may act to suppress the immune response and accelerate tumor growth (Clayton et al., 2007; Liu et al., 2006a). Breast cancer microvesicles may stimulate angiogenesis, and platelet-derived microvesicles may promote tumor progression and metastasis of lung cancer cells (Janowska-Wieczorek et al., 2005; Millimaggi et al., 2007).

[0008] Cancers arise through accumulation of genetic alterations that promote unrestricted cell growth. It has been stated that each tumor harbors, on average, around 50-80 mutations that are absent in non-tumor cells (Jones et al., 2008; Parsons et al., 2008; Wood et al., 2007). Current techniques to detect these mutation profiles include the analysis of biopsy samples and the non-invasive analysis of mutant tumor DNA fragments circulating in bodily fluids such as blood (Diehl et al., 2008). The former method is invasive, complicated and

possibly harmful to subjects. The latter method inherently lacks sensitivity due to the extremely low copy number of mutant cancer DNA in bodily fluid (Gormally et al., 2007). Therefore, one challenge facing cancer diagnosis is to develop a diagnostic method that can detect tumor cells at different stages non-invasively, yet with high sensitivity and specificity. It has also been stated that gene expression profiles (encoding mRNA or microRNA) can distinguish cancerous and non-cancerous tissue (Jones et al., 2008; Parsons et al., 2008; Schetter et al., 2008). However, current diagnostic techniques to detect gene expression profiles require intrusive biopsy of tissues. Some biopsy procedures cause high risk and are potentially harmful. Moreover, in a biopsy procedure, tissue samples are taken from a limited area and may give false positives or false negatives, especially in tumors which are heterogeneous and/or dispersed within normal tissue. Therefore, a non-intrusive and sensitive diagnostic method for detecting biomarkers would be highly desirable.

SUMMARY OF THE INVENTION

[0009] In general, the invention is a novel method for detecting in a subject the presence or absence of a variety of biomarkers contained in microvesicles, thereby aiding the diagnosis, monitoring and evaluation of diseases, other medical conditions, and treatment efficacy associated with microvesicle biomarkers.

[0010] One aspect of the invention are methods for aiding in the diagnosis or monitoring of a disease or other medical condition in a subject, comprising the steps of: a) isolating a microvesicle fraction from a biological sample from the subject; and b) detecting the presence or absence of a biomarker within the microvesicle fraction, wherein the biomarker is associated with the disease or other medical condition. The methods may further comprise the step or steps of comparing the result of the detection step to a control (e.g., comparing the amount of one or more biomarkers detected in the sample to one or more control levels), wherein the subject is diagnosed as having the disease or other medical condition (e.g., cancer) if there is a measurable difference in the result of the detection step as compared to a control.

[0011] Another aspect of the invention is a method for aiding in the evaluation of treatment efficacy in a subject, comprising the steps of: a) isolating a microvesicle fraction from a biological sample from the subject; and b) detecting the presence or absence of a biomarker within the microvesicle fraction, wherein the biomarker is associated with the treatment efficacy for a disease or other medical condition. The method may further

comprise the step of providing a series of a biological samples over a period of time from the subject. Additionally, the method may further comprise the step or steps of determining any measurable change in the results of the detection step (e.g., the amount of one or more detected biomarkers) in each of the biological samples from the series to thereby evaluate treatment efficacy for the disease or other medical condition.

[0012] In certain preferred embodiments of the foregoing aspects of the invention, the biological sample from the subject is a sample of bodily fluid. Particularly preferred body fluids are blood and urine.

[0013] In certain preferred embodiments of the foregoing aspects of the invention, the methods further comprise the isolation of a selective microvesicle fraction derived from cells of a specific type (e.g., cancer or tumor cells). Additionally, the selective microvesicle fraction may consist essentially of urinary microvesicles.

[0014] In certain embodiments of the foregoing aspects of the invention, the biomarker associated with a disease or other medical condition is i) a species of nucleic acid; ii) a level of expression of one or more nucleic acids; iii) a nucleic acid variant; or iv) a combination of any of the foregoing. Preferred embodiments of such biomarkers include messenger RNA, microRNA, DNA, single stranded DNA, complementary DNA and noncoding DNA.

[0015] In certain embodiments of the foregoing aspects of the invention, the disease or other medical condition is a neoplastic disease or condition (e.g., glioblastoma, pancreatic cancer, breast cancer, melanoma and colorectal cancer), a metabolic disease or condition (e.g., diabetes, inflammation, perinatal conditions or a disease or condition associated with iron metabolism), a post transplantation condition, or a fetal condition.

[0016] Another aspect of the invention is a method for aiding in the diagnosis or monitoring of a disease or other medical condition in a subject, comprising the steps of a) obtaining a biological sample from the subject; and b) determining the concentration of microvesicles within the biological sample.

[0017] Yet another aspect of this invention is a method for delivering a nucleic acid or protein to a target cell in an individual comprising the steps of administering microvesicles which contain the nucleic acid or protein, or one or more cells that produce such microvesicles, to the individual such that the microvesicles enter the target cell of the individual. In a preferred embodiment of this aspect of the invention, microvesicles are delivered to brain cells.

[0018] A further aspect of this invention is a method for performing bodily fluid transfusion (e.g., blood, serum or plasma), comprising the steps of obtaining a fraction of donor body fluid free of all or substantially all microvesicles, or free of all or substantially all microvesicles from a particular cell type (e.g., tumor cells), and introducing the microvesicle-free fraction to a patient. A related aspect of this invention is a composition of matter comprising a sample of body fluid (e.g., blood, serum or plasma) free of all or substantially all microvesicles, or free of all or substantially all microvesicles from a particular cell type.

[0019] Another aspect of this invention is a method for performing bodily fluid transfusion (e.g., blood, serum or plasma), comprising the steps of obtaining a microvesicle-enriched fraction of donor body fluid and introducing the microvesicle-enriched fraction to a patient. In a preferred embodiment, the fraction is enriched with microvesicles derived from a particular cell type. A related aspect of this invention is a composition of matter comprising a sample of bodily fluid (e.g., blood, serum or plasma) enriched with microvesicles.

[0020] A further aspect of this invention is a method for aiding in the identification of new biomarkers associated with a disease or other medical condition, comprising the steps of obtaining a biological sample from a subject; isolating a microvesicle fraction from the sample; and detecting within the microvesicle fraction species of nucleic acid; their respective expression levels or concentrations; nucleic acid variants; or combinations thereof.

[0021] Various aspects and embodiments of the invention will now be described in detail. It will be appreciated that modification of the details may be made without departing from the scope of the invention. Further, unless otherwise required by context, singular terms shall include pluralities and plural terms shall include the singular.

[0022] All patents, patent applications, and publications identified are expressly incorporated herein by reference for the purpose of describing and disclosing, for example, the methodologies described in such publications that might be used in connection with the present invention. These publications are provided solely for their disclosure prior to the filing date of the present application. Nothing in this regard should be construed as an admission that the inventors are not entitled to antedate such disclosure by virtue of prior invention or for any other reason. All statements as to the date or representations as to the contents of these documents are based on the information available to the applicants and do not constitute any admission as to the correctness of the dates or contents of these documents.

BRIEF DESCRIPTION OF THE DRAWINGS

FIGURE 1. Glioblastoma cells produce microvesicles containing RNA. (a) Scanning electron microscopy image of a primary glioblastoma cell (bar = 10 μm). (b) Higher magnification showing the microvesicles on the cell surface. The vesicles vary in size with diameters between around 50 nm and around 500 nm (bar = 1 μm). (c) Graph showing the amount of total RNA extracted from microvesicles that were either treated or not treated with RNase A. The amounts are indicated as the absorption (Abs, y-axis) of 260nm wavelength (x-axis). The experiments were repeated 5 times and a representative graph is shown. (d) Bioanalyzer data showing the size distribution of total RNA extracted from primary glioblastoma cells and (e) Bioanalyzer data showing the size distribution of total RNA extracted from microvesicles isolated from primary glioblastoma cells. The 25 nt peak represents an internal standard. The two prominent peaks in (d) (arrows) represent 18S (left arrow) and 28S (right arrow) ribosomal RNA. The ribosomal peaks are absent from RNA extracted from microvesicles (e). (f) Transmission electron microscopy image of microvesicles secreted by primary glioblastoma cells (bar = 100 nm).

FIGURE 2. Analysis of microvesicle RNA. FIGS. 2 (a) and 2 (b) are scatter plots of mRNA levels in microvesicles and mRNA levels in donor glioblastoma cells from two different experiments. Linear regressions show that mRNA levels in donor cells versus microvesicles were not well correlated. FIGS. 2 (c) and 2 (d) are mRNA levels in two different donor cells or two different microvesicle preparations. In contrast to FIGS. 2 (a) and 2 (b), linear regressions show that mRNA levels between donor cells FIG 2 (c) or between microvesicles FIG 2 (d) were closely correlated.

FIGURE 3. Analysis of microvesicle DNA.

a) GAPDH gene amplification with DNA templates from exosomes treated with DNase prior to nucleic acid extraction. The lanes are identified as follows:

1. 100bp MW ladder

2. Negative control

3. Genomic DNA control from GBM 20/3 cells

4. DNA from normal serum exosomes (tumor cell-free control)

5. Exosome DNA from normal human fibroblasts (NHF19)

6. Exosome DNA from primary medulloblastoma cells (D425)

b) GAPDH gene amplification with DNA templates from exosomes without prior DNase treatment. The lanes are identified as follows:

1. 100bp MW ladder
2. DNA from primary melanoma cell 0105
3. Exosome DNA from melanoma 0105
4. Negative Control
5. cDNA from primary GBM 20/3 (positive control)

c) Human Endogenous Retrovirus K gene amplification. The lanes are identified as follows:

1. 100 bp MW ladder
2. Exosome DNA from medulloblastoma D425 a
3. Exosome DNA from medulloblasotma D425 b
4. Exosome DNA from normal human fibroblasts (NHF19)
5. Exosome DNA from normal human serum
6. Genomic DNA from GBM 20/3.
7. Negative Control

d) Tenascin C gene amplification. The lanes are listed identified follows:

1. 100bp MW ladder
2. Exosomes from normal human fibroblasts (NHF19)
3. Exosomes from serum (tumor cell free individual A)
4. Exosomes from serum (tumor cell free individual B)
5. Exosomes from primary medulloblastoma cell D425

e) Transposable Line 1 element amplification. The lanes are identified as follows:

1. 100bp MW ladder.
2. Exosome DNA from normal human serum.
3. Exosome DNA from normal human fibroblasts
4. Exosome DNA from medulloblastoma D425 a

5. Exosome DNA from medulloblastoma D425 b

f) DNA is present in exosomes from D425 medulloblastoma cell. The lanes are identified as follows:

1. 100bp marker
2. D425 no DNase
3. D425 with DNase
4. 1kb marker

g) Single stranded nucleic acid analysis using a RNA pico chip. Upper panel: purified DNA was not treated with DNase; lower panel: purified DNA was treated with DNase. The arrow in the upper panel refers to the detected nucleic acids. The peak at 25 nt is an internal standard.

h) Analysis of nucleic acids contained in exosomes from primary medulloblastoma D425. Upper panel: single stranded nucleic acids detected by a RNA pico chip. Lower panel: double stranded nucleic acids detected by a DNA 1000 chip. The arrow in the upper panel refers to the detected nucleic acids. The two peaks (15 and 1500 bp) are internal standards.

i) Analysis of exosome DNA from different origins using a RNA pico chip. Upper panel: DNA was extracted from exosomes from glioblastoma cells. Lower panel: DNA was extracted from exosomes from normal human fibroblasts.

FIGURE 4. Extracellular RNA extraction from serum is more efficient when a serum exosome isolation step is included. a) Exosome RNA from serum. b) Direct whole serum extraction. c) Empty well. Arrows refer to the detected RNA in the samples.

FIGURE 5. Comparison of gene expression levels between microvesicles and cells of origin. 3426 genes were found to be more than 5-fold differentially distributed in the microvesicles as compared to the cells from which they were derived (p-value <0.01).

FIGURE 6. Ontological analysis of microvesicular RNAs. (a) Pie chart displays the biological process ontology of the 500 most abundant mRNA species in the microvesicles. (b) Graph showing the intensity of microvesicle RNAs belonging to ontologies related to tumor growth. The x-axis represents the number of mRNA transcripts present in the ontology. The median intensity levels on the arrays were 182.

FIGURE 7. Clustering diagram of mRNA levels. Microarray data on the mRNA expression profiles in cell lines and exosomes isolated from the culture media of these cell lines were analyzed and clusters of expression profiles were generated. The labels of the RNA species are as follows:

20/3C-1: Glioblastoma 20/3 Cell RNA, array replicate 1

20/3C-2: Glioblastoma 20/3 Cell RNA, array replicate 2

11/5C: Glioblastoma 11/5 Cell RNA

0105C: Melanoma 0105 Cell RNA

0664C: Melanoma 0664 Cell RNA

0664 E-1: Melanoma 0664 exosome RNA, array replicate 1

0664 E-2: Melanoma 0664 exosome RNA, array replicate 2

0105E: Melanoma 0105 Exosome RNA

20/3E: Glioblastoma 20/3 Exosome RNA

11/5E-1: Glioblastoma 11/5 Exosomes, array replicate 1

11/5E-2: Glioblastoma 11/5 Exosomes, array replicate 2

GBM: glioblastoma. The scale refers to the distance between clusters.

FIGURE 8. Microvesicles from serum contain microRNAs. Levels of mature miRNAs extracted from microvesicles and from glioblastoma cells from two different patients (GBM1 and GBM2) were analysed using quantitative miRNA RT-PCR. The cycle threshold (Ct) value is presented as the mean ± SEM (n = 4).

FIGURE 9. Clustering diagram of microRNA levels. Microarray data on the microRNA expression profiles in cell lines and exosomes isolated from the culture media of these cell lines were analyzed and clusters of expression profiles were generated. The labels of the RNA species are as follows:

0664C-1: Melanoma 0664 Cell RNA, array replicate 1

0664C-2: Melanoma 0664 Cell RNA, array replicate 2

0105C-1: Melanoma 0105 Cell RNA, array replicate 1

0105C-2: Melanoma 0105 Cell RNA, array replicate 2

20/3C-1: Glioblastoma 20/3 Cell RNA, array replicate 1

20/3C-2: Glioblastoma 20/3 Cell RNA, array replicate 2

11/5C-1: Glioblastoma 11/5 Cell RNA, array replicate 1

11/5C-2: Glioblastoma 11/5 Cell RNA, array replicate 2

11/5E-1: Glioblastoma 11/5 Exosomes, array replicate 1

11/5E-2: Glioblastoma 11/5 Exosomes, array replicate 2

20/3E-1: Glioblastoma 20/3 Exosome RNA, array replicate 1

20/3E-2: Glioblastoma 20/3 Exosome RNA, array replicate 2

0664 E: Melanoma 0664 exosome RNA

0105E-1: Melanoma 0105 Exosome RNA, array replicate 1

0105E-2: Melanoma 0105 Exosome RNA, array replicate 2

GBM: Glioblastoma. The scale refers to the distance between clusters.

FIGURE 10. The expression level of microRNA-21 in serum microvesicles is associated with glioma. Shown is a bar graph, normal control serum on the left, glioma serum on the right. Quantitative RT-PCR was used to measure the levels of microRNA-21 (miR-21) in exosomes from serum of glioblastoma patients as well as normal patient controls. Glioblastoma serum showed a 5.4 reduction of the Ct-value, corresponding to an approximately 40 ($2^{\Delta Ct}$)-fold increase of miR21. The miR21 levels were normalized to GAPDH in each sample (n=3).

FIGURE 11. Nested RT-PCR was used to detect EGFRvIII mRNA in tumor samples and corresponding serum exosomes. The wild type EGFR PCR product appears as a band at 1153 bp and the EGFRvIII PCR product appears as a band at 352 bp. RT PCR of GAPDH mRNA was included as a positive control (226 bp). Samples considered as positive for EGFRvIII are indicated with an asterisk. Patients 11, 12 and 14 showed only a weak amplification of EGFRvIII in the tumor sample, but it was evident when more samples were loaded.

FIGURE 12. Nested RT PCR of EGFRvIII was performed on microvesicles from 52 normal control serums. EGFRvIII (352 bp) was never found in the normal control serums. PCR of GAPDH (226 bp) was included as a control.

FIGURE 13. Hepcidin mRNA can be detected within exosomes from human serum. A) Pseudo-gel generated by an Agilent Bioanalyzer. B) Raw plot generated by an Agilent Bioanalyser for the positive control (Sample 1). C) Raw plot generated by an Agilent Bioanalyser for the negative control (Sample 2). D) Raw plot generated by an Agilent Bioanalyser for the exosomes (Sample 3).

FIGURE 14. Urinary exosome isolation and nucleic acid identification within urinary exosomes. (a) Electron microscopy image of a multivesicular body (MVB) containing many small "exosomes" in a kidney tubule cell. (b) Electron microscopy image of isolated urinary exosomes. (c) Analysis of RNA transcripts contained within urinary exosomes by an Agilent Bioanalyzer. A broad range of RNA species were identified but both 18S and 28S ribosomal RNAs were absent. (d) Identification of various RNA transcripts in urinary exosomes by PCR. The transcripts thus identified were: Aquaporin 1 (AQP1); Aquaporin 2 (AQP2); Cubulin (CUBN); Megalin (LRP2); Arginine Vasopressin Receptor 2 (AVPR2); Sodium/Hydrogen Exchanger 3 (SLC9A3); V-ATPase B1 subunit (ATP6V1B1); Nephrin (NPHS1); Podocin (NPHS2); and Chloride Channel 3 (CLCN3). From top to bottom, the five bands in the molecular weight (MW) lane correspond to 1000, 850, 650, 500, 400, 300 base pair fragments. (e) Bioanalyzer diagrams of exosomal nucleic acids from urine samples. The numbers refer to the numbering of human individuals. (f) Pseudogels depicting PCR products generated with different primer pairs using the nucleic acid extracts described in (e). House refers to actin gene and the actin primers were from Ambion (TX, USA). The +ve control refers to PCRs using human kidney cDNA from Ambion (TX, USA) as templates and the –ve control refers to PCRs without nucleic acid templates. (g) Pseudo-gel picture showing positive identification of actin gene cDNA via PCR with and without the DNase treatment of exosomes prior to nucleic acid extraction. (h) Bioanalyzer diagrams showing the amount of nucleic acids isolated from human urinary exosomes.

FIGURE 15. Analysis of prostate cancer biomarkers in urinary exosomes. (a) Gel pictures showing PCR products of the *TMPRSS2-ERG* gene and digested fragments of the PCR products. P1 and P2 refer to urine samples from patient 1 and patient 2, respectively. For each sample, the undigested product is in the left lane and the digested product is in the right lane. MWM indicates lanes with MW markers. The sizes of the bands (both undigested and digested) are indicated on the right of the panel. (b) Gel pictures showing PCR products of the *PCA3* gene and digested fragments of the PCR products. P1, P2, P3 and P4 refer to urine samples from patient 1, patient 2, patient 3 and patient 4, respectively. For each

sample, the undigested product is in the left lane and the digested product is in the right lane. MWM indicates lanes with MW markers. The sizes of the bands (both undigested and digested) are indicated on the right of the panel. (c) A summary of the information of the patients and the data presented in (a) and (b). TMERG refers to the *TMPRSS2-ERG* fusion gene.

FIGURE 16. BRAF mRNA is contained within microvesicles shed by melanoma cells. (a) An electrophoresis gel picture showing RT-PCR products of BRAF gene amplification. (b) An electrophoresis gel picture showing RT-PCR products of GAPDH gene amplification. The lanes and their corresponding samples are as follows: Lane #1 - 100 bp Molecular Weight marker; Lane #2 - YUMEL-01-06 exo; Lane # 3 - YUMEL-01-06 cell; Lane # 4 YUMEL-06-64 exo; Lane # 5. YUMEL-06-64 cell; Lane # 6. M34 exo; Lane # 7 - M34 cell; Lane # 8 - Fibroblast cell; Lane # 9 - Negative control. The reference term "exo" means that the RNA was extracted from exosomes in the culture media. The reference term "cell" means that the RNA was extracted from the cultured cells. The numbers following YUMEL refers to the identification of a specific batch of YUMEL cell line. (c) Sequencing results of PCR products from YUMEL-01-06 exo. The results from YUMEL-01-06 cell, YUMEL-06-64 exo and YUMEL-06-64 cell are the same as those from YUMEL-01-06 exo. (d) Sequencing results of PCR products from M34 exo. The results from M34 cell are the same as those from M34 exo.

FIGURE 17. Glioblastoma microvesicles can deliver functional RNA to HBMVECs. (a) Purified microvesicles were labelled with membrane dye PKH67 (green) and added to HBMVECs. The microvesicles were internalised into endosome-like structures within an hour. (b) Microvesicles were isolated from glioblastoma cells stably expressing Gluc. RNA extraction and RT-PCR of Gluc and GAPDH mRNAs showed that both were incorporated into microvesicles. (c) Microvesicles were then added to HBMVECs and incubated for 24 hours. The Gluc activity was measured in the medium at 0, 15 and 24 hours after microvesicle addition and normalized to Gluc activity in microvesicles. The results are presented as the mean ± SEM (n = 4).

FIGURE 18. Glioblastoma microvesicles stimulate angiogenesis *in vitro* and contain angiogenic proteins. (a) HBMVECs were cultured on Matrigel™ in basal medium (EBM) alone, or supplemented with GBM microvesicles (EBM+MV) or angiogenic factors (EGM). Tubule formation was measured after 16 hours as average tubule length ± SEM compared to cells grown in EBM (n = 6). (b) Total protein from primary glioblastoma cells and

microvesicles (MV) from these cells (1 mg each) were analysed on a human angiogenesis antibody array. (c) The arrays were scanned and the intensities analysed with the Image J software (n = 4).

FIGURE 19. Microvesicles isolated from primary glioblastoma cells promote proliferation of the U87 glioblastoma cell line. 100,000 U87 cells were seeded in wells of a 24 well plate and allowed to grow for three days in (a) normal growth medium (DMEM-5% FBS) or (b) normal growth medium supplemented with 125 μg microvesicles. (c) After three days, the non-supplemented cells had expanded to 480,000 cells, whereas the microvesicle-supplemented cells had expanded to 810,000 cells. NC refers to cells grown in normal control medium and MV refers to cells grown in medium supplemented with microvesicles. The result is presented as the mean ± SEM (n=6).

DETAILED DESCRIPTION OF THE INVENTION

[0023] Microvesicles are shed by eukaryotic cells, or budded off of the plasma membrane, to the exterior of the cell. These membrane vesicles are heterogeneous in size with diameters ranging from about 10nm to about 5000 nm. The small microvesicles (approximately 10 to 1000nm, and more often 30 to 200 nm in diameter) that are released by exocytosis of intracellular multivesicular bodies are referred to in the art as "exosomes". The methods and compositions described herein are equally applicable to microvesicles of all sizes; preferably 30 to 800 nm; and more preferably 30 to 200 nm.

[0024] In some of the literature, the term "exosome" also refers to protein complexes containing exoribonucleases which are involved in mRNA degradation and the processing of small nucleolar RNAs (snoRNAs), small nuclear RNAs (snRNAs) and ribosomal RNAs (rRNA) (Liu et al., 2006b; van Dijk et al., 2007). Such protein complexes do not have membranes and are not "microvesicles" or "exosomes" as those terms are used here in.

Exosomes As Diagnostic And/Or Prognostic Tools

[0025] Certain aspects of the present invention are based on the surprising finding that glioblastoma derived microvesicles can be isolated from the serum of glioblastoma patients. This is the first discovery of microvesicles derived from cells in the brain, present in a bodily fluid of a subject. Prior to this discovery it was not known whether glioblastoma cells produced microvesicles or whether such microvesicles could cross the blood brain barrier into the rest of the body. These microvesicles were found to contain mutant mRNA associated with tumor cells. The microvesicles also contained microRNAs (miRNAs) which

were found to be abundant in glioblastomas. Glioblastoma-derived microvesicles were also found to potently promote angiogenic features in primary human brain microvascular endothelial cells (HBMVEC) in culture. This angiogenic effect was mediated at least in part through angiogenic proteins present in the microvesicles. The nucleic acids found within these microvesicles, as well as other contents of the microvesicles such as angiogenic proteins, can be used as valuable biomarkers for tumor diagnosis, characterization and prognosis by providing a genetic profile. Contents within these microvesicles can also be used to monitor tumor progression over time by analyzing if other mutations are acquired during tumor progression as well as if the levels of certain mutations are becoming increased or decreased over time or over a course of treatment

[0026] Certain aspects of the present invention are based on the finding that microvesicles are secreted by tumor cells and circulating in bodily fluids. The number of microvesicles increases as the tumor grows. The concentration of the microvesicles in bodily fluids is proportional to the corresponding tumor load. The bigger the tumor load, the higher the concentration of microvesicles in bodily fluids.

[0027] Certain aspects of the present invention are based on another surprising finding that most of the extracellular RNAs in bodily fluid of a subject are contained within microvesicles and thus protected from degradation by ribonucleases. As shown in Example 3, more than 90% of extracellular RNA in total serum can be recovered in microvesicles.

[0028] One aspect of the present invention relates to methods for detecting, diagnosing, monitoring, treating or evaluating a disease or other medical condition in a subject by determining the concentration of microvesicles in a biological sample. The determination may be performed using the biological sample without first isolating the microvesicles or by isolating the microvesicles first.

[0029] Another aspect of the present invention relates to methods for detecting, diagnosing, monitoring, treating or evaluating a disease or other medical condition in a subject comprising the steps of, isolating exosomes from a bodily fluid of a subject, and analyzing one or more nucleic acids contained within the exosomes. The nucleic acids are analyzed qualitatively and/or quantitatively, and the results are compared to results expected or obtained for one or more other subjects who have or do not have the disease or other medical condition. The presence of a difference in microvesicular nucleic acid content of the subject, as compared to that of one or more other individuals, can indicate the presence or

absence of, the progression of (e.g., changes of tumor size and tumor malignancy), or the susceptibility to a disease or other medical condition in the subject.

[0030] Indeed, the isolation methods and techniques described herein provide the following heretofore unrealized advantages: 1) the opportunity to selectively analyze disease- or tumor-specific nucleic acids, which may be realized by isolating disease- or tumor-specific microvesicles apart from other microvesicles within the fluid sample; 2) significantly higher yield of nucleic acid species with higher sequence integrity as compared to the yield/integrity obtained by extracting nucleic acids directly from the fluid sample; 3) scalability, e.g. to detect nucleic acids expressed at low levels, the sensitivity can be increased by pelleting more microvesicles from a larger volume of serum; 4) purer nucleic acids in that protein and lipids, debris from dead cells, and other potential contaminants and PCR inhibitors are excluded from the microvesicle pellets before the nucleic acid extraction step; and 5) more choices in nucleic acid extraction methods as microvesicle pellets are of much smaller volume than that of the starting serum, making it possible to extract nucleic acids from these microvesicle pellets using small volume column filters.

[0031] The microvesicles are preferably isolated from a sample taken of a bodily fluid from a subject. As used herein, a "bodily fluid" refers to a sample of fluid isolated from anywhere in the body of the subject, preferably a peripheral location, including but not limited to, for example, blood, plasma, serum, urine, sputum, spinal fluid, pleural fluid, nipple aspirates, lymph fluid, fluid of the respiratory, intestinal, and genitourinary tracts, tear fluid, saliva, breast milk, fluid from the lymphatic system, semen, cerebrospinal fluid, intra-organ system fluid, ascitic fluid, tumor cyst fluid, amniotic fluid and combinations thereof.

[0032] The term "subject" is intended to include all animals shown to or expected to have microvesicles. In particular embodiments, the subject is a mammal, a human or nonhuman primate, a dog, a cat, a horse, a cow, other farm animals, or a rodent (e.g. mice, rats, guinea pig. etc.). The term "subject" and "individual" are used interchangeably herein.

[0033] Methods of isolating microvesicles from a biological sample are known in the art. For example, a method of differential centrifugation is described in a paper by Raposo et al. (Raposo et al., 1996) , and similar methods are detailed in the Examples section herein. Methods of anion exchange and/or gel permeation chromatography are described in US Patent Nos. 6,899,863 and 6,812,023. Methods of sucrose density gradients or organelle electrophoresis are described in U.S. Patent No. 7,198,923. A method of magnetic activated

cell sorting (MACS) is described in (Taylor and Gercel-Taylor, 2008). A method of nanomembrane ultrafiltration concentrator is described in (Cheruvanky et al., 2007). Preferably, microvesicles can be identified and isolated from bodily fluid of a subject by a newly developed microchip technology that uses a unique microfluidic platform to efficiently and selectively separate tumor derived microvesicles. This technology, as described in a paper by Nagrath et al. (Nagrath et al., 2007), can be adapted to identify and separate microvesicles using similar principles of capture and separation as taught in the paper. Each of the foregoing references is incorporated by reference herein for its teaching of these methods.

[0034] In one embodiment, the microvesicles isolated from a bodily fluid are enriched for those originating from a specific cell type, for example, lung, pancreas, stomach, intestine, bladder, kidney, ovary, testis, skin, colorectal, breast, prostate, brain, esophagus, liver, placenta, fetus cells. Because the microvesicles often carry surface molecules such as antigens from their donor cells, surface molecules may be used to identify, isolate and/or enrich for microvesicles from a specific donor cell type (Al-Nedawi et al., 2008; Taylor and Gercel-Taylor, 2008). In this way, microvesicles originating from distinct cell populations can be analyzed for their nucleic acid content. For example, tumor (malignant and non-malignant) microvesicles carry tumor-associated surface antigens and may be detected, isolated and/or enriched via these specific tumor-associated surface antigens. In one example, the surface antigen is epithelial-cell-adhesion-molecule (EpCAM), which is specific to microvesicles from carcinomas of lung, colorectal, breast, prostate, head and neck, and hepatic origin, but not of hematological cell origin (Balzar et al., 1999; Went et al., 2004). In another example, the surface antigen is CD24, which is a glycoprotein specific to urine microvesicles (Keller et al., 2007). In yet another example, the surface antigen is selected from a group of molecules CD70, carcinoembryonic antigen (CEA), EGFR, EGFRvIII and other variants, Fas ligand, TRAIL, tranferrin receptor, p38.5, p97 and HSP72. Additionally, tumor specific microvesicles may be characterized by the lack of surface markers, such as CD80 and CD86.

[0035] The isolation of microvesicles from specific cell types can be accomplished, for example, by using antibodies, aptamers, aptamer analogs or molecularly imprinted polymers specific for a desired surface antigen. In one embodiment, the surface antigen is specific for a cancer type. In another embodiment, the surface antigen is specific for a cell type which is not necessarily cancerous. One example of a method of microvesicle separation based on cell

surface antigen is provided in U.S. Patent No. 7,198,923. As described in, e.g., U.S. Patent Nos. 5,840,867 and 5,582,981, WO/2003/050290 and a publication by Johnson et al. (Johnson et al., 2008), aptamers and their analogs specifically bind surface molecules and can be used as a separation tool for retrieving cell type-specific microvesicles. Molecularly imprinted polymers also specifically recognize surface molecules as described in, e.g., US Patent Nos. 6,525,154, 7,332,553 and 7,384,589 and a publication by Bossi et al. (Bossi et al., 2007) and are a tool for retrieving and isolating cell type-specific microvesicles. Each of the foregoing reference is incorporated herein for its teaching of these methods.

[0036] It may be beneficial or otherwise desirable to extract the nucleic acid from the exosomes prior to the analysis. Nucleic acid molecules can be isolated from a microvesicle using any number of procedures, which are well-known in the art, the particular isolation procedure chosen being appropriate for the particular biological sample. Examples of methods for extraction are provided in the Examples section herein. In some instances, with some techniques, it may also be possible to analyze the nucleic acid without extraction from the microvesicle.

[0037] In one embodiment, the extracted nucleic acids, including DNA and/or RNA, are analyzed directly without an amplification step. Direct analysis may be performed with different methods including, but not limited to, the nanostring technology. NanoString technology enables identification and quantification of individual target molecules in a biological sample by attaching a color coded fluorescent reporter to each target molecule. This approach is similar to the concept of measuring inventory by scanning barcodes. Reporters can be made with hundreds or even thousands of different codes allowing for highly multiplexed analysis. The technology is described in a publication by Geiss et al. (Geiss et al., 2008) and is incorporated herein by reference for this teaching.

[0038] In another embodiment, it may be beneficial or otherwise desirable to amplify the nucleic acid of the microvesicle prior to analyzing it. Methods of nucleic acid amplification are commonly used and generally known in the art, many examples of which are described herein. If desired, the amplification can be performed such that it is quantitative. Quantitative amplification will allow quantitative determination of relative amounts of the various nucleic acids, to generate a profile as described below.

[0039] In one embodiment, the extracted nucleic acid is RNA. RNAs are then preferably reverse-transcribed into complementary DNAs before further amplification. Such

reverse transcription may be performed alone or in combination with an amplification step. One example of a method combining reverse transcription and amplification steps is reverse transcription polymerase chain reaction (RT-PCR), which may be further modified to be quantitative, e.g., quantitative RT-PCR as described in US Patent No. 5,639,606, which is incorporated herein by reference for this teaching.

[0040] Nucleic acid amplification methods include, without limitation, polymerase chain reaction (PCR) (US Patent No. 5,219,727) and its variants such as in situ polymerase chain reaction (US Patent No. 5,538,871), quantitative polymerase chain reaction (US Patent No. 5,219,727), nested polymerase chain reaction (US Patent No. 5,556,773), self sustained sequence replication and its variants (Guatelli et al., 1990), transcriptional amplification system and its variants (Kwoh et al., 1989), Qb Replicase and its variants (Miele et al., 1983), cold-PCR (Li et al., 2008) or any other nucleic acid amplification methods, followed by the detection of the amplified molecules using techniques well known to those of skill in the art. Especially useful are those detection schemes designed for the detection of nucleic acid molecules if such molecules are present in very low numbers. The foregoing references are incorporated herein for their teachings of these methods.

[0041] The analysis of nucleic acids present in the microvesicles is quantitative and/or qualitative. For quantitative analysis, the amounts (expression levels), either relative or absolute, of specific nucleic acids of interest within the microvesicles are measured with methods known in the art (described below). For qualitative analysis, the species of specific nucleic acids of interest within the microvesicles, whether wild type or variants, are identified with methods known in the art (described below).

[0042] "Genetic aberrations" is used herein to refer to the nucleic acid amounts as well as nucleic acid variants within the microvesicles. Specifically, genetic aberrations include, without limitation, over-expression of a gene (e.g., oncogenes) or a panel of genes, under-expression of a gene (e.g., tumor suppressor genes such as p53 or RB) or a panel of genes, alternative production of splice variants of a gene or a panel of genes, gene copy number variants (CNV) (e.g. DNA double minutes) (Hahn, 1993), nucleic acid modifications (e.g., methylation, acetylation and phosphorylations), single nucleotide polymorphisms (SNPs), chromosomal rearrangements (e.g., inversions, deletions and duplications), and mutations (insertions, deletions, duplications, missense, nonsense, synonymous or any other nucleotide changes) of a gene or a panel of genes, which mutations, in many cases, ultimately affect the

activity and function of the gene products, lead to alternative transcriptional splicing variants and/or changes of gene expression level.

[0043] The determination of such genetic aberrations can be performed by a variety of techniques known to the skilled practitioner. For example, expression levels of nucleic acids, alternative splicing variants, chromosome rearrangement and gene copy numbers can be determined by microarray analysis (US Patent Nos. 6,913,879, 7,364,848, 7,378,245, 6,893,837 and 6,004,755) and quantitative PCR. Particularly, copy number changes may be detected with the Illumina Infinium II whole genome genotyping assay or Agilent Human Genome CGH Microarray (Steemers et al., 2006). Nucleic acid modifications can be assayed by methods described in, e.g., US Patent No. 7,186,512 and patent publication WO/2003/023065. Particularly, methylation profiles may be determined by Illumina DNA Methylation OMA003 Cancer Panel. SNPs and mutations can be detected by hybridization with allele-specific probes, enzymatic mutation detection, chemical cleavage of mismatched heteroduplex (Cotton et al., 1988), ribonuclease cleavage of mismatched bases (Myers et al., 1985), mass spectrometry (US Patent Nos. 6,994,960, 7,074,563, and 7,198,893), nucleic acid sequencing, single strand conformation polymorphism (SSCP) (Orita et al., 1989), denaturing gradient gel electrophoresis (DGGE)(Fischer and Lerman, 1979a; Fischer and Lerman, 1979b), temperature gradient gel electrophoresis (TGGE) (Fischer and Lerman, 1979a; Fischer and Lerman, 1979b), restriction fragment length polymorphisms (RFLP) (Kan and Dozy, 1978a; Kan and Dozy, 1978b), oligonucleotide ligation assay (OLA), allele-specific PCR (ASPCR) (US Patent No. 5,639,611), ligation chain reaction (LCR) and its variants (Abravaya et al., 1995; Landegren et al., 1988; Nakazawa et al., 1994), flow-cytometric heteroduplex analysis (WO/2006/113590) and combinations/modifications thereof. Notably, gene expression levels may be determined by the serial analysis of gene expression (SAGE) technique (Velculescu et al., 1995). In general, the methods for analyzing genetic aberrations are reported in numerous publications, not limited to those cited herein, and are available to skilled practitioners. The appropriate method of analysis will depend upon the specific goals of the analysis, the condition/history of the patient, and the specific cancer(s), diseases or other medical conditions to be detected, monitored or treated. The forgoing references are incorporated herein for their teachings of these methods.

[0044] A variety of genetic aberrations have been identified to occur and/or contribute to the initial generation or progression of cancer. Examples of genes which are commonly up-regulated (i.e. over-expressed) in cancer are provided in Table 4 (cancers of different

types) and Table 6 (pancreatic cancer). Examples of microRNAs which are up-regulated in brain tumor are provided in Table 8. In one embodiment of the invention, there is an increase in the nucleic acid expression level of a gene listed in Table 4 and/or Table 6 and/or of a microRNA listed in Table 8. Examples of genes which are commonly down-regulated (e.g. under-expressed) in cancer are provided in Table 5 (cancers of different types) and Table 7 (pancreatic cancer). Examples of microRNAs which are down-regulated in brain tumor are provided in Table 9. In one embodiment of the invention, there is a decrease in the nucleic acid expression level of a gene listed in Table 5 and/or Table 7 and/or a microRNA listed in Table 9. Examples of genes which are commonly under expressed, or over expressed in brain tumors are reviewed in (Furnari et al., 2007) , and this subject matter is incorporated herein by reference. With respect to the development of brain tumors, RB and p53 are often down-regulated to otherwise decrease their tumor suppressive activity. Therefore, in these embodiments, the presence or absence of an increase or decrease in the nucleic acid expression level of a gene(s) and/or a microRNA(s) whose disregulated expression level is specific to a type of cancer can be used to indicate the presence or absence of the type of cancer in the subject.

[0045] Likewise, nucleic acid variants, e.g., DNA or RNA modifications, single nucleotide polymorphisms (SNPs) and mutations (e.g., missense, nonsense, insertions, deletions, duplications) may also be analyzed within microvesicles from bodily fluid of a subject, including pregnant females where microvesicles derived from the fetus may be in serum as well as amniotic fluid. Non-limiting examples are provided in Table 3. In yet a further embodiment, the nucleotide variant is in the EGFR gene. In a still further embodiment, the nucleotide variant is the EGFRvIII mutation/variant. The terms "EGFR","epidermal growth factor receptor" and "ErbB1"are used interchangeably in the art, for example as described in a paper by Carpenter (Carpenter, 1987). With respect to the development of brain tumors, RB, PTEN, p16, p21 and p53 are often mutated to otherwise decrease their tumor suppressive activity. Examples of specific mutations in specific forms of brain tumors are discussed in a paper by Furnari et al. (Furnari et al., 2007), and this subject matter is incorporated herein by reference.

[0046] In addition, more genetic aberrations associated with cancers have been identified recently in a few ongoing research projects. For example, the Cancer Genome Atlas (TCGA) program is exploring a spectrum of genomic changes involved in human cancers. The results of this project and other similar research efforts are published and

incorporated herein by reference (Jones et al., 2008; McLendon et al., 2008; Parsons et al., 2008; Wood et al., 2007). Specifically, these research projects have identified genetic aberrations, such as mutations (e.g., missense, nonsense, insertions, deletions and duplications), gene expression level variations (mRNA or microRNA), copy number variations and nucleic acid modification (e.g. methylation), in human glioblastoma, pancreatic cancer, breast cancer and/or colorectal cancer. The genes most frequently mutated in these cancers are listed in Table 11 and Table 12 (glioblastoma), Table 13 (pancreatic cancer), Table 14 (breast cancer) and Table 15 (colorectal cancer). The genetic aberrations in these genes, and in fact any genes which contain any genetic aberrations in a cancer, are targets that may be selected for use in diagnosing and/or monitoring cancer by the methods described herein.

[0047] Detection of one or more nucleotide variants can be accomplished by performing a nucleotide variant screen on the nucleic acids within the microvesicles. Such a screen can be as wide or narrow as determined necessary or desirable by the skilled practitioner. It can be a wide screen (set up to detect all possible nucleotide variants in genes known to be associated with one or more cancers or disease states). Where one specific cancer or disease is suspected or known to exist, the screen can be specific to that cancer or disease. One example is a brain tumor/brain cancer screen (e.g., set up to detect all possible nucleotide variants in genes associated with various clinically distinct subtypes of brain cancer or known drug-resistant or drug-sensitive mutations of that cancer).

[0048] In one embodiment, the analysis is of a profile of the amounts (levels) of specific nucleic acids present in the microvesicle, herein referred to as a "quantitative nucleic acid profile" of the microvesicles. In another embodiment, the analysis is of a profile of the species of specific nucleic acids present in the microvesicles (both wild type as well as variants), herein referred to as a "nucleic acid species profile." A term used herein to refer to a combination of these types of profiles is "genetic profile" which refers to the determination of the presence or absence of nucleotide species, variants and also increases or decreases in nucleic acid levels.

[0049] Once generated, these genetic profiles of the microvesicles are compared to those expected in, or otherwise derived from a healthy normal individual. A profile can be a genome wide profile (set up to detect all possible expressed genes or DNA sequences). It can be narrower as well, such as a cancer wide profile (set up to detect all possible genes or nucleic acids derived therefrom, or known to be associated with one or more cancers).

Where one specific cancer is suspected or known to exist, the profile can be specific to that cancer (e.g., set up to detect all possible genes or nucleic acids derived therefrom, associated with various clinically distinct subtypes of that cancer or known drug-resistant or sensitive mutations of that cancer).

[0050] Which nucleic acids are to be amplified and/or analyzed can be selected by the skilled practitioner. The entire nucleic acid content of the exosomes or only a subset of specific nucleic acids which are likely or suspected of being influenced by the presence of a disease or other medical condition such as cancer, can be amplified and/or analyzed. The identification of a nucleic acid aberration(s) in the analyzed microvesicle nucleic acid can be used to diagnose the subject for the presence of a disease such as cancer, hereditary diseases or viral infection with which that aberration(s) is associated. For instance, analysis for the presence or absence of one or more nucleic acid variants of a gene specific to a cancer (e.g. the EGFRvIII mutation) can indicate the cancer's presence in the individual. Alternatively, or in addition, analysis of nucleic acids for an increase or decrease in nucleic acid levels specific to a cancer can indicate the presence of the cancer in the individual (e.g., a relative increase in EGFR nucleic acid, or a relative decrease in a tumor suppressor gene such as p53).

[0051] In one embodiment, mutations of a gene which is associated with a disease such as cancer (e.g. via nucleotide variants, over-expression or under-expression) are detected by analysis of nucleic acids in microvesicles, which nucleic acids are derived from the genome itself in the cell of origin or exogenous genes introduced through viruses. The nucleic acid sequences may be complete or partial, as both are expected to yield useful information in diagnosis and prognosis of a disease. The sequences may be sense or anti-sense to the actual gene or transcribed sequences. The skilled practitioner will be able to devise detection methods for a nucleotide variance from either the sense or anti-sense nucleic acids which may be present in a microvesicle. Many such methods involve the use of probes which are specific for the nucleotide sequences which directly flank, or contain the nucleotide variances. Such probes can be designed by the skilled practitioner given the knowledge of the gene sequences and the location of the nucleic acid variants within the gene. Such probes can be used to isolate, amplify, and/or actually hybridize to detect the nucleic acid variants, as described in the art and herein.

[0052] Determining the presence or absence of a particular nucleotide variant or plurality of variants in the nucleic acid within microvesicles from a subject can be performed

in a variety of ways. A variety of methods are available for such analysis, including, but not limited to, PCR, hybridization with allele-specific probes, enzymatic mutation detection, chemical cleavage of mismatches, mass spectrometry or DNA sequencing, including minisequencing. In particular embodiments, hybridization with allele specific probes can be conducted in two formats: 1) allele specific oligonucleotides bound to a solid phase (glass, silicon, nylon membranes) and the labeled sample in solution, as in many DNA chip applications, or 2) bound sample (often cloned DNA or PCR amplified DNA) and labeled oligonucleotides in solution (either allele specific or short so as to allow sequencing by hybridization). Diagnostic tests may involve a panel of variances, often on a solid support, which enables the simultaneous determination of more than one variance. In another embodiment, determining the presence of at least one nucleic acid variance in the microvesicle nucleic acid entails a haplotyping test. Methods of determining haplotypes are known to those of skill in the art, as for example, in WO 00/04194.

[0053] In one embodiment, the determination of the presence or absence of a nucleic acid variant(s) involves determining the sequence of the variant site or sites (the exact location within the sequence where the nucleic acid variation from the norm occurs) by methods such as polymerase chain reaction (PCR), chain terminating DNA sequencing (US Patent No. 5547859), minisequencing (Fiorentino et al., 2003), oligonucleotide hybridization, pyrosequencing, Illumina genome analyzer, deep sequencing, mass spectrometry or other nucleic acid sequence detection methods. Methods for detecting nucleic acid variants are well known in the art and disclosed in WO 00/04194, incorporated herein by reference. In an exemplary method, the diagnostic test comprises amplifying a segment of DNA or RNA (generally after converting the RNA to complementary DNA) spanning one or more known variants in the desired gene sequence. This amplified segment is then sequenced and/or subjected to electrophoresis in order to identify nucleotide variants in the amplified segment.

[0054] In one embodiment, the invention provides a method of screening for nucleotide variants in the nucleic acid of microvesicles isolated as described herein. This can be achieved, for example, by PCR or, alternatively, in a ligation chain reaction (LCR) (Abravaya et al., 1995; Landegren et al., 1988; Nakazawa et al., 1994). LCR can be particularly useful for detecting point mutations in a gene of interest (Abravaya et al., 1995). The LCR method comprises the steps of designing degenerate primers for amplifying the target sequence, the primers corresponding to one or more conserved regions of the nucleic acid corresponding to the gene of interest, amplifying PCR products with the primers using, as a template, a nucleic

acid obtained from a microvesicle, and analyzing the PCR products. Comparison of the PCR products of the microvesicle nucleic acid to a control sample (either having the nucleotide variant or not) indicates variants in the microvesicle nucleic acid. The change can be either an absence or presence of a nucleotide variant in the microvesicle nucleic acid, depending upon the control.

[0055] Analysis of amplification products can be performed using any method capable of separating the amplification products according to their size, including automated and manual gel electrophoresis, mass spectrometry, and the like.

[0056] Alternatively, the amplification products can be analyzed based on sequence differences, using SSCP, DGGE, TGGE, chemical cleavage, OLA, restriction fragment length polymorphisms as well as hybridization, for example, nucleic acid microarrays.

[0057] The methods of nucleic acid isolation, amplification and analysis are routine for one skilled in the art and examples of protocols can be found, for example, in *Molecular Cloning: A Laboratory Manual* (3-Volume Set) Ed. Joseph Sambrook, David W. Russel, and Joe Sambrook, Cold Spring Harbor Laboratory, 3rd edition (January 15, 2001), ISBN: 0879695773. A particular useful protocol source for methods used in PCR amplification is *PCR Basics: From Background to Bench* by Springer Verlag; 1st edition (October 15, 2000), ISBN: 0387916008.

[0058] Many methods of diagnosis performed on a tumor biopsy sample can be performed with microvesicles since tumor cells, as well as some normal cells are known to shed microvesicles into bodily fluid and the genetic aberrations within these microvesicles reflect those within tumor cells as demonstrated herein. Furthermore, methods of diagnosis using microvesicles have characteristics that are absent in methods of diagnosis performed directly on a tumor biopsy sample. For example, one particular advantage of the analysis of microvesicular nucleic acids, as opposed to other forms of sampling of tumor/cancer nucleic acid, is the availability for analysis of tumor/cancer nucleic acids derived from all foci of a tumor or genetically heterogeneous tumors present in an individual. Biopsy samples are limited in that they provide information only about the specific focus of the tumor from which the biopsy is obtained. Different tumorous/cancerous foci found within the body, or even within a single tumor often have different genetic profiles and are not analyzed in a standard biopsy. However, analysis of the microvesicular nucleic acids from an individual presumably provides a sampling of all foci within an individual. This provides valuable

information with respect to recommended treatments, treatment effectiveness, disease prognosis, and analysis of disease recurrence, which cannot be provided by a simple biopsy.

[0059] Identification of genetic aberrations associated with specific diseases and/or medical conditions by the methods described herein can also be used for prognosis and treatment decisions of an individual diagnosed with a disease or other medical condition such as cancer. Identification of the genetic basis of a disease and/or medical condition provides useful information guiding the treatment of the disease and/or medical condition. For example, many forms of chemotherapy have been shown to be more effective on cancers with specific genetic abnormalities/aberrations. One example is the effectiveness of EGFR-targeting treatments with medicines, such as the kinase inhibitors gefitinib and erlotinib. Such treatment have been shown to be more effective on cancer cells whose EGFR gene harbors specific nucleotide mutations in the kinase domain of EGFR protein (U.S. Patent publication 20060147959). In other words, the presence of at least one of the identified nucleotide variants in the kinase domain of EGFR nucleic acid message indicates that a patient will likely benefit from treatment with the EGFR-targeting compound gefitinib or erlotinib. Such nucleotide variants can be identified in nucleic acids present in microvesicles by the methods described herein, as it has been demonstrated that EGFR transcripts of tumor origin are isolated from microvesicles in bodily fluid.

[0060] Genetic aberrations in other genes have also been found to influence the effectiveness of treatments. As disclosed in the publication by Furnari et al. (Furnari et al., 2007), mutations in a variety of genes affect the effectiveness of specific medicines used in chemotherapy for treating brain tumors. The identification of these genetic aberrations in the nucleic acids within microvesicles will guide the selection of proper treatment plans.

[0061] As such, aspects of the present invention relate to a method for monitoring disease (e.g. cancer) progression in a subject, and also to a method for monitoring disease recurrence in an individual. These methods comprise the steps of isolating microvesicles from a bodily fluid of an individual, as discussed herein, and analyzing nucleic acid within the microvesicles as discussed herein (e.g. to create a genetic profile of the microvesicles). The presence/absence of a certain genetic aberration/profile is used to indicate the presence/absence of the disease (e.g. cancer) in the subject as discussed herein. The process is performed periodically over time, and the results reviewed, to monitor the progression or regression of the disease, or to determine recurrence of the disease. Put another way, a change in the genetic profile indicates a change in the disease state in the subject. The period

of time to elapse between sampling of microvesicles from the subject, for performance of the isolation and analysis of the microvesicle, will depend upon the circumstances of the subject, and is to be determined by the skilled practitioner. Such a method would prove extremely beneficial when analyzing a nucleic acid from a gene that is associated with the therapy undergone by the subject. For example, a gene which is targeted by the therapy can be monitored for the development of mutations which make it resistant to the therapy, upon which time the therapy can be modified accordingly. The monitored gene may also be one which indicates specific responsiveness to a specific therapy.

[0062] Aspects of the present invention also relate to the fact that a variety of non-cancer diseases and/or medical conditions also have genetic links and/or causes, and such diseases and/or medical conditions can likewise be diagnosed and/or monitored by the methods described herein. Many such diseases are metabolic, infectious or degenerative in nature. One such disease is diabetes (e.g. diabetes insipidus) in which the vasopressin type 2 receptor (V2R) is modified. Another such disease is kidney fibrosis in which the genetic profiles for the genes of collagens, fibronectin and TGF-β are changed. Changes in the genetic profile due to substance abuse (e.g. a steroid or drug use), viral and/or bacterial infection, and hereditary disease states can likewise be detected by the methods described herein.

[0063] Diseases or other medical conditions for which the inventions described herein are applicable include, but are not limited to, nephropathy, diabetes insipidus, diabetes type I, diabetes II, renal disease glomerulonephritis, bacterial or viral glomerulonephritides, IgA nephropathy, Henoch-Schonlein Purpura, membranoproliferative glomerulonephritis, membranous nephropathy, Sjogren's syndrome, nephrotic syndrome minimal change disease, focal glomerulosclerosis and related disorders, acute renal failure, acute tubulointerstitial nephritis, pyelonephritis, GU tract inflammatory disease, Pre-clampsia, renal graft rejection, leprosy, reflux nephropathy, nephrolithiasis, genetic renal disease, medullary cystic, medullar sponge, polycystic kidney disease, autosomal dominant polycystic kidney disease, autosomal recessive polycystic kidney disease, tuberous sclerosis, von Hippel-Lindau disease, familial thin-glomerular basement membrane disease, collagen III glomerulopathy, fibronectin glomerulopathy, Alport's syndrome, Fabry's disease, Nail-Patella Syndrome, congenital urologic anomalies, monoclonal gammopathies, multiple myeloma, amyloidosis and related disorders, febrile illness, familial Mediterranean fever, HIV infection-AIDS, inflammatory disease, systemic vasculitides, polyarteritis nodosa, Wegener's granulomatosis, polyarteritis,

necrotizing and crecentic glomerulonephritis, polymyositis-dermatomyositis, pancreatitis, rheumatoid arthritis, systemic lupus erythematosus, gout, blood disorders, sickle cell disease, thrombotic thrombocytopenia purpura, Fanconi's syndrome, transplantation, acute kidney injury, irritable bowel syndrome, hemolytic-uremic syndrome, acute corticol necrosis, renal thromboembolism, trauma and surgery, extensive injury, burns, abdominal and vascular surgery, induction of anesthesia, side effect of use of drugs or drug abuse, circulatory disease myocardial infarction, cardiac failure, peripheral vascular disease, hypertension, coronary heart disease, non-atherosclerotic cardiovascular disease, atherosclerotic cardiovascular disease, skin disease, soriasis, systemic sclerosis, respiratory disease, COPD, obstructive sleep apnoea, hypoia at high altitude or erdocrine disease, acromegaly, diabetes mellitus, or diabetes insipidus.

[0064] Selection of an individual from whom the microvesicles are isolated is performed by the skilled practitioner based upon analysis of one or more of a variety of factors. Such factors for consideration are whether the subject has a family history of a specific disease (e.g. a cancer), has a genetic predisposition for such a disease, has an increased risk for such a disease due to family history, genetic predisposition, other disease or physical symptoms which indicate a predisposition, or environmental reasons. Environmental reasons include lifestyle, exposure to agents which cause or contribute to the disease such as in the air, land, water or diet. In addition, having previously had the disease, being currently diagnosed with the disease prior to therapy or after therapy, being currently treated for the disease (undergoing therapy), being in remission or recovery from the disease, are other reasons to select an individual for performing the methods.

[0065] The methods described herein are optionally performed with the additional step of selecting a gene or nucleic acid for analysis, prior to the analysis step. This selection can be based on any predispositions of the subject, or any previous exposures or diagnosis, or therapeutic treatments experienced or concurrently undergone by the subject.

[0066] The cancer diagnosed, monitored or otherwise profiled, can be any kind of cancer. This includes, without limitation, epithelial cell cancers such as lung, ovarian, cervical, endometrial, breast, brain, colon and prostate cancers. Also included are gastrointestinal cancer, head and neck cancer, non-small cell lung cancer, cancer of the nervous system, kidney cancer, retina cancer, skin cancer, liver cancer, pancreatic cancer, genital-urinary cancer and bladder cancer, melanoma, and leukemia. In addition, the methods and compositions of the present invention are equally applicable to detection, diagnosis and

prognosis of non-malignant tumors in an individual (e.g. neurofibromas, meningiomas and schwannomas).

[0067] In one embodiment, the cancer is brain cancer. Types of brain tumors and cancer are well known in the art. Glioma is a general name for tumors that arise from the glial (supportive) tissue of the brain. Gliomas are the most common primary brain tumors. Astrocytomas, ependymomas, oligodendrogliomas, and tumors with mixtures of two or more cell types, called mixed gliomas, are the most common gliomas. The following are other common types of brain tumors: Acoustic Neuroma (Neurilemmoma, Schwannoma. Neurinoma), Adenoma, Astracytoma, Low-Grade Astrocytoma, giant cell astrocytomas, Mid- and High-Grade Astrocytoma, Recurrent tumors, Brain Stem Glioma, Chordoma, Choroid Plexus Papilloma, CNS Lymphoma (Primary Malignant Lymphoma), Cysts, Dermoid cysts, Epidermoid cysts, Craniopharyngioma, Ependymoma Anaplastic ependymoma, Gangliocytoma (Ganglioneuroma), Ganglioglioma, Glioblastoma Multiforme (GBM), Malignant Astracytoma, Glioma, Hemangioblastoma, Inoperable Brain Tumors, Lymphoma, Medulloblastoma (MDL), Meningioma, Metastatic Brain Tumors, Mixed Glioma, Neurofibromatosis, Oligodendroglioma. Optic Nerve Glioma, Pineal Region Tumors, Pituitary Adenoma, PNET (Primitive Neuroectodermal Tumor), Spinal Tumors, Subependymoma, and Tuberous Sclerosis (Bourneville's Disease).

[0068] In addition to identifying previously known nucleic acid aberrations (as associated with diseases), the methods of the present invention can be used to identify previously unidentified nucleic acid sequences/modifications (e.g. post transcriptional modifications) whose aberrations are associated with a certain disease and/or medical condition. This is accomplished, for example, by analysis of the nucleic acid within microvesicles from a bodily fluid of one or more subjects with a given disease/medical condition (e.g. a clinical type or subtype of cancer) and comparison to the nucleic acid within microvesicles of one or more subjects without the given disease/medical condition, to identify differences in their nucleic acid content. The differences may be any genetic aberrations including, without limitation, expression level of the nucleic acid, alternative splice variants, gene copy number variants (CNV), modifications of the nucleic acid , single nucleotide polymorphisms (SNPs), and mutations (insertions, deletions or single nucleotide changes) of the nucleic acid. Once a difference in a genetic parameter of a particular nucleic acid is identified for a certain disease, further studies involving a clinically and statistically significant number of subjects may be carried out to establish the correlation between the

genetic aberration of the particular nucleic acid and the disease. The analysis of genetic aberrations can be done by one or more methods described herein, as determined appropriate by the skilled practitioner.

Exosomes As Delivery Vehicles

[0069] Aspects of the present invention also relate to the actual microvesicles described herein. In one embodiment, the invention is an isolated microvesicle as described herein, isolated from an individual. In one embodiment, the microvesicle is produced by a cell within the brain of the individual (e.g. a tumor or non-tumor cell). In another embodiment, the microvesicle is isolated from a bodily fluid of an individual, as described herein. Methods of isolation are described herein.

[0070] Another aspect of the invention relates to the finding that isolated microvesicles from human glioblastoma cells contain mRNAs, miRNAs and angiogenic proteins. Such glioblastoma microvesicles were taken up by primary human brain endothelial cells, likely via an endocytotic mechanism, and a reporter protein mRNA incorporated into the microvesicles was translated in those cells. This indicates that messages delivered by microvesicles can change the genetic and/or translational profile of a target cell (a cell which takes up a microvesicle). The microvesicles also contained miRNAs which are known to be abundant in glioblastomas (Krichevsky et al, manuscript in preparation). Thus microvesicles derived from glioblastoma tumors function as delivery vehicles for mRNA, miRNA and proteins which can change the translational state of other cells via delivery of specific mRNA species, promote angiogenesis of endothelial cells, and stimulate tumor growth.

[0071] In one embodiment, microvesicles are depleted from a bodily fluid from a donor subject before said bodily fluid is delivered to a recipient subject. The donor subject may be a subject with an undetectable tumor and the microvesicles in the bodily fluid are derived from the tumor. The tumor microvesicles in the donor bodily fluid, if unremoved, would be harmful since the genetic materials and proteins in the microvesicle may promote unrestricted growth of cells in the recipient subject.

[0072] As such, another aspect of the invention is the use of the microvesicles identified herein to deliver a nucleic acid to a cell. In one embodiment, the cell is within the body of an individual. The method comprises administering a microvesicle(s) which contains the nucleic acid, or a cell that produces such microvesicles, to the individual such that the

microvesicles contacts and/or enters the cell of the individual. The cell to which the nucleic acid gets delivered is referred to as the target cell.

[0073] The microvesicle can be engineered to contain a nucleic acid that it would not naturally contain (i.e. which is exogenous to the normal content of the microvesicle). This can be accomplished by physically inserting the nucleic acid into the microvesicles. Alternatively, a cell (e.g. grown in culture) can be engineered to target one or more specific nucleic acid into the exosome, and the exosome can be isolated from the cell. Alternatively, the engineered cell itself can be administered to the individual.

[0074] In one embodiment, the cell which produces the exosome for administration is of the same or similar origin or location in the body as the target cell. That is to say, for delivery of a microvesicle to a brain cell, the cell which produces the microvesicle would be a brain cell (e.g. a primary cell grown in culture). In another embodiment, the cell which produces the exosome is of a different cell type than the target cell. In one embodiment, the cell which produces the exosome is a type that is located proximally in the body to the target cell.

[0075] A nucleic acid sequence which can be delivered to a cell via an exosome can be RNA or DNA, and can be single or double stranded, and can be selected from a group comprising: nucleic acid encoding a protein of interest, oligonucleotides, nucleic acid analogues, for example peptide-nucleic acid (PNA), pseudo-complementary PNA (pc-PNA), locked nucleic acid (LNA) etc. Such nucleic acid sequences include, for example, but are not limited to, nucleic acid sequences encoding proteins, for example that act as transcriptional repressors, antisense molecules, ribozymes, small inhibitory nucleic acid sequences, for example but are not limited to RNAi, shRNA, siRNA, miRNA, antisense oligonucleotides, and combinations thereof.

[0076] Microvesicles isolated from a cell type are delivered to a recipient subject. Said microvesicles may benefit the recipient subject medically. For example, the angiogenesis and pro-proliferation effects of tumor exosomes may help the regeneration of injured tissues in the recipient subject. In one embodiment, the delivery means is by bodily fluid transfusion wherein microvesicles are added into a bodily fluid from a donor subject before said bodily fluid is delivered to a recipient subject.

[0077] In another embodiment, the microvesicle is an ingredient (e.g. the active ingredient in a pharmaceutically acceptable formulation suitable for administration to the

subject (e.g. in the methods described herein). Generally this comprises a pharmaceutically acceptable carrier for the active ingredient. The specific carrier will depend upon a number of factors (e.g.. the route of administration).

[0078] The "pharmaceutically acceptable carrier" means any pharmaceutically acceptable means to mix and/or deliver the targeted delivery composition to a subject. This includes a pharmaceutically acceptable material, composition or vehicle, such as a liquid or solid filler, diluent, excipient, solvent or encapsulating material, involved in carrying or transporting the subject agents from one organ, or portion of the body, to another organ, or portion of the body. Each carrier must be "acceptable" in the sense of being compatible with the other ingredients of the formulation and is compatible with administration to a subject, for example a human.

[0079] Administration to the subject can be either systemic or localized. This includes, without limitation, dispensing, delivering or applying an active compound (e.g. in a pharmaceutical formulation) to the subject by any suitable route for delivery of the active compound to the desired location in the subject, including delivery by either the parenteral or oral route, intramuscular injection, subcutaneous/intradermal injection, intravenous injection, buccal administration, transdermal delivery and administration by the rectal, colonic, vaginal, intranasal or respiratory tract route.

[0080] It should be understood that this invention is not limited to the particular methodologies, protocols and reagents, described herein and as such may vary. The terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention, which is defined solely by the claims.

[0081] In one respect, the present invention relates to the herein described compositions, methods, and respective components thereof, as essential to the invention, yet open to the inclusion of unspecified elements, essential or not ("comprising"). In some embodiments, other elements to be included in the description of the composition, method or respective component thereof are limited to those that do not materially affect the basic and novel characteristic(s) of the invention ("consisting essentially of"). This applies equally to steps within a described method as well as compositions and components therein. In other embodiments, the inventions, compositions, methods, and respective components thereof, described herein are intended to be exclusive of any element not deemed an essential element to the component, composition or method ("consisting of").

EXAMPLES

Examples 1-7. Tumor cells shed microvesicles, which contain RNAs, including mRNAs and microRNAs, and that microvesicles contain more than 90% of the extracellular RNA in bodily fluids.

Example 1: Microvesicles are shed from primary human glioblastoma cells.

[0082] Glioblastoma tissue was obtained from surgical resections and tumor cells were dissociated and cultured as monolayers. Specifically, brain tumor specimens from patients diagnosed by a neuropathologist as glioblastoma multiforme were taken directly from surgery and placed in cold sterile Neurobasal media (Invitrogen, Carlsbad, CA, USA). The specimens were dissociated into single cells within 1 hr from the time of surgery using a Neural Tissue Dissociation Kit (Miltenyi Biotech, Berisch Gladbach, Germany) and plated in DMEM 5% dFBS supplemented with penicillin-streptomycin (10 IU ml^{-1} and 10 $\mu g\ ml^{-1}$, respectively, Sigma-Aldrich, St Louis, MO, USA). Because microvesicles can be found in the fetal bovine serum (FBS) traditionally used to cultivate cells, and these microvesicles contain substantial amounts of mRNA and miRNA, it was important to grow the tumor cells in media containing microvesicle-depleted FBS (dFBS). Cultured primary cells obtained from three glioblastoma tumors were found to produce microvesicles at both early and later passages (a passage is a cellular generation defined by the splitting of cells, which is a common cell culture technique and is necessary to keep the cells alive). The microvesicles were able to be detected by scanning electronmicroscopy (FIGS 1a and 1b) and transmission electronmicroscopy (FIG 1f). Briefly, human glioblastoma cells were placed on ornithine-coated cover-slips, fixed in 0.5x Karnovskys fixative and then washed 2x5min (2 times with 5 min each) with PBS. The cells were dehydrated in 35% EtOH 10 min, 50% EtOH 2x 10 min, 70% EtOH 2x 10 min, 95% EtOH 2x 10 min, and 100% EtOH 4 x 10 min. The cells were then transferred to critical point drying in a Tousimis SAMDR1-795 semi-automatic Critical Point Dryer followed by coating with chromium in a GATAN Model 681 High Resolution Ion Beam Coater. As shown in FIGS. 1a and 1b, tumor cells were covered with microvesicles varying in size from about 50 - 500 nm.

Example 2: Glioblastoma microvesicles contain RNA.

[0083] To isolate microvesicles, glioblastoma cells at passage 1-15 were cultured in microvesicle-free media (DMEM containing 5% dFBS prepared by ultracentrifugation at 110,000 x g for 16 hours to remove bovine microvesicles). The conditioned medium from 40

million cells was harvested after 48 hours. The microvesicles were purified by differential centrifugation. Specifically, glioblastoma conditioned medium was centrifuged for 10 min at 300 x g to eliminate any cell contamination. Supernatants were further centrifuged for 20 min at 16,500 x g and filtered through a 0.22 μm filter. Microvesicles were then pelleted by ultracentrifugation at 110,000 x g for 70 min. The microvesicle pellets were washed in 13 ml PBS, pelleted again and resuspended in PBS.

[0084] Isolated microvesicles were measured for their total protein content using DC Protein Assay (Bio-Rad, Hercules, CA, USA).

[0085] For the extraction of RNA from microvesicles, RNase A (Fermentas, Glen Burnie, MD, USA) at a final concentration of 100 μg/ml was added to suspensions of microvesicles and incubated for 15 min at 37°C to get rid of RNA outside of the microvesicles and thus ensure that the extracted RNA would come from inside the microvesicles. Total RNA was then extracted from the microvesicles using the MirVana RNA isolation kit (Ambion, Austin TX, USA) according to the manufacturer's protocol. After treatment with DNAse according to the manufacturer's protocol, the total RNA was quantified using a nanodrop ND-1000 instrument (Thermo Fischer Scientific, Wilmington, DE, USA).

[0086] Glioblastoma microvesicles were found to contain RNA and protein in a ratio of approximately 1:80 (μg RNA:μg protein). The average yield of proteins and RNAs isolated from microvesicles over a 48-hour period in culture was around 4 μg protein and 50 ng RNA/million cells.

[0087] To confirm that the RNA was contained inside the microvesicles, microvesicles were either exposed to RNase A or mock treatment before RNA extraction (FIG. 1c). There was never more than a 7% decrease in RNA content following RNase treatment. Thus, it appears that almost all of the extracellular RNA from the media is contained within the microvesicles and is thereby protected from external RNases by the surrounding vesicular membrane.

[0088] Total RNA from microvesicles and their donor cells were analyzed with a Bioanalyzer, showing that the microvesicles contain a broad range of RNA sizes consistent with a variety of mRNAs and miRNAs, but lack 18S and 28S the ribosomal RNA peaks characteristic of cellular RNA (FIGS. 1d and 1e).

Example 3: Microvesicles contain DNA.

[0089] To test if microvesicles also contain DNA, exosomes were isolated as mentioned in Example 2 and then treated with DNase before being lysed to release contents. The DNase treatment step was to remove DNA outside of the exosomes so that only DNA residing inside the exosomes was extracted. Specifically, the DNase treatment was performed using the DNA-free kit from Ambion according to manufacturer's recommendations (Catalog#AM1906). For the DNA purification step, an aliquot of isolated exosomes was lysed in 300µl lysis buffer that was part of the MirVana RNA isolation kit (Ambion) and the DNAs were purified from the lysed mixture using a DNA purification kit (Qiagen) according to the manufacturer's recommendation.

[0090] To examine whether the extracted DNA contains common genes, PCRs were performed using primer pairs specific to GAPDH, Human endogenous retrovirus K, Tenascin-c and Line-1. For the GAPDH gene, the following primers were used: Forw3GAPDHnew (SEQ ID NO: 1) and Rev3GAPDHnew (SEQ ID NO: 2). The primer pair amplifies a 112bp amplicon if the template is a spliced GAPDH cDNA and a 216bp amplicon if the template is an un-spliced genomic GAPDH DNA. In one experiment, isolated exosomes were treated with DNase before being lysed for DNA extraction (FIG. 3a). The 112bp fragments were amplified as expected from the exosomes from the tumor serum (See Lane 4 in FIG. 3a) and the primary tumor cells (See Lane 6 in FIG. 3a) but not from the exosomes from normal human fibroblasts (See Lane 5 in FIG. 3a). The 216bp fragment could not be amplified from exosomes of all three origins. However, fragments of both 112bp and 216bp were amplified when the genomic DNA isolated from the glioblastoma cell was used as templates (See Lane 3 in FIG. 3a). Thus, spliced GAPDH DNA exists within exosomes isolated from tumor cells but not within exosomes isolated from normal fibroblast cells.

[0091] In contrast, in another experiment, isolated exosomes were not treated with DNase before being lysed for DNA extraction (FIG. 3b). Not only the 112bp fragments but also the 216bp fragments were amplified from exosomes isolated from primary melanoma cells (See Lane 3 in FIG. 3b), suggesting that non-spliced GAPDH DNA or partially spliced cDNA that has been reverse transcribed exists outside of the exosomes.

[0092] For the Human Endogenous Retrovirus K (HERV-K) gene, the following primers were used: HERVK_6Forw (SEQ ID NO: 3) and HERVK_6Rev (SEQ ID NO: 4). The primer pair amplifies a 172bp amplicon. DNA was extracted from exosomes that were isolated and treated with DNase, and used as the template for PCR amplification. As shown

in FIG. 3c, 172bp fragments were amplified in all tumor and normal human serum exosomes but not in exosomes from normal human fibroblasts. These data suggest that unlike exosomes from normal human fibroblasts, tumor and normal human serum exosomes contain endogenous retrovirus DNA sequences. To examine if tumor exosomes also contain transposable elements, the following LINE-1 specific primers were used for PCR amplifications: Line1_Forw (SEQ ID NO: 5) and Line1_Rev (SEQ ID NO: 6). These two primers are designed to detect LINE-1 in all species since each primer contains equal amounts of two different oligos. For the Line1_Forw primer, one oligo contains a C and the other oligo contains a G at the position designated with "s". For the Line1_Rev primer, one oligo contains an A and the other oligo contains a G at the position designated with "r". The primer pair amplifies a 290bp amplicon. The template was the DNA extracted from exosomes that were treated with DNase (as described above). As shown in FIG. 3e, 290bp LINE-1 fragments could be amplified from the exosomes from tumor cells and normal human serum but not from exosomes from the normal human fibroblasts.

[0093] To test if exosomes also contain Tenascin-C DNA, the following primer pair was used to perform PCR: Tenascin C Forw (SEQ ID NO: 7) and Tenascin C Rev (SEQ ID NO: 8). The primer pair amplifies a 197bp amplicon. The template was the DNA extracted from exosomes that were isolated and then treated with DNase before lysis. As shown in FIG. 3d, 197bp Tenascin C fragments were amplified in exosomes from tumor cells or normal human serum but not in exosomes from normal human fibroblasts. Thus, Tenascin-C DNA exists in tumor and normal human serum exosomes but not in exosomes from normal human fibroblasts.

[0094] To further confirm the presence of DNA in exosomes, exosomal DNA was extracted from D425 medulloblastoma cells using the method described above. Specifically, the exosomes were isolated and treated with DNase before lysis. Equal volumes of the final DNA extract were either treated with DNase or not treated with DNase before being visualized by Ethidium Bromide staining in 1% agarose gel. Ethidium Bromide is a dye that specifically stains nucleic acids and can be visualized under ultraviolet light. As shown in FIG. 3f, Ethidium Bromide staining disappeared after DNase treatment (See Lane 3 in FIG. 3f) while strong staining could be visualized in the un-treated aliquot (See Lane 2 in FIG. 3f). The DNase treated and non-treated extracts were also analyzed on a RNA pico chip (Agilent Technologies). As shown in FIG. 3g, single stranded DNA could be readily detected in the

DNase-non-treated extract (See upper panel in FIG. 3g) but could barely be detected in the DNase-treated extract (See lower panel in FIG. 3g).

[0095] To test whether the extracted DNA was single-stranded, nucleic acids were extracted from the treated exosomes as described in the previous paragraph and further treated with RNAse to eliminate any RNA contamination. The treated nucleic acids were then analyzed on a RNA pico Bioanalyzer chip and in a DNA 1000 chip. The RNA pico chip only detects single stranded nucleic acids. The DNA 1000 chip detected double stranded nucleic acids. As shown in FIG. 3h, single stranded nucleic acids were detected (See upper panel) but double stranded nucleic acids were not detected (See lower panel). Thus, the DNA contained within tumor exosomes are mostly single stranded.

[0096] To demonstrate that single stranded DNA exists in tumor cells but not in normal human fibroblasts, nucleic acids were extracted from exosomes from either glioblastoma patient serum or normal human fibroblasts. The exosomes were treated with DNase before lysis and the purified nucleic acids were treated with RNase before analysis. As shown in FIG. 3i, exosomal nucleic acids extracted from glioblastoma patient serum could be detected by a RNA pico chip. In contrast, only a very small amount of single stranded DNA was extracted from normal human fibroblasts.

[0097] Accordingly, exosomes from tumor cells and normal human serum were found to contain contain single-stranded DNA. The single-stranded DNA is a reverse transcription product since the amplification products do not contain introns (FIG. 3a and FIG. 3b). It is known that tumor cells as well as normal progenitor cells/stem cells have active reverse transcriptase (RT) activity although the activity in normal progenitor cells/stem cells is relatively much lower. This RT activity makes it plausible that RNA transcripts in the cell can be reverse transcribed and packaged into exosomes as cDNA. Interestingly, exosomes from tumor cells contain more cDNAs corresponding to tumor-specific gene transcripts since tumor cells usually have up-regulated reverse transcriptase activity. Therefore, tumor specific cDNA in exosomes may be used as biomarkers for the diagnosis or prognosis of different tumor types. The use of cDNAs as biomarkers would skip the step of reverse transcription compared to the used of mRNA as biomarkers for tumors. In addition, the use of exosomal cDNA is advantageous over the use of whole serum/plasma DNA because serum/plasma contains genomic DNA released from dying cells. When testing amplified whole serum/plasma DNA, there will be more background.

Example 4: Most extracellular RNA in human serum is contained within exosomes.

[0098] To determine the amount of RNA circulating in serum as "free RNA"/RNA-protein complex versus the amount of RNA contained within the exosomes, we isolated serum from a healthy human subject, and evenly split the serum into two samples with equal volume. For sample 1, the serum was ultracentrifuged to remove most microvesicles. Then the serum supernatant was collected and RNA left in the supernatant was extracted using Trizol LS. For sample 2, the serum was not ultracentrifuged and total RNA was extracted from the serum using Trizol LS. The amount of RNA in the sample 1 supernatant and sample 2 serum was measured. As a result, it was found that the amount of free RNA in sample 1 supernatant was less than 10% of the amount of total RNA isolated from the serum sample 2. Therefore, a majority of the RNA in serum is associated with the exosomes.

Example 5: High efficiency of serum extracellular nucleic acid extraction is achieved by incorporating a serum exosome isolation step.

[0099] Whole serum and plasma contain large amounts of circulating DNA and possibly also RNA protected in protein complexes, while free RNA have a half-life of a few minutes in serum. Extracellular nucleic acid profiles in serum vary between normal and diseased mammals and thus may be biomarkers for certain diseases. To examine the profiles, nucleic acids need to be extracted. However, direct extraction of nucleic acids from serum and plasma is not practical, especially from large serum/plasma volumes. In this case, large volumes of Trizol LS (a RNA extraction reagent) are used to instantly inactivate all serum nucleases before extracting the exosomal nucleic acids. Subsequently, contaminants precipitate into the sample and affect subsequent analyses. As shown in Example 4, most extracellular RNAs in serum are contained in serum exosomes. Therefore, we tested whether it is more efficient to isolate extracellular nucleic acids by isolating the serum exosomes before nucleic acid extraction.

[00100] Four milliliter (ml) blood serum from a patient was split into 2 aliquots of 2 ml each. Serum exosomes from one aliquot were isolated prior to RNA extraction. The methods of exosome isolation and RNA extraction are the same as mentioned in Example 2. For the other aliquot, RNA was extracted directly using Trizol LS according to manufacturer's recommendation. The nucleic acids from these two extractions were analyzed on a Bioanalyzer RNA chip (Agilent Technologies). As shown in Figure 4, the amount of RNA extracted with the former method is significantly more than that obtained from the latter

method. Further, the quality of RNA extracted with the latter method is relatively poor compared to that with the former method. Thus, the step of exosome isolation contributes to the efficiency of extracellular RNA extraction from serum.

Example 6: Microarray analysis of mRNA.

[00101] Microarray analysis of the mRNA population in glioblastoma cells and microvesicles derived from them was performed by Miltenyi Biotech (Auburn, CA, USA) using the Agilent Whole Human Genome Microarray, 4x44K, two color array. The microarray analysis was performed on two different RNA preparations from primary glioblastoma cells and their corresponding microvesicles RNA preparations prepared as described in Examples 1 and 2. The data was analyzed using the GeneSifter software (Vizxlabs, Seattle, WA, USA). The Intersector software (Vizxlabs) was used to extract the genes readily detected on both arrays. The microarray data have been deposited in NCBI's Gene Expression Omnibus and are accessible through GEO series accession number GSE13470.

[00102] We found approximately 22,000 gene transcripts in the cells and 27,000 gene transcripts in the microvesicles that were detected well above background levels (99% confidence interval) on both arrays. Approximately 4,700 different mRNAs were detected exclusively in microvesicles on both arrays, indicating a selective enrichment process within the microvesicles. Consistent with this, there was a poor overall correlation in levels of mRNAs in the microvesicles as compared to their cells of origin from two tumor cell preparations (FIGS. 2a and 2b). In contrast, there was a good correlation in levels of mRNA from one cell culture (A) versus the second cell culture (B) (FIG. 2c) and a similar correlation in levels of mRNA from the corresponding microvesicles (A) and (B) (FIG. 2d). Accordingly, there is a consistency of mRNA distribution within the tumor cells and microvesicles. In comparing the ratio of transcripts in the microvesicles versus their cells of origin, we found 3,426 transcripts differentially distributed more than 5-fold (p-value <0.01). Of these, 2,238 transcripts were enriched (up to 380 fold) and 1,188 transcripts were less abundant (up to 90 fold) than in the cells (FIG. 5). The intensities and ratios of all gene transcripts were documented. The ontologies of mRNA transcripts enriched or reduced more than 10-fold were recorded and reviewed.

[00103] The mRNA transcripts that were highly enriched in the microvesicles were not always the ones that were most abundant in the microvesicles. The most abundant transcripts

would be more likely to generate an effect in the recipient cell upon delivery, and therefore the 500 most abundant mRNA transcripts present in microvesicles were divided into different biological processes based on their ontology descriptions (FIG. 6a). Of the various ontologies, angiogenesis, cell proliferation, immune response, cell migration and histone modification were selected for further study as they represent specific functions that could be involved in remodeling the tumor stroma and enhancing tumor growth. Glioblastoma microvesicle mRNAs belonging to these five ontologies were plotted to compare their levels and contribution to the mRNA spectrum (FIG. 6b). All five ontologies contained mRNAs with very high expression levels compared to the median signal intensity level of the array.

[00104] A thorough analysis of mRNAs that are enriched in the microvesicles versus donor cells, suggests that there may be a cellular mechanism for localizing these messages into microvesicles, possibly via a "zip code" in the 3'UTR as described for mRNAs translated in specific cellular locations, such as that for beta actin (Kislauskis et al., 1994). The conformation of the mRNAs in the microvesicles is not known, but they may be present as ribonuclear particles (RNPs) (Mallardo et al., 2003) which would then prevent degradation and premature translation in the donor cell.

[00105] Microarray analysis of the mRNA populations in glioblastoma cells and microvesicles derived from glioblastoma cells, melanoma cells, and microvesicles derived from melanoma cells was performed by Illumina Inc. (San Diego, CA, USA) using the Whole-Genome cDNA-mediated Annealing, Selection, Extension, and Ligation (DASL) Assay. The Whole-Genome DASL Assay combines the PCR and labeling steps of Illumina's DASL Assay with the gene-based hybridization and whole-genome probe set of Illumina's HumanRef-8 BeadChip. This BeadChip covers more than 24,000 annotated genes derived from RefSeq (Build 36.2, Release 22). The microarray analysis was performed on two different RNA preparations from primary glioblastoma cells, microvesicles from glioblastomas cells (derived with the method as described in Examples 1 and 2), melanoma cells, and microvesicles from melanoma cells (derived with the method as described in Examples 1 and 2).

[00106] The expression data for each RNA preparation were pooled together and used to generate a cluster diagram. As shown in FIG. 7, mRNA expression profiles for glioblastoma cells, microvesicles from glioblastomas cells, melanoma cells, and microvesicles from melanoma cells are clustered together, respectively. Expression profiles of the two primary glioblastoma cell lines 20/3C and 11/5c are clustered with a distance of about 0.06.

Expression profiles of the two primary melanoma cell lines 0105C and 0664C are clustered with a distance of about 0.09. Expression profiles of exosomes from the two primary melanoma cell lines 0105C and 0664C are clustered together with a distance of around 0.15. Expression profiles of exosomes from the two primary glioblastomas cell lines 20/3C and 11/5c are clustered together with a distance of around 0.098. Thus, exosomes from glioblastoma and melanoma have distinctive mRNA expression signatures and the gene expression signature of exosomes differs from that of their original cells. These data demonstrate that mRNA expression profiles from microvesicles may be used in the methods described herein for the diagnosis and prognosis of cancers.

Example 7: Glioblastoma microvesicles contain miRNA

[00107] Mature miRNA from microvesicles and from donor cells was detected using a quantitative miRNA reverse transcription PCR. Specifically, total RNA was isolated from microvesicles and from donor cells using the mirVana RNA isolation kit (Applied Biosystems, Foster City, CA, USA). Using the TaqMan® MicroRNA Assay kits (Applied Biosystems, Foster City, CA, USA), 30 ng total RNA was converted into cDNA using specific miR-primers and further amplified according to the manufacturer's protocol.

[00108] A subset of 11 miRNAs among those known to be up-regulated and abundant in gliomas was analyzed in microvesicles purified from two different primary glioblastomas (GBM 1 and GBM 2). These subset contained let-7a, miR-15b, miR-16, miR-19b, miR-21, miR-26a, miR-27a, miR-92, miR-93, miR-320 and miR-20. All of these miRNA were readily detected in donor cells and in microvesicles (FIG. 8). The levels were generally lower in microvesicles per μg total RNA than in parental cells (10%, corresponding to approximately 3 Ct-values), but the levels were well correlated, indicating that these 11 miRNA species are not enriched in microvesicles.

[00109] Microarray analysis of the microRNA populations in glioblastoma cells and microvesicles derived from glioblastoma cells, melanoma cells, and microvesicles derived from melanoma cells was performed by Illumina Inc. (San Diego, CA, USA) using the MicroRNA Expression Profiling Panel, powered by the DASL Assay. The human MicroRNA Panels include 1146 microRNA species. The microarray analysis was performed on two different RNA preparations from primary glioblastoma cells, microvesicles from glioblastomas cells (derived using the method described in Examples 1 and 2), melanoma

cells, and microvesicles from melanoma cells (derived using the method described in Examples 1 and 2).

[00110] The expression data for each RNA preparation were pooled together and used to generate a cluster diagram. As shown in FIG. 9, microRNA expression profiles for glioblastoma cells, microvesicles from glioblastomas cells, melanoma cells, and microvesicles from melanoma cells are clustered together, respectively. Expression profiles of the two primary melanoma cell lines 0105C and 0664C are clustered with a distance of about 0.13. Expression profiles of the two primary glioblastomas cell lines 20/3C and 11/5c are clustered with a distance of about 0.12. Expression profiles of exosomes from the two primary glioblastomas cell lines 20/3C and 11/5c are clustered together with a distance of around 0.12. Expression profiles of exosomes from the two primary melanoma cell lines 0105C and 0664C are clustered together with a distance of around 0.17. Thus, exosomes from glioblastoma and melanoma have distinctive microRNA expression signatures and that the gene expression signature of exosomes differs from that of their original cells. Furthermore, as demonstrated herein, microRNA expression profiles from microvesicles may be used in the methods described herein for the diagnosis and prognosis of cancers.

[00111] The finding of miRNAs in microvesicles suggests that tumor-derived microvesicles can modify the surrounding normal cells by changing their transcriptional/translational profiles. Furthermore, as demonstrated herein, miRNA expression profile from microvesicles may be used in the methods described herein for the diagnosis and prognosis of cancers, including but not limited to glioblastoma.

Examples 8-15. These examples show that nucleic acids within exosomes from bodily fluids can be used as biomarkers for diseases or other medical conditions.

Example 8: Expression profiles of miRNAs in microvesicles can be used as sensitive biomarkers for glioblastoma.

[00112] To determine if microRNAs within exosomes may be used as biomarkers for a disease and/or medical condition, we examined the existence of a correlation between the expression level of microRNA and disease status. Since microRNA-21 is expressed at high levels in glioblastoma cells and is readily detectable in exosomes isolated from serum of glioblastoma patients, we measured quantitatively microRNA-21 copy numbers within exosomes from the sera of glioblastoma patients by quantitative RT-PCR. Specifically, exosomes were isolated from 4 ml serum samples from 9 normal human subjects and 9

glioblastoma patients. The RNA extraction procedure was similar to the RNA extraction procedure as described in Example 2. The level of miR-21 was analyzed using singleplex qPCR (Applied Biosystems) and normalized to GAPDH expression level.

[00113] As shown in FIG. 10, the average Ct-value was 5.98 lower in the glioblastoma serum sample, suggesting that the exosomal miRNA-21 expression level in glioblastoma patients is approximately 63 fold higher than that in a normal human subject. The difference is statistically significant with a p value of 0.01. Therefore, there is a correlation between microRNA-21 expression level and glioblastoma disease status, which demonstrates that validity and applicability of the non-invasive diagnostic methods disclosed herein. For example, in one aspect, the method comprised the steps of isolating exosomes from the bodily fluid of a subject and analyzing microRNA-21 expression levels within the exosomes by measuring the copy number of microRNA-21 and comparing the number to that within exosomes from a normal subject or to a standard number generated by analyzing micro-RNA-21 contents within exosomes from a group of normal subjects. An increased copy number indicates the existence of glioblastoma in the subject; while the absence of an increased copy number indicates the absence of glioblastoma in the subject. This basic method may be extrapolated to diagnose/monitor other diseases and/or medical conditions associated with other species of microRNAs.

Example 9: mRNAs in microvesicles can be used as sensitive biomarkers for diagnosis

[00114] Nucleic acids are of high value as biomarkers because of their ability to be detected with high sensitivity by PCR methods. Accordingly, the following tests were designed and carried out to determine whether the mRNA in microvesicles could be used as biomarkers for a medical disease or condition, in this case glioblastoma tumors. The epidermal growth factor receptor (EGFR) mRNA was selected because the expression of the EGFRvIII mutation is specific to some tumors and defines a clinically distinct subtype of glioma (Pelloski et al., 2007). In addition, EGFRvIII mutations traditionally cannot be detected using tissues other than the lesion tissues since these mutations are somatic mutations but not germ line mutations. Therefore, a biopsy from lesion tissues such as glioma tumor is conventionally required for detecting EGFRvIII mutations. As detailed below, nested RT-PCR was used to identify EGFRvIII mRNA in glioma tumor biopsy samples and the results compared with the mRNA species found in microvesicles purified from a serum sample from the same patient.

[00115] Microvesicles were purified from primary human glioblastoma cells followed by RNA extraction from both the microvesicles and donor cells (biopsy). The samples were coded and the PCRs were performed in a blind fashion. Gli-36EGFRvIII (human glioma cell stably expressing EGFRvIII) was included as a positive control. The microvesicles from 0.5-2 ml of frozen serum samples were pelleted as described in Example 2 and the RNA was extracted using the MirVana Microvesicles RNA isolation kit. Nested RT-PCR was then used to amplify both the wild type EGFR (1153 bp) and EGFRvIII (352 bp) transcripts from both the microvesicles and donor cells using the same set of primers. Specifically, the RNA was converted to cDNA using the Omniscript RT kit (Qiagen Inc, Valencia, CA, USA) according to the manufacturer's recommended protocol. GAPDH primers were GAPDH Forward (SEQ ID NO: 9) and GAPDH Reverse (SEQ ID NO: 10). The EGFR/EGFRvIII PCR1 primers were SEQ ID NO: 11 and SEQ ID NO: 12. The EGFR/EGFRvIII PCR2 primers were SEQ ID NO: 13 and SEQ ID NO: 14. The PCR cycling protocol was 94 °C for 3 minutes; 94 °C for 45 seconds, 60 °C for 45 seconds, 72 °C for 2 minutes for 35 cycles; and a final step 72 °C for 7 minutes.

[00116] We analyzed the biopsy sample to determine whether the EGFRvIII mRNA was present and compared the result with RNA extracted from exosomes purified from a frozen serum sample from the same patient. Fourteen of the 30 tumor samples (47%) contained the EGFRvIII transcript, which is consistent with the percentage of glioblastomas found to contain this mutation in other studies (Nishikawa et al., 2004). EGFRvIII could be amplified from exosomes in seven of the 25 patients (28%) from whom serum was drawn around the time of surgery (FIG. 11 and Table 1). When a new pair of primers EGFR/EGFRvIII PCR3: SEQ ID NO: 15 and SEQ ID NO: 16, were used as the second primer pair for the above nested PCR amplification, more individuals were found to harbor EGFRvIII mutations (Table 1). EGFRvIII could be amplified from exosomes in the six patients who was identified as negatives with the old pair of primers EGFRvIII PCR2: SEQ ID NO: 13 AND SEQ ID NO: 14. Notably, exosomes from individual 13, whose biopsy did not show EGFRvIII mutation, was shown to contain EGFRvIII mutation, suggesting an increased sensivity of EGFRvIII mutation detection using exosomes technology. From the exosomes isolated from 52 normal control serum samples, EGFRvIII could not be amplified (FIG. 12). Interestingly, two patients with an EGFRvIII negative tumor sample turned out to be EGFRvIII positive in the serum exosomes, supporting heterogeneous foci of EGFRvIII expression in the glioma tumor. Furthermore, our data also showed that intact RNAs in microvesicles were, unexpectedly,

able to be isolated from frozen bodily serum of glioblastoma patients. These blind serum samples from confirmed glioblastoma patients were obtained from the Cancer Research Center (VU medical center, Amsterdam, the Netherlands) and were kept at -80°C until use. The identification of tumor specific RNAs in serum microvesicles allows the detection of somatic mutations which are present in the tumor cells. Such technology should result in improved diagnosis and therapeutic decisions.

[00117] The RNA found in the microvesicles contains a "snapshot" of a substantial array of the cellular gene expression profile at a given time. Among the mRNA found in glioblastoma-derived microvesicles, the EGFR mRNA is of special interest since the EGFRvIII splice variant is specifically associated with glioblastomas (Nishikawa et al., 2004). Here it is demonstrated that brain tumors release microvesicles into the bloodstream across the blood-brain-barrier (BBB), which has not been shown before. It is further demonstrated that mRNA variants, such as EGFRvIII in brain tumors, are able to be detected by a method comprising the steps of isolating exosomes from a small amount of patient serum and analyzing the RNA in said microvesicles.

[00118] Knowledge of the EGFRvIII mutation in tumors is important in choosing an optimal treatment regimen. EGFRvIII-positive gliomas are over 50 times more likely to respond to treatment with EGFR-inhibitors like erlotinib or gefitinib (Mellinghoff et al., 2005).

Example 10: Diagnosis of iron metabolism disorders

[00119] The exosome diagnostics method can be adapted for other purposes as shown by the following example.

[00120] Hepcidin, an antimicrobial peptide, is the master hormonal regulator of iron metabolism. This peptide is produced mainly in mammalian liver and is controlled by the erythropoietic activity of the bone-marrow, the amount of circulating and stored body iron, and inflammation. Upon stimulation, hepcidin is secreted into the circulation or urine where it may act on target ferroportin-expressing cells. Ferroportin is the sole iron exporter identified to date and when bound to hepcidin, it is internalized and degraded. The resulting destruction of ferroportin leads to iron retention in ferroportin expressing cells such as macrophages and enterocytes. This pathophysiological mechanism underlies anemia of chronic diseases. More specifically, inappropriately high levels of hepcidin and elevated iron content within the reticuloendothelial system characterize anemia. Indeed, anemia may be

associated with many diseases and/or medical conditions such as infections (acute and chronic), cancer, autoimmune, chronic rejection after solid-organ transplantation, and chronic kidney disease and inflammation (Weiss and Goodnough, 2005). On the other hand, in a genetic iron overload disease such as hereditary hemochromatosis, inappropriately low expression levels of hepcidin encourage a potentially fatal excessive efflux of iron from within the reticuloendothelial system. So, hepcidin is up-regulated in anemia associated with chronic disease, but down-regulated in hemochromatosis.

[00121] Currently, there is no suitable assay to quantitatively measure hepcidin levels in circulation or urine (Kemna et al., 2008) except time-of-flight mass spectrometry (TOF MS), which needs highly specialized equipment, and therefore is not readily accessible. Recently, the method of Enzyme Linked ImmunoSorbent Assay (ELISA) has been proposed to quantitatively measure hepcidin hormone levels but this method is not consistent because of the lack of clear correlations with hepcidin (Kemna et al., 2005; Kemna et al., 2007) and other iron related parameters (Brookes et al., 2005; Roe et al., 2007).

[00122] Hepcidin mRNA was detected in exosomes from human serum, as follows. Exosomes were first isolated from human serum and their mRNA contents extracted before conversion to cDNA and PCR amplification. PCR primers were designed to amplify a 129 nucleotide fragment of human Hepcidin. The sequences of the primers are SEQ ID NO: 57 and SEQ ID NO: 58. A hepcidin transcript of 129 nucleotides (the middle peak in FIG. 13D) was readily detected by Bioanalyzer. As a positive control (FIG. 13B), RNA from a human hepatoma cell line Huh-7 was extracted and converted to cDNA. The negative control (FIG. 13C) is without mRNA. These Bioanalyzer data are also shown in the pseudogel in FIG. 13A.

[00123] Hepcidin mRNA in microvesicles in circulation correlates with hepcidin mRNA in liver cells. Hence, measuring hepcidin mRNA within microvesicles in a bodily fluid sample would allow one to diagnose or monitor anemia or hemochromatosis in the subject.

[00124] Thus, it is possible to diagnose and/or monitor anemia and hemochromatosis in a subject by isolating microvesicles from a bodily fluid and comparing the hepcidin mRNA in said microvesicles with the mRNA from from a normal subject. With an anemic subject, the copy number of mRNA is increased over the normal, non-anemic level. In a subject suffering from hemochromatosis, the copy number is decreased relative to the mRNA in a normal subject.

Example 11: Non-invasive transcriptional profiling of exosomes for diabetic nephropathy diagnosis

[00125] Diabetic nephropathy (DN) is a life threatening complication that currently lacks specific treatments. Thus, there is a need to develop sensitive diagnostics to identify patients developing or at risk of developing DN, enabling early intervention and monitoring.

[00126] Urine analysis provides a way to examine kidney function without having to take a biopsy. To date, this analysis has been limited to the study of protein in the urine. This Example sets forth a method to obtain from urine transcriptional profiles derived from cells that normally could only be obtained by kidney biopsy. Specifically, the method comprises the steps of isolating urine exosomes and analyzing the RNAs within said exosomes to obtain transcriptional profiles, which can be used to examine molecular changes being made by kidney cells in diabetic individuals and provide a 'snap shot' of any new proteins being made by the kidney. State-of-the-art technologies to obtain exosomal transcription profiles include, but are not limited to, contemporary hybridization arrays, PCR based technologies, and next generation sequencing methods. Since direct sequencing does not require pre-designed primers or spotted DNA oligos, it will provide a non-biased description of exosomal RNA profiles. An example of next generation sequencing technology is provided by the Illumina Genome Analyzer, which utilizes massively parallel sequencing technology which allows it to sequence the equivalent of 1/3 a human genome per run. The data obtainable from this analysis would enable one to rapidly and comprehensively examine the urinary exosomal transcriptional profile and allow comparison to the whole kidney. Using such a method, one could obtain much needed information regarding the transcription profile of urinary exosomes. A comparison of transcripts in control versus diabetes-derived urinary exosomes could further provide one with a comprehensive list of both predicted and new biomarkers for diabetic nephropathy.

[00127] In order to prove the feasibility of the diagnostic method described above, an experiment was designed and carried out to isolate urinary exosomes and to confirm the presence of renal specific biomarkers within these exosomes. In this experiment, a fresh morning urine sample of 220 ml was collected from a 28-year old healthy male subject and processed via differential centrifugation to isolate urinary exosomes. Specifically, urine was first spun at 300 x g spin for 10 minutes to remove any cells from the sample. The supernatant was collected and then underwent a 20-minute 16,500 x g spin to bring down any cell debris or protein aggregates. The supernatant was then passed through a 0.22 uM

membrane filter to remove debris with diameters larger than 0.22uM. Finally, the sample underwent ultra-centrifugation at 100,000 x g for 1 hour to pellet the exosomes (Thery et al., 2006). The pellet was gently washed in phosphate buffered saline (PBS) and RNA was extracted using a Qiagen RNeasy kit pursuant to the manufacturer's instructions. The isolated RNA was converted to cDNA using the Omniscript RT kit (Qiagen) followed by PCR amplification of renal specific genes.

[00128] The renal specific genes examined and their corresponding renal area where the gene is expressed are as follows: AQP1 – proximal tubules; AQP2 – distal tubule (principal cells); CUBN – proximal tubules; LRP2 – proximal tubules; AVPR2 – proximal and distal tubules; SLC9A3 (NHE-3) - Proximal tubule; ATP6V1B1 – distal tubule (intercalated cells); NPHS1 – glomerulus (podocyte cells); NPHS2 – glomerulus (podocyte cells); and CLCN3 – Type B intercalated cells of collecting ducts. The sequences of the primers designed to amplify each gene are AQP1-F (SEQ ID NO: 17) and AQP1-R (SEQ ID NO: 18); AQP2-F (SEQ ID NO: 19) and AQP2-R (SEQ ID NO: 20); CUBN-F (SEQ ID NO: 21) and CUBN-R (SEQ ID NO: 22); LRP2-F (SEQ ID NO: 23) and LRP2-R (SEQ ID NO: 24); AVPR2-F (SEQ ID NO: 25) and AVPR2-R (SEQ ID NO: 26); SLC9A3-F (SEQ ID NO: 27) and SLC9A3-R (SEQ ID NO: 28); ATP6V1B1-F (SEQ ID NO: 29) and ATP6V1B1-R (SEQ ID NO: 30); NPHS1-F (SEQ ID NO: 31) and NPHS1-R (SEQ ID NO: 32); NPHS2-F (SEQ ID NO: 33) and NPHS2-R (SEQ ID NO: 34); CLCN5-F (SEQ ID NO: 35) and CLCN5-R (SEQ ID NO: 36).

[00129] The expected sizes of the PCR products for each gene are AQP1-226bp, AQP2-208bp, CUBN-285bp, LRP2-220bp, AVPR2-290bp, SLC9A3-200bp, ATP6V1B1-226bp, NPHS1-201bp, NPHS2-266bp and CLCN5-204bp. The PCR cycling protocol was 95 °C for 8 minutes; 95 °C for 30 seconds, 60 °C for 30 seconds, 72 °C for 45 seconds for 30 cycles; and a final step 72 °C for 10 minutes.

[00130] As shown in FIG. 14a, kidney tubule cells contain multivesicular bodies, which is an intermediate step during exosome generation. Exosomes isolated from these cells can be identified by electron microscopy (FIG. 14b). Analysis of total RNA extracted from urinary exosomes indicates the presence of RNA species with a broad range of sizes (FIG. 14c). 18S and 28S ribosomal RNAs were not found. PCR analysis confirmed the presence of renal specific transcripts within urinary exosomes (FIG. 14d). These data show that kidney cells shed exosomes into urine and these urinary exosomes contain transcripts of renal

origin, and that the exosome method can detect renal biomarkers associated with certain renal diseases and/or other medical conditions.

[00131] To further confirm the presence of renal specific mRNA transcripts in urinary exosomes, an independent set of experiments were performed using urine samples from six individuals. Exosomal nucleic acids were extracted from 200ml morning urine samples from each indivisual following a procedure as mentioned above. Specifically, urine samples underwent differential centrifugation starting with a 1000 ×g centrifugation to spin down whole cells and cell debris. The supernatant was carefully removed and centrifuged at 16,500 ×g for 20 minutes. The follow-on supernatant was then removed and filtered through a 0.8μm filter to remove residual debris from the exosome containing supernatant. The final supernatant then underwent ultracentrifugation at 100,000 ×g for 1hr 10min. The pellet was washed in nuclease free PBS and re-centrifuged at 100,000 ×g for 1hr 10min to obtain the exosomes pellet which is ready for nucleic acid extraction. Nucleic acids were extracted from the pelleted exosomes using the Arcturus PicoPure RNA Isolation kit and the nucleic acid concentration and integrity was analyzed using a Bioanalyzer (Agilent) Pico chip. As shown in FIG. 14e, nucleic acids isolated from urinary exosomes vary from individual to individual. To test whether the presence of renal biomarkers also varies from individual to individual, PCR amplifications were carried out for Aquaporin1, Aquaporin2 and Cubilin gene using a new set of primer pairs: AQP1 new primer pair: SEQ ID NO: 37 and SEQ ID NO: 38; AQP2 new primer pair: SEQ ID NO: 39 and SEQ ID NO: 40; CUBN new primer pair: SEQ ID NO: 41 and SEQ ID NO: 42. These primer pairs were designed specifically to amplify the spliced and reverse transcribed cDNA fragments. Reverse transcription was performed using the Qiagen Sensiscript kit. As shown in FIG. 14f, no amplification was seen in individual 1, probably due to failed nucleic acid extraction. AQP1 was amplified only in individual 2. CUBN was amplified in indivisual 2 and 3. And AQP2 was amplified in individual 2, 3, 4 and 5. In comparison actin gene (indicated by "House" in FIG. 14f) was amplified in individual 2, 3, 4, 5 and 6. These data provide more evidence that urinary exosomes contain renal specific mRNA transcripts although the expression levels are different between different individuals.

[00132] To test the presence of cDNAs in urinary exosomes, a 200ml human urine sample was split into two 100ml urine samples. Urinary exosomes were isolated from each sample. Exosomes from one sample were treated with DNase and those from the other sample were mock treated. Exosomes from each sample were then lysed for nucleic acid

extraction using PicoPure RNA isolation kit (Acturus). The nucleic acids were used as templates for nested-PCR amplification (PCR protocols described in Example 9) without prior reverse transcription. The primer pairs to amplify the actin gene were Actin-FOR (SEQ ID NO: 43) and Actin-REV (SEQ ID NO: 44); Actin-nest-FOR (SEQ ID NO: 45) and Actin-nest-REV (SEQ ID NO: 46) with an expected final amplicon of 100bp based on the actin gene cDNA sequence. As shown in FIG. 14g, the 100bp fragments were present in the positive control (human kidney cDNA as templates), DNase treated and non-treated exosomes, but absent in the negative control lane (without templates). Accordingly, actin cDNA is present in both the DNase treated and non-treated urinary exosomes.

[00133] To test whether most nucleic acids extracted using the method were present within exosomes, the nucleic acids extracted from the DNase treated and non-treated exosomes were dissolved in equal volumes and analyzed using a RNA Pico chip (Agilent Technologies). As shown in FIG. 14h, the concentration of the isolated nucleic acids from the DNase treated sample was 1,131 pg/ul and that from the non-treated sample was 1,378 pg/ul. Thus, more than 80% nucleic acids extracted from urinary exosomes using the above method were from inside exosomes.

[00134] To identify the content of urinary exosomes systematically, nucleic acids were extracted from urinary exosomes and submitted to the Broad Institute for sequencing. Approximately 14 million sequence reads were generated, each 76 nucleotides in length. These sequence reads correspond to fragments of DNA/RNA transcripts present within urinary exosomes. Using an extremely strict alignment parameter (100% identity over full length sequence), approximately 15% of the reads were aligned to the human genome. This percentage would likely increase if less stringent alignment criteria was used. A majority of these 15% reads did not align with protein coding genes but rather with non-coding genomic elements such are transposons and various LINE & SINE repeat elements. Notably, for those reads that are not aligned to the human genome, many are aligned to viral sequences. To the extent that the compositions and levels of nucleic acids contained in urinary exosomes change with respect to a disease status, profiles of the nucleic acids could be used according to the present methods as biomarkers for disease diagnosis.

[00135] This example demonstrates that the exosome method of analyzing urine exosomes can be used to determine cellular changes in the kidney in diabetes-related kidney disease without having to take a high-risk, invasive renal biopsy. The method provides a new and sensitive diagnostic tool using exosomes for early detection of kidney diseases such as

diabetic nephropathy. This will allow immediate intervention and treatment. In sum, the exosome diagnostic method and technology described herein provides a means of much-needed diagnostics for diabetic nephropathy and other diseases which are associated with certain profiles of nucleic acids contained in urinary exosomes.

Example 12: Prostate cancer diagnosis and urinary exosomes

[00136] Prostate cancer is the most common cancer in men today. The risk of prostate cancer is approximately 16%. More than 218,000 men in the United States were diagnosed in 2008. The earlier prostate cancer is detected, the greater are the chances of successful treatment. According to the American Cancer Society, if prostate cancers are found while they are still in the prostate itself or nearby areas, the five-year relative survival rate is over 98%.

[00137] One established diagnostic method is carried out by measuring the level of prostate specific antigen (PSA) in the blood, combined with a digital rectal examination. However, both the sensitivity and specificity of the PSA test requires significant improvement. This low specificity results in a high number of false positives, which generate numerous unnecessary and expensive biopsies. Other diagnostic methods are carried out by detecting the genetic profiles of newly identified biomarkers including, but not limited to, prostate cancer gene 3 (PCA3) (Groskopf et al., 2006; Nakanishi et al., 2008), a fusion gene between transmembrane protease serine 2 and ETS-related gene (TMPRSS2-ERG) (Tomlins et al., 2005), glutathione S-transferase pi (Goessl et al., 2000; Gonzalgo et al., 2004), and alpha-methylacyl CoA racemase (AMACR) (Zehentner et al., 2006; Zielie et al., 2004) in prostate cancer cells found in bodily fluids such as serum and urine (Groskopf et al., 2006; Wright and Lange, 2007). Although these biomarkers may give increased specificity due to overexpression in prostate cancer cells (e.g., PCA3 expression is increased 60- to 100-fold in prostate cancer cells), a digital rectal examination is required to milk prostate cells into the urine just before specimen collection (Nakanishi et al., 2008). Such rectal examinations have inherent disadvantages such as the bias on collecting those cancer cells that are easily milked into urine and the involvement of medical doctors which is costly and time consuming.

[00138] Here, a new method of detecting the genetic profiles of these biomarkers is proposed to overcome the limitation mentioned above. The method comprises the steps of isolating exosomes from a bodily fluid and analyzing the nucleic acid from said exosomes. The procedures of the method are similar to those detailed in Example 9. In this example, the

urine samples were from four diagnosed prostate cancer patients. As shown in FIG. 15c, the cancer stages were characterized in terms of grade, Gleason stage and PSA levels. In addition, the nucleic acids analyzed by nested-RT-PCR as detailed in Example 7 were TMPRSS2-ERG and PCA3, two of the newly identified biomarkers of prostate cancer. For amplification of TMPRSS2-ERG, the primer pair for the first amplification step was TMPRSS2-ERG F1 (SEQ ID NO: 47) and TMPRSS2-ERG R1 (SEQ ID NO: 48); and the primer pair for the second amplification step was TMPRSS2-ERG F2 (SEQ ID NO: 49) and TMPRSS2-ERG R2 (SEQ ID NO: 50). The expected amplicon is 122 base pairs (bp) and gives two fragments (one is 68 bp, the other is 54 bp) after digestion with the restriction enzyme HaeII. For amplification of PCA3, the primer pair for the first amplification step was PCA3 F1 (SEQ ID NO: 51) and PCA3 R1 (SEQ ID NO: 52); and the primer pair for the second amplification step was PCA3 F2 (SEQ ID NO: 53) and PCA3 R2 (SEQ ID NO: 54). The expected amplicon is 152 bp in length and gives two fragments (one is 90 bp, the other is 62 bp) after digestion with the restriction enzyme ScaI.

[00139] As shown in FIG. 15a, in both patient 1 and 2, but not in patient 3 and 4, the expected amplicon of TMPRSS2-ERG could be detected and digested into two fragments of expected sizes. As shown in FIG. 15b, in all four patients, the expected amplicon of PCA3 could be detected and digested into two fragments of expected sizes. Therefore, PCA3 expression could be detected in urine samples from all four patients, while TMPRSS2-ERG expression could only be detected in urine samples from patient 1 and 2 (FIG. 15c). These data, although not conclusive due to the small sample size, demonstrate the applicability of the new method in detecting biomarkers of prostate cancer. Further, the exosome method is not limited to diagnosis but can be employed for prognosis and/or monitoring other medical conditions related to prostate cancer.

Example 13: Microvesicles in non-invasive prenatal diagnosis

[00140] Prenatal diagnosis is now part of established obstetric practice all over the world. Conventional methods of obtaining fetal tissues for genetic analysis includes amniocentesis and chorionic villus sampling, both of which are invasive and confer risk to the unborn fetus. There is a long-felt need in clinical genetics to develop methods of non-invasive diagnosis. One approach that has been investigated extensively is based on the discovery of circulating fetal cells in maternal plasma. However, there are a number of barriers that hinder its application in clinical settings. Such barriers include the scarcity of fetal cells (only 1.2 cells/ml maternal blood), which requires relatively large volume blood

samples, and the long half life of residual fetal cells from previous pregnancy, which may cause false positives. Another approach is based on the discovery of fetal DNA in maternal plasma. Sufficient fetal DNA amounts and short clearance time overcome the barriers associated with the fetal cell method. Nevertheless, DNA only confers inheritable genetic and some epigenetic information, both of which may not represent the dynamic gene expression profiles that are linked to fetal medical conditions. The discovery of circulating fetal RNA in maternal plasma (Ng et al., 2003b; Wong et al., 2005) may be the method of choice for non-invasive prenatal diagnosis.

[00141] Several studies suggest that fetal RNAs are of high diagnostic value. For example, elevated expression of fetal corticotropin-releasing hormone (CRH) transcript is associated with pre-eclampsia (a clinical condition manifested by hypertension, edema and proteinuria) during pregnancy (Ng et al., 2003a). In addition, the placenta-specific 4 (PLAC4) mRNA in maternal plasma was successfully used in a non-invasive test for aneuploid pregnancy (such as trisomy 21, Down syndrome) (Lo et al., 2007). Furthermore, fetal human chorionic gonadotropin (hCG) transcript in maternal plasma may be a marker of gestational trophoblastic diseases (GTDs), which is a tumorous growth of fetal tissues in a maternal host. Circulating fetal RNAs are mainly of placenta origin (Ng et al., 2003b). These fetal RNAs can be detected as early as the 4th week of gestation and such RNA is cleared rapidly postpartum.

[00142] Prenatal diagnosis using circulating fetal RNAs in maternal plasma, nevertheless, has several limitations. The first limitation is that circulating fetal RNA is mixed with circulating maternal RNA and is not effectively separable. Currently, fetal transcripts are identified, based on an assumption, as those that are detected in pregnant women antepartum as well as in their infant's cord blood, yet are significantly reduced or absent in maternal blood within 24 or 36 hours postpartum (Maron et al., 2007). The second limitation is that no method is established to enrich the circulating fetal RNA for enhanced diagnostic sensitivity since it is still unknown how fetal RNA is packaged and released. The way to overcome these limitations may lie in the isolation of microvesicles and the analysis of the fetal RNAs therein.

[00143] Several facts suggest that microvesicles, which are shed by eukaryotic cells, are the vehicles for circulating fetal RNAs in maternal plasma. First, circulating RNAs within microvesicles are protected from RNase degradation. Second, circulating fetal RNAs have been shown to remain in the non-cellular fraction of maternal plasma, which is consistent

with the notion that microvesicles encompassing these fetal RNAs are able to be filtered through 0.22 um membrane. Third, similar to tumorous tissues which are know to shed microvesicles, placental cells, which are a pseudo-malignant fetal tissue, are most likely capable of shedding microvesicles. Thus, a novel method of non-invasive prenatal diagnosis is comprised of isolating fetal microvesicles from maternal blood plasma and then analyzing the nucleic acids within the microvesicles for any genetic variants associated with certain diseases and/or other medical conditions.

[00144] A hypothetical case of non-invasive prenatal diagnosis is as follows: peripheral blood samples are collected from pregnant women and undergo magnetic activated cell sorting (MACS) or other affinity purification to isolate and enrich fetus-specific microvesicles. The microvesicular pellet is resuspended in PBS and used immediately or stored at -20°C for further processing. RNA is extracted from the isolated microvesicles using the Qiagen RNA extraction kit as per the manufacturer's instructions. RNA content is analyzed for the expression level of fetal human chorionic gonadotropin (hCG) transcript. An increased expression level of hCG compared to the standard range points to the development of gestational trophoblastic diseases (GTDs) and entail further the need for clinical treatment for this abnormal growth in the fetus. The sensitivity of microvesicle technology makes it possible to detect the GTDs at a very early stage before any symptomatic manifestation or structural changes become detectable under ultrasonic examination. The standard range of hCG transcript levels may be determined by examining a statistically significant number of circulating fetal RNA samples from normal pregnancies.

[00145] This prenatal diagnostic method may be extrapolated to the prenatal diagnosis and/or monitoring of other diseases or medical conditions by examining those transcripts associated with these diseases or medical conditions. For example, extraction and analysis of anaplastic lymphoma kinase (ALK) nucleic acid from microvesicles of fetus origin from maternal blood is a non-invasive prenatal diagnosis of neuroblastoma, which is closely associated with mutations within the kinase domain or elevated expression of *ALK* (Mosse et al., 2008). Accordingly, the microvesicle methods and technology described herein may lead to a new era of much-needed, non-invasive prenatal genetic diagnosis.

Example 14: Melanoma diagnosis

[00146] Melanoma is a malignant tumor of melanocytes (pigment cells) and is found predominantly in skin. It is a serious form of skin cancer and accounts for 75 percent of all deaths associated with skin cancer. Somatic activating mutations (e.g. V600E) of BRAF are the earliest and most common genetic abnormality detected in the genesis of human melanoma. Activated BRAF promotes melanoma cell cycle progression and/or survival.

[00147] Currently, the diagnosis of melanoma is made on the basis of physical examination and excisional biopsy. However, a biopsy can sample only a limited number of foci within the lesion and may give false positives or false negatives. The exosome method provides a more accurate means for diagnosing melanoma. As discussed above, the method is comprised of the steps of isolating exosomes from a bodily fluid of a subject and analyzing the nucleic acid from said exosomes.

[00148] To determine whether exosomes shed by melanoma cells contain BRAF mRNA, we cultured primary melanoma cells in DMEM media supplemented with exosome-depleted FBS and harvested the exosomes in the media using a similar procedure as detailed in Example 2. The primary cell lines were Yumel and M34. The Yumel cells do not have the V600E mutation in BRAF, while M34 cells have the V600E mutation in BRAF. RNAs were extracted from the exosomes and then analyzed for the presence of BRAF mRNA by RT-PCR. The primers used for PCR amplification were: BRAF forward (SEQ ID NO: 55) and BRAF reverse (SEQ ID NO: 56). The amplicon is 118 base pairs (bp) long and covers the part of BRAF cDNA sequence where the V600E mutation is located. As shown in FIG. 16a, a band of 118 bp was detected in exosomes from primary melanoma cells (Yumel and M34 cells), but not in exosomes from human fibroblast cells or negative controls. The negative detection of a band of 118 bp PCR product is not due to a mistaken RNA extraction since GAPDH transcripts could be detected in exosomes from both melanoma cell and human fibroblast cells (FIG. 16b). The 118 bp PCR products were further sequenced to detect the V600E mutation. As shown in FIGS. 16c and 16d, PCR products from YUMEL cells, as expected, contain wild type BRAF mRNA. In contrast, PCR products from M34 cells, as expected, contain mutant BRAF mRNA with a T-A point mutation, which causes the amino acid Valine (V) to be replaced by Glutamic acid (E) at the amino acid position 600 of the BRAF protein. Furthermore, BRAF mRNA cannot be detected in exosomes from normal human fibroblast cells, suggesting the BRAF mRNA is not contained in exosomes of all tissue origins.

[00149] These data suggest that melanoma cells shed exosomes into the blood circulation and thus melanoma can be diagnosed by isolating these exosomes from blood serum and analyzing the nucleic acid therefrom for the presence or absence of mutations (e.g., V600E) in BRAF. The method described above can also be employed to diagnose melanomas that are associated with other BRAF mutations and mutations in other genes. The method can also be employed to diagnose melanomas that are associated with the expression profiles of BRAF and other nucleic acids.

Example 15: Detection of MMP levels from exosomes to monitor post transplantation conditions.

[00150] Organ transplants are usually effective treatments for organ failures. Kidney failure, heart disease, end-stage lung disease and cirrhosis of the liver are all conditions that can be effectively treated by a transplant. However, organ rejections caused by post-transplantation complications are major obstacles for long-term survival of the allograft recipients. For example, in lung transplantations, bronchiolitis obliterans syndrome is a severe complication affecting survival rates. In kidney transplants, chronic allograft nephropathy remains one of the major causes of renal allograft failure. Ischemia-reperfusion injury damages the donor heart after heart transplantation, as well as the donor liver after orthotopic liver transplantation. These post-transplantation complications may be ameliorated once detected at early stages. Therefore, it is essential to monitor post-transplantation conditions in order to alleviate adverse complications.

[00151] Alterations in the extracellular matrix contribute to the interstitial remodeling in post-transplantation complications. Matrix metalloproteinases (MMPs) are involved in both the turnover and degradation of extracellular matrix (ECM) proteins. MMPs are a family of proteolytic, zinc-dependent enzymes, with 27 members described to date, displaying multidomain structures and substrate specificities, and functioning in the processing, activation, or deactivation of a variety of soluble factors. Serum MMP levels may indicate the status of post-transplantation conditions. Indeed, circulating MMP-2 is associated with cystatin C, post-transplant duration, and diabetes mellitus in kidney transplant recipients (Chang et al., 2008). Disproportional expression of MMP-9 is linked to the development of bronchiolitis obliterans syndrome after lung transplantation (Hubner et al., 2005).

[00152] MMP mRNAs (MMP1, 8, 12, 15, 20, 21, 24, 26 and 27) can be detected in exosomes shed by glioblastoma cells as shown in Example 4 and Table 10. The present

exosome method, isolating exosomes from a bodily fluid and analyzing nucleic acids from said exosomes, can be used to monitor transplantation conditions. The exosome isolation procedure is similar to that detailed in Example 2. The present procedures to analyze nucleic acid contained within exosomes are detailed in Example 9. A significant increase in the expression level of MMP-2 after kidney transplantation will indicate the onset and/or deterioration of post-transplantation complications. Similarly, a significantly elevated level of MMP-9 after lung transplantation, suggests the onset and/or deterioration of bronchiolitis obliterans syndrome.

[00153] Therefore, the exosome method can be used to monitor post-transplantation conditions by determining the expression levels of MMP proteins associated with post-transplantation complications. It is also expected that the method can be extrapolated to monitor post-transplantation conditions by determining the expression of other marker genes as well as monitor other medical conditions by determining the genetic profile of nucleic acids associated with these medical conditions.

Examples 16-18. Microvesicles can be therapeutic agents or delivery vehicles of therapeutic agents.

Example 16: Microvesicle proteins induce angiogenesis in vitro.

[00154] A study was designed and carried out to demonstrate glioblastoma microvesicles contribute to angiogenesis. HBMVECs (30,000 cells), a brain endothelial cell line, (Cell Systems, Catalogue #ACBRI-376, Kirkland, WA, USA) were cultured on Matrigel-coated wells in a 24-well plate in basal medium only (EBM) (Lonza Biologics Inc., Portsmouth, NH, USA), basal medium supplemented with glioblastoma microvesicles (EBM+ MV) (7 μg/well), or basal medium supplemented with a cocktail of angiogenic factors (EGM; hydrocortisone, EGF, FGF, VEGF, IGF, ascorbic acid, FBS, and heparin; Singlequots (EBM positive control). Tubule formation was measured after 16 hours and analyzed with the Image J software. HBMVECs cultured in the presence of glioblastoma microvesicles demonstrated a doubling of tubule length within l 6 hours. The result was comparable to the result obtained with HBMCECs cultured in the presence of angiogenic factors (FIG. 18a). These results show that glioblastoma-derived microvesicles play a role in initiating angiogenesis in brain endothelial cells.

[00155] Levels of angiogenic proteins in microvesicles were also analyzed and compared with levels in glioblastoma donor cells. Using a human angiogenesis antibody

array, we were able to detect 19 proteins involved in angiogenesis. Specifically, total protein from either primary glioblastoma cells or purified microvesicles isolated from said cells were lysed in lysis buffer (Promega, Madison, WI, USA) and added to the human angiogenesis antibody array (Panomics, Fremont CA, USA) according to manufacturer's recommendations. The arrays were scanned and analyzed with the Image J software. As shown in FIG. 18b, of the seven of the 19 angiogenic proteins were readily detected in the microvesicles, 6 (angiogenin, IL-6, lL-8, TIMP-I, VEGF and TIMP-2) were present at higher levels on a total protein basis as compared to the glioblastoma cells (FIG. 18c). The three angiogenic proteins most enriched in microvesicles compared to tumor cells were angiogenin, IL-6 and 1L-8, all of which have been implicated in glioma angiogenesis with higher levels associated with increased malignancy (25-27).

[00156] Microvesicles isolated from primary glioblastoma cells were also found to promote proliferation of a human U87 glioma cell line. In these studies, 100 000 U87 cells were seeded in wells of a 24-well plate and allowed to grow for three days (DMEM-5%FBS) or DMEM-5%FBS supplemented with 125 μg microvesicles isolated from primary glioblastoma cells. After three days, untreated U87 cells (FIG. 19a) were found to be fewer in number as determined using a Burker chamber, than those supplemented with microvesicles (FIG. 19b). Both non-supplemented and supplemented U87 cells had increased 5-and 8-fold in number over this period, respectively (FIG. 19c). Thus, glioblastoma microvesicles appear to stimulate proliferation of other glioma cells.

Example 17: Glioblastoma microvesicles are taken up by HBMVECs.

[00157] To demonstrate that glioblastoma microvesicles are able to be taken up by human brain microvesicular endothelial cells (HBMVECs), purified glioblastoma microvesicles were labeled with PKH67 Green Fluorescent labeling kit (Sigma-Aldrich, St Louis, MO, USA). The labeled microvesicles were incubated with HBMVEC in culture (5 μg/50,000 cells) for 20 min at 4°C. The cells were washed and incubated at 37°C for 1 hour. Within 30 min the PKH67-labeled microvesicles were internalized into endosome-like structures within the HBMVECs (FIG. 17a). These results show that glioblastoma microvesicles can be internalized by brain endothelial cells.

[00158] Similar results were obtained when adding the fluorescently labeled microvesicles to primary glioblastoma cells.

Example 18: mRNA delivered by glioblastoma microvesicles can be translated in recipient cells.

[00159] To determine whether glioblastoma-derived microvesicles mRNA could be delivered to and expressed in recipient cells, primary human glioblastoma cells were infected with a self-inactivating lentivirus vector expressing secreted Gaussia luciferase (Gluc) using a CMV promoter at an infection efficiency of >95%. The cells were stably transduced and generated microvesicles during the subsequent passages (2-10 passages were analyzed). Microvesicles were isolated from the cells and purified as described above. RT-PCR analysis showed that the mRNA for Gluc (555 bp) as well as GAPDH (226 bp) were present in the microvesicles (FIG. 17b). The level of *Gluc* mRNA was even higher than that for GAPDH as evaluated with quantitative RT-PCR.

[00160] Fifty micrograms of the purified microvesicles were added to 50,000 HBMVE cells and incubated for 24 hrs. The Gluc activity in the supernatant was measured directly after microvesicle addition (0 hrs), and after 15 hrs and 24 hrs. The Gluc activity in the supernatant was normalized to the Gluc protein activity associated with the microvesicles. The results are presented as the mean ± SEM (n=4). Specifically, the activity in the recipient HBMVE cells demonstrated a continual translation of the microvesicular *Gluc* mRNA. Thus, mRNA incorporated into the tumor microvesicles can be delivered into recipient cells and generate a functional protein.

[00161] The statistical analyses in all examples were performed using the Student's t-test.

References

1. Abravaya, K., J.J. Carrino, S. Muldoon, and H.H. Lee. 1995. Detection of point mutations with a modified ligase chain reaction (Gap-LCR). *Nucleic Acids Res.* 23:675-82.

2. Al-Nedawi, K., B. Meehan, J. Micallef, V. Lhotak, L. May, A. Guha, and J. Rak. 2008. Intercellular transfer of the oncogenic receptor EGFRvIII by microvesicles derived from tumour cells. *Nat Cell Biol.* 10:619-24.

3. Baj-Krzyworzeka, M., R. Szatanek, K. Weglarczyk, J. Baran, B. Urbanowicz, P. Branski, M.Z. Ratajczak, and M. Zembala. 2006. Tumour-derived microvesicles carry several surface determinants and mRNA of tumour cells and transfer some of these determinants to monocytes. *Cancer Immunol Immunother.* 55:808-18.

4. Balzar, M., M.J. Winter, C.J. de Boer, and S.V. Litvinov. 1999. The biology of the 17-1A antigen (Ep-CAM). *J Mol Med.* 77:699-712.

5. Booth, A.M., Y. Fang, J.K. Fallon, J.M. Yang, J.E. Hildreth, and S.J. Gould. 2006. Exosomes and HIV Gag bud from endosome-like domains of the T cell plasma membrane. *J Cell Biol.* 172:923-35.

6. Bossi, A., F. Bonini, A.P. Turner, and S.A. Piletsky. 2007. Molecularly imprinted polymers for the recognition of proteins: the state of the art. *Biosens Bioelectron.* 22:1131-7.

7. Brookes, M.J., N.K. Sharma, C. Tselepis, and T.H. Iqbal. 2005. Serum pro-hepcidin: measuring active hepcidin or a non-functional precursor? *Gut.* 54:169-70.

8. Carmeliet, P., and R.K. Jain. 2000. Angiogenesis in cancer and other diseases. *Nature.* 407:249-57.

9. Carpenter, G. 1987. Receptors for epidermal growth factor and other polypeptide mitogens. *Annu Rev Biochem.* 56:881-914.

10. Chang, H.R., W.H. Kuo, Y.S. Hsieh, S.F. Yang, C.C. Lin, M.L. Lee, J.D. Lian, and S.C. Chu. 2008. Circulating matrix metalloproteinase-2 is associated with cystatin C level, posttransplant duration, and diabetes mellitus in kidney transplant recipients. *Transl Res.* 151:217-23.

11. Chaput, N., J. Taieb, F. Andre, and L. Zitvogel. 2005. The potential of exosomes in immunotherapy. *Expert Opin Biol Ther.* 5:737-47.

12. Cheruvanky, A., H. Zhou, T. Pisitkun, J.B. Kopp, M.A. Knepper, P.S. Yuen, and R.A. Star. 2007. Rapid isolation of urinary exosomal biomarkers using a nanomembrane ultrafiltration concentrator. *Am J Physiol Renal Physiol.* 292:F1657-61.

13. Clayton, A., J.P. Mitchell, J. Court, M.D. Mason, and Z. Tabi. 2007. Human tumor-derived exosomes selectively impair lymphocyte responses to interleukin-2. *Cancer Res.* 67:7458-66.

14. Cotton, R.G., N.R. Rodrigues, and R.D. Campbell. 1988. Reactivity of cytosine and thymine in single-base-pair mismatches with hydroxylamine and osmium tetroxide and its application to the study of mutations. *Proc Natl Acad Sci U S A.* 85:4397-401.

15. Delves, G.H., A.B. Stewart, A.J. Cooper, and B.A. Lwaleed. 2007. Prostasomes, angiogenesis, and tissue factor. *Semin Thromb Hemost.* 33:75-9.

16. Diehl, F., K. Schmidt, M.A. Choti, K. Romans, S. Goodman, M. Li, K. Thornton, N. Agrawal, L. Sokoll, S.A. Szabo, K.W. Kinzler, B. Vogelstein, and L.A. Diaz, Jr. 2008. Circulating mutant DNA to assess tumor dynamics. *Nat Med.* 14:985-90.

17. Fiorentino, F., M.C. Magli, D. Podini, A.P. Ferraretti, A. Nuccitelli, N. Vitale, M. Baldi, and L. Gianaroli. 2003. The minisequencing method: an alternative strategy for preimplantation genetic diagnosis of single gene disorders. *Mol Hum Reprod.* 9:399-410.

18. Fischer, S.G., and L.S. Lerman. 1979a. Length-independent separation of DNA restriction fragments in two-dimensional gel electrophoresis. *Cell.* 16:191-200.

19. Fischer, S.G., and L.S. Lerman. 1979b. Two-dimensional electrophoretic separation of restriction enzyme fragments of DNA. *Methods Enzymol.* 68:183-91.

20. Furnari, F.B., T. Fenton, R.M. Bachoo, A. Mukasa, J.M. Stommel, A. Stegh, W.C. Hahn, K.L. Ligon, D.N. Louis, C. Brennan, L. Chin, R.A. DePinho, and W.K. Cavenee. 2007. Malignant astrocytic glioma: genetics, biology, and paths to treatment. *Genes Dev.* 21:2683-710.

21. Gabrilovich, D.I. 2007. Molecular mechanisms and therapeutic reversal of immune suppression in cancer. *Curr Cancer Drug Targets.* 7:1.

22. Geiss, G.K., R.E. Bumgarner, B. Birditt, T. Dahl, N. Dowidar, D.L. Dunaway, H.P. Fell, S. Ferree, R.D. George, T. Grogan, J.J. James, M. Maysuria, J.D. Mitton, P. Oliveri, J.L. Osborn, T. Peng, A.L. Ratcliffe, P.J. Webster, E.H. Davidson, and L. Hood. 2008. Direct multiplexed measurement of gene expression with color-coded probe pairs. *Nat Biotechnol.* 26:317-25.

23. Goessl, C., H. Krause, M. Muller, R. Heicappell, M. Schrader, J. Sachsinger, and K. Miller. 2000. Fluorescent methylation-specific polymerase chain reaction for DNA-based detection of prostate cancer in bodily fluids. *Cancer Res.* 60:5941-5.

24. Gonzalgo, M.L., M. Nakayama, S.M. Lee, A.M. De Marzo, and W.G. Nelson. 2004. Detection of GSTP1 methylation in prostatic secretions using combinatorial MSP analysis. *Urology.* 63:414-8.

25. Gormally, E., E. Caboux, P. Vineis, and P. Hainaut. 2007. Circulating free DNA in plasma or serum as biomarker of carcinogenesis: practical aspects and biological significance. *Mutat Res.* 635:105-17.

26. Greco, V., M. Hannus, and S. Eaton. 2001. Argosomes: a potential vehicle for the spread of morphogens through epithelia. *Cell.* 106:633-45.

27. Groskopf, J., S.M. Aubin, I.L. Deras, A. Blase, S. Bodrug, C. Clark, S. Brentano, J. Mathis, J. Pham, T. Meyer, M. Cass, P. Hodge, M.L. Macairan, L.S. Marks, and H. Rittenhouse. 2006. APTIMA PCA3 molecular urine test: development of a method to aid in the diagnosis of prostate cancer. *Clin Chem.* 52:1089-95.

28. Guatelli, J.C., K.M. Whitfield, D.Y. Kwoh, K.J. Barringer, D.D. Richman, and T.R. Gingeras. 1990. Isothermal, in vitro amplification of nucleic acids by a multienzyme reaction modeled after retroviral replication. *Proc Natl Acad Sci U S A.* 87:1874-8.

29. Hahn, P.J. 1993. Molecular biology of double-minute chromosomes. *Bioessays.* 15:477-84.

30. Hubner, R.H., S. Meffert, U. Mundt, H. Bottcher, S. Freitag, N.E. El Mokhtari, T. Pufe, S. Hirt, U.R. Folsch, and B. Bewig. 2005. Matrix metalloproteinase-9 in bronchiolitis obliterans syndrome after lung transplantation. *Eur Respir J.* 25:494-501.

31. Janowska-Wieczorek, A., M. Wysoczynski, J. Kijowski, L. Marquez-Curtis, B. Machalinski, J. Ratajczak, and M.Z. Ratajczak. 2005. Microvesicles derived from activated platelets induce metastasis and angiogenesis in lung cancer. *Int J Cancer.* 113:752-60.

32. Johnson, S., D. Evans, S. Laurenson, D. Paul, A.G. Davies, P.K. Ferrigno, and C. Walti. 2008. Surface-immobilized peptide aptamers as probe molecules for protein detection. *Anal Chem.* 80:978-83.

33. Jones, S., X. Zhang, D.W. Parsons, J.C. Lin, R.J. Leary, P. Angenendt, P. Mankoo, H. Carter, H. Kamiyama, A. Jimeno, S.M. Hong, B. Fu, M.T. Lin, E.S. Calhoun, M. Kamiyama, K. Walter, T. Nikolskaya, Y. Nikolsky, J. Hartigan, D.R. Smith, M. Hidalgo, S.D. Leach, A.P. Klein, E.M. Jaffee, M. Goggins, A. Maitra, C. Iacobuzio-Donahue, J.R. Eshleman, S.E. Kern, R.H. Hruban, R. Karchin, N. Papadopoulos, G. Parmigiani, B. Vogelstein, V.E. Velculescu, and K.W. Kinzler. 2008. Core Signaling

Pathways in Human Pancreatic Cancers Revealed by Global Genomic Analyses. *Science*.

34. Kan, Y.W., and A.M. Dozy. 1978a. Antenatal diagnosis of sickle-cell anaemia by D.N.A. analysis of amniotic-fluid cells. *Lancet*. 2:910-2.

35. Kan, Y.W., and A.M. Dozy. 1978b. Polymorphism of DNA sequence adjacent to human beta-globin structural gene: relationship to sickle mutation. *Proc Natl Acad Sci U S A*. 75:5631-5.

36. Keller, S., C. Rupp, A. Stoeck, S. Runz, M. Fogel, S. Lugert, H.D. Hager, M.S. Abdel-Bakky, P. Gutwein, and P. Altevogt. 2007. CD24 is a marker of exosomes secreted into urine and amniotic fluid. *Kidney Int*. 72:1095-102.

37. Kemna, E., P. Pickkers, E. Nemeth, H. van der Hoeven, and D. Swinkels. 2005. Time-course analysis of hepcidin, serum iron, and plasma cytokine levels in humans injected with LPS. *Blood*. 106:1864-6.

38. Kemna, E.H., H. Tjalsma, V.N. Podust, and D.W. Swinkels. 2007. Mass spectrometry-based hepcidin measurements in serum and urine: analytical aspects and clinical implications. *Clin Chem*. 53:620-8.

39. Kemna, E.H., H. Tjalsma, H.L. Willems, and D.W. Swinkels. 2008. Hepcidin: from discovery to differential diagnosis. *Haematologica*. 93:90-7.

40. Kislauskis, E.H., X. Zhu, and R.H. Singer. 1994. Sequences responsible for intracellular localization of beta-actin messenger RNA also affect cell phenotype. *J Cell Biol*. 127:441-51.

41. Kwoh, D.Y., G.R. Davis, K.M. Whitfield, H.L. Chappelle, L.J. DiMichele, and T.R. Gingeras. 1989. Transcription-based amplification system and detection of amplified human immunodeficiency virus type 1 with a bead-based sandwich hybridization format. *Proc Natl Acad Sci U S A*. 86:1173-7.

42. Landegren, U., R. Kaiser, J. Sanders, and L. Hood. 1988. A ligase-mediated gene detection technique. *Science*. 241:1077-80.

43. Li, J., L. Wang, H. Mamon, M.H. Kulke, R. Berbeco, and G.M. Makrigiorgos. 2008. Replacing PCR with COLD-PCR enriches variant DNA sequences and redefines the sensitivity of genetic testing. *Nat Med*. 14:579-84.

44. Liu, C., S. Yu, K. Zinn, J. Wang, L. Zhang, Y. Jia, J.C. Kappes, S. Barnes, R.P. Kimberly, W.E. Grizzle, and H.G. Zhang. 2006a. Murine mammary carcinoma exosomes promote tumor growth by suppression of NK cell function. *J Immunol.* 176:1375-85.

45. Liu, Q., J.C. Greimann, and C.D. Lima. 2006b. Reconstitution, activities, and structure of the eukaryotic RNA exosome. *Cell.* 127:1223-37.

46. Lo, Y.M., N.B. Tsui, R.W. Chiu, T.K. Lau, T.N. Leung, M.M. Heung, A. Gerovassili, Y. Jin, K.H. Nicolaides, C.R. Cantor, and C. Ding. 2007. Plasma placental RNA allelic ratio permits noninvasive prenatal chromosomal aneuploidy detection. *Nat Med.* 13:218-23.

47. Louis, D.N., H. Ohgaki, O.D. Wiestler, W.K. Cavenee, P.C. Burger, A. Jouvet, B.W. Scheithauer, and P. Kleihues. 2007. The 2007 WHO classification of tumours of the central nervous system. *Acta Neuropathol.* 114:97-109.

48. Mack, M., A. Kleinschmidt, H. Bruhl, C. Klier, P.J. Nelson, J. Cihak, J. Plachy, M. Stangassinger, V. Erfle, and D. Schlondorff. 2000. Transfer of the chemokine receptor CCR5 between cells by membrane-derived microparticles: a mechanism for cellular human immunodeficiency virus 1 infection. *Nat Med.* 6:769-75.

49. Mallardo, M., A. Deitinghoff, J. Muller, B. Goetze, P. Macchi, C. Peters, and M.A. Kiebler. 2003. Isolation and characterization of Staufen-containing ribonucleoprotein particles from rat brain. *Proc Natl Acad Sci U S A.* 100:2100-5.

50. Maron, J.L., K.L. Johnson, D. Slonim, C.Q. Lai, M. Ramoni, G. Alterovitz, Z. Jarrah, Z. Yang, and D.W. Bianchi. 2007. Gene expression analysis in pregnant women and their infants identifies unique fetal biomarkers that circulate in maternal blood. *J Clin Invest.* 117:3007-19.

51. Mazzocca, A., R. Coppari, R. De Franco, J.Y. Cho, T.A. Libermann, M. Pinzani, and A. Toker. 2005. A secreted form of ADAM9 promotes carcinoma invasion through tumor-stromal interactions. *Cancer Res.* 65:4728-38.

52. McLendon, R., A. Friedman, D. Bigner, E.G. Van Meir, D.J. Brat, G. Marie Mastrogianakis, J.J. Olson, T. Mikkelsen, N. Lehman, K. Aldape, W.K. Alfred Yung, O. Bogler, S. Vandenberg, M. Berger, M. Prados, D. Muzny, M. Morgan, S. Scherer, A. Sabo, L. Nazareth, L. Lewis, O. Hall, Y. Zhu, Y. Ren, O. Alvi, J. Yao, A. Hawes,

S. Jhangiani, G. Fowler, A. San Lucas, C. Kovar, A. Cree, H. Dinh, J. Santibanez, V. Joshi, M.L. Gonzalez-Garay, C.A. Miller, A. Milosavljevic, L. Donehower, D.A. Wheeler, R.A. Gibbs, K. Cibulskis, C. Sougnez, T. Fennell, S. Mahan, J. Wilkinson, L. Ziaugra, R. Onofrio, T. Bloom, R. Nicol, K. Ardlie, J. Baldwin, S. Gabriel, E.S. Lander, L. Ding, R.S. Fulton, M.D. McLellan, J. Wallis, D.E. Larson, X. Shi, R. Abbott, L. Fulton, K. Chen, D.C. Koboldt, M.C. Wendl, R. Meyer, Y. Tang, L. Lin, J.R. Osborne, B.H. Dunford-Shore, T.L. Miner, K. Delehaunty, C. Markovic, G. Swift, W. Courtney, C. Pohl, S. Abbott, A. Hawkins, S. Leong, C. Haipek, H. Schmidt, M. Wiechert, T. Vickery, S. Scott, D.J. Dooling, A. Chinwalla, G.M. Weinstock, E.R. Mardis, R.K. Wilson, G. Getz, W. Winckler, R.G. Verhaak, M.S. Lawrence, M. O'Kelly, J. Robinson, G. Alexe, R. Beroukhim, S. Carter, D. Chiang, J. Gould, et al. 2008. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature*.

53. Mellinghoff, I.K., M.Y. Wang, I. Vivanco, D.A. Haas-Kogan, S. Zhu, E.Q. Dia, K.V. Lu, K. Yoshimoto, J.H. Huang, D.J. Chute, B.L. Riggs, S. Horvath, L.M. Liau, W.K. Cavenee, P.N. Rao, R. Beroukhim, T.C. Peck, J.C. Lee, W.R. Sellers, D. Stokoe, M. Prados, T.F. Cloughesy, C.L. Sawyers, and P.S. Mischel. 2005. Molecular determinants of the response of glioblastomas to EGFR kinase inhibitors. *N Engl J Med.* 353:2012-24.

54. Miele, E.A., D.R. Mills, and F.R. Kramer. 1983. Autocatalytic replication of a recombinant RNA. *J Mol Biol.* 171:281-95.

55. Millimaggi, D., M. Mari, S. D'Ascenzo, E. Carosa, E.A. Jannini, S. Zucker, G. Carta, A. Pavan, and V. Dolo. 2007. Tumor vesicle-associated CD147 modulates the angiogenic capability of endothelial cells. *Neoplasia.* 9:349-57.

56. Mosse, Y.P., M. Laudenslager, L. Longo, K.A. Cole, A. Wood, E.F. Attiyeh, M.J. Laquaglia, R. Sennett, J.E. Lynch, P. Perri, G. Laureys, F. Speleman, C. Kim, C. Hou, H. Hakonarson, A. Torkamani, N.J. Schork, G.M. Brodeur, G.P. Tonini, E. Rappaport, M. Devoto, and J.M. Maris. 2008. Identification of ALK as a major familial neuroblastoma predisposition gene. *Nature*.

57. Muerkoster, S., K. Wegehenkel, A. Arlt, M. Witt, B. Sipos, M.L. Kruse, T. Sebens, G. Kloppel, H. Kalthoff, U.R. Folsch, and H. Schafer. 2004. Tumor stroma interactions induce chemoresistance in pancreatic ductal carcinoma cells involving increased

secretion and paracrine effects of nitric oxide and interleukin-1beta. *Cancer Res.* 64:1331-7.

58. Myers, R.M., Z. Larin, and T. Maniatis. 1985. Detection of single base substitutions by ribonuclease cleavage at mismatches in RNA:DNA duplexes. *Science*. 230:1242-6.

59. Nagrath, S., L.V. Sequist, S. Maheswaran, D.W. Bell, D. Irimia, L. Ulkus, M.R. Smith, E.L. Kwak, S. Digumarthy, A. Muzikansky, P. Ryan, U.J. Balis, R.G. Tompkins, D.A. Haber, and M. Toner. 2007. Isolation of rare circulating tumour cells in cancer patients by microchip technology. *Nature*. 450:1235-9.

60. Nakanishi, H., J. Groskopf, H.A. Fritsche, V. Bhadkamkar, A. Blase, S.V. Kumar, J.W. Davis, P. Troncoso, H. Rittenhouse, and R.J. Babaian. 2008. PCA3 molecular urine assay correlates with prostate cancer tumor volume: implication in selecting candidates for active surveillance. *J Urol.* 179:1804-9; discussion 1809-10.

61. Nakazawa, H., D. English, P.L. Randell, K. Nakazawa, N. Martel, B.K. Armstrong, and H. Yamasaki. 1994. UV and skin cancer: specific p53 gene mutation in normal skin as a biologically relevant exposure measurement. *Proc Natl Acad Sci U S A.* 91:360-4.

62. Ng, E.K., T.N. Leung, N.B. Tsui, T.K. Lau, N.S. Panesar, R.W. Chiu, and Y.M. Lo. 2003a. The concentration of circulating corticotropin-releasing hormone mRNA in maternal plasma is increased in preeclampsia. *Clin Chem.* 49:727-31.

63. Ng, E.K., N.B. Tsui, T.K. Lau, T.N. Leung, R.W. Chiu, N.S. Panesar, L.C. Lit, K.W. Chan, and Y.M. Lo. 2003b. mRNA of placental origin is readily detectable in maternal plasma. *Proc Natl Acad Sci U S A.* 100:4748-53.

64. Nishikawa, R., T. Sugiyama, Y. Narita, F. Furnari, W.K. Cavenee, and M. Matsutani. 2004. Immunohistochemical analysis of the mutant epidermal growth factor, deltaEGFR, in glioblastoma. *Brain Tumor Pathol.* 21:53-6.

65. Orita, M., H. Iwahana, H. Kanazawa, K. Hayashi, and T. Sekiya. 1989. Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. *Proc Natl Acad Sci U S A.* 86:2766-70.

66. Pan, B.T., and R.M. Johnstone. 1983. Fate of the transferrin receptor during maturation of sheep reticulocytes in vitro: selective externalization of the receptor. *Cell.* 33:967-78.

67. Parsons, D.W., S. Jones, X. Zhang, J.C. Lin, R.J. Leary, P. Angenendt, P. Mankoo, H. Carter, I.M. Siu, G.L. Gallia, A. Olivi, R. McLendon, B.A. Rasheed, S. Keir, T. Nikolskaya, Y. Nikolsky, D.A. Busam, H. Tekleab, L.A. Diaz, Jr., J. Hartigan, D.R. Smith, R.L. Strausberg, S.K. Marie, S.M. Shinjo, H. Yan, G.J. Riggins, D.D. Bigner, R. Karchin, N. Papadopoulos, G. Parmigiani, B. Vogelstein, V.E. Velculescu, and K.W. Kinzler. 2008. An Integrated Genomic Analysis of Human Glioblastoma Multiforme. *Science*.

68. Pelloski, C.E., K.V. Ballman, A.F. Furth, L. Zhang, E. Lin, E.P. Sulman, K. Bhat, J.M. McDonald, W.K. Yung, H. Colman, S.Y. Woo, A.B. Heimberger, D. Suki, M.D. Prados, S.M. Chang, F.G. Barker, 2nd, J.C. Buckner, C.D. James, and K. Aldape. 2007. Epidermal growth factor receptor variant III status defines clinically distinct subtypes of glioblastoma. *J Clin Oncol.* 25:2288-94.

69. Raposo, G., H.W. Nijman, W. Stoorvogel, R. Liejendekker, C.V. Harding, C.J. Melief, and H.J. Geuze. 1996. B lymphocytes secrete antigen-presenting vesicles. *J Exp Med.* 183:1161-72.

70. Roe, M.A., C. Spinks, A.L. Heath, L.J. Harvey, R. Foxall, J. Wimperis, C. Wolf, and S.J. Fairweather-Tait. 2007. Serum prohepcidin concentration: no association with iron absorption in healthy men; and no relationship with iron status in men carrying HFE mutations, hereditary haemochromatosis patients undergoing phlebotomy treatment, or pregnant women. *Br J Nutr.* 97:544-9.

71. Schetter, A.J., S.Y. Leung, J.J. Sohn, K.A. Zanetti, E.D. Bowman, N. Yanaihara, S.T. Yuen, T.L. Chan, D.L. Kwong, G.K. Au, C.G. Liu, G.A. Calin, C.M. Croce, and C.C. Harris. 2008. MicroRNA expression profiles associated with prognosis and therapeutic outcome in colon adenocarcinoma. *JAMA*. 299:425-36.

72. Singer, C.F., D. Gschwantler-Kaulich, A. Fink-Retter, C. Haas, G. Hudelist, K. Czerwenka, and E. Kubista. 2007. Differential gene expression profile in breast cancer-derived stromal fibroblasts. *Breast Cancer Res Treat.*

73. Steemers, F.J., W. Chang, G. Lee, D.L. Barker, R. Shen, and K.L. Gunderson. 2006. Whole-genome genotyping with the single-base extension assay. *Nat Methods.* 3:31-3.

74. Stupp, R., W.P. Mason, M.J. van den Bent, M. Weller, B. Fisher, M.J. Taphoorn, K. Belanger, A.A. Brandes, C. Marosi, U. Bogdahn, J. Curschmann, R.C. Janzer, S.K.

Ludwin, T. Gorlia, A. Allgeier, D. Lacombe, J.G. Cairncross, E. Eisenhauer, and R.O. Mirimanoff. 2005. Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. *N Engl J Med.* 352:987-96.

75. Taylor, D.D., and C. Gercel-Taylor. 2008. MicroRNA signatures of tumor-derived exosomes as diagnostic biomarkers of ovarian cancer. *Gynecol Oncol.* 110:13-21.

76. Thery, C., S. Amigorena, G. Raposo, and A. Clayton. 2006. Isolation and characterization of exosomes from cell culture supernatants and biological fluids. *Curr Protoc Cell Biol.* Chapter 3:Unit 3 22.

77. Thery, C., L. Zitvogel, and S. Amigorena. 2002. Exosomes: composition, biogenesis and function. *Nat Rev Immunol.* 2:569-79.

78. Tomlins, S.A., D.R. Rhodes, S. Perner, S.M. Dhanasekaran, R. Mehra, X.W. Sun, S. Varambally, X. Cao, J. Tchinda, R. Kuefer, C. Lee, J.E. Montie, R.B. Shah, K.J. Pienta, M.A. Rubin, and A.M. Chinnaiyan. 2005. Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. *Science.* 310:644-8.

79. Valadi, H., K. Ekstrom, A. Bossios, M. Sjostrand, J.J. Lee, and J.O. Lotvall. 2007. Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells. *Nat Cell Biol.* 9:654-9.

80. van Dijk, E.L., G. Schilders, and G.J. Pruijn. 2007. Human cell growth requires a functional cytoplasmic exosome, which is involved in various mRNA decay pathways. *RNA.* 13:1027-35.

81. Velculescu, V.E., L. Zhang, B. Vogelstein, and K.W. Kinzler. 1995. Serial analysis of gene expression. *Science.* 270:484-7.

82. Weiss, G., and L.T. Goodnough. 2005. Anemia of chronic disease. *N Engl J Med.* 352:1011-23.

83. Went, P.T., A. Lugli, S. Meier, M. Bundi, M. Mirlacher, G. Sauter, and S. Dirnhofer. 2004. Frequent EpCam protein expression in human carcinomas. *Hum Pathol.* 35:122-8.

84. Wieckowski, E., and T.L. Whiteside. 2006. Human tumor-derived vs dendritic cell-derived exosomes have distinct biologic roles and molecular profiles. *Immunol Res.* 36:247-54.

85. Wong, B.C., R.W. Chiu, N.B. Tsui, K.C. Chan, L.W. Chan, T.K. Lau, T.N. Leung, and Y.M. Lo. 2005. Circulating placental RNA in maternal plasma is associated with a preponderance of 5' mRNA fragments: implications for noninvasive prenatal diagnosis and monitoring. *Clin Chem.* 51:1786-95.

86. Wood, L.D., D.W. Parsons, S. Jones, J. Lin, T. Sjoblom, R.J. Leary, D. Shen, S.M. Boca, T. Barber, J. Ptak, N. Silliman, S. Szabo, Z. Dezso, V. Ustyanksky, T. Nikolskaya, Y. Nikolsky, R. Karchin, P.A. Wilson, J.S. Kaminker, Z. Zhang, R. Croshaw, J. Willis, D. Dawson, M. Shipitsin, J.K. Willson, S. Sukumar, K. Polyak, B.H. Park, C.L. Pethiyagoda, P.V. Pant, D.G. Ballinger, A.B. Sparks, J. Hartigan, D.R. Smith, E. Suh, N. Papadopoulos, P. Buckhaults, S.D. Markowitz, G. Parmigiani, K.W. Kinzler, V.E. Velculescu, and B. Vogelstein. 2007. The genomic landscapes of human breast and colorectal cancers. *Science.* 318:1108-13.

87. Wright, J.L., and P.H. Lange. 2007. Newer potential biomarkers in prostate cancer. *Rev Urol.* 9:207-13.

88. Zehentner, B.K., H. Secrist, X. Zhang, D.C. Hayes, R. Ostenson, G. Goodman, J. Xu, M. Kiviat, N. Kiviat, D.H. Persing, and R.L. Houghton. 2006. Detection of alpha-methylacyl-coenzyme-A racemase transcripts in blood and urine samples of prostate cancer patients. *Mol Diagn Ther.* 10:397-403.

89. Zielie, P.J., J.A. Mobley, R.G. Ebb, Z. Jiang, R.D. Blute, and S.M. Ho. 2004. A novel diagnostic test for prostate cancer emerges from the determination of alpha-methylacyl-coenzyme a racemase in prostatic secretions. *J Urol.* 172:1130-3.

Table 1. RNA in glioblastoma microvesicles can be used as sensitive biomarkers.

Nested RT-PCR was used to monitor EGFRvIII mRNA in glioma biopsy tissue as well as exosomes purified from a frozen serum sample from the same patient. Samples from 30 patients were analysed in a blinded fashion and PCR reactions were repeated at least three times for each sample. No EGFRvIII mRNA was found in serum microvesicles from 30 normal controls. PP1 refers to primer pair composed of SEQ ID NOs: 13 and 14. PP2 refers to primer pair composed of SEQ ID NOS: 15 and 16. "-" refers to "not available".

Patient#	Time of serum collection*	Serum volume	Biopsy EGFRvIII	Serum exosome EGFRvIII(PP1)	Serum exosome EGFRvIII(PP2)
1	0	3 ml	Yes	Yes	-
2	0	2 ml	No	No	-
3	0	2.5 ml	No	No	-
4	0	1 ml	Yes	No	Yes
5	0	1 ml	Yes	No	Yes
6	0	1 ml	No	No	-
7	0	0.6 ml	Yes	Yes	-
8	0	1 ml	No	No	-
9	0	1 ml	Yes	Yes	-
10	0	1 ml	No	Yes	-
11	0	2 ml	Yes	No	Yes
12	0	2 ml	Yes	Yes	-
13	0	2 ml	No	Yes	-
14	0	2 ml	Yes	Yes	-
15	0	2 ml	No	No	-
16	0	2 ml	No	No	-
17	0	1 ml	Yes	No	-
18	0	0.8 ml	Yes	No	-
19	0	1 ml	No	No	-
20	0	1 ml	No	No	-
21	0	1 ml	No	No	-
22	0	1 ml	No	No	-
23	0	1 ml	No	No	-
24	0	1 ml	No	No	-
25	0	1 ml	No	No	-
26	14	0.6 ml	Yes	No	Yes
27	14	1.2 ml	No	No	No
28	14	0.8 ml	Yes	No	Yes
29	14	0.9 ml	Yes	No	No
30	14	0.6 ml	Yes	No	Yes

*Days post-surgery of tumor removal

Table 2 Abbreviations used in Table 3.

Abbreviation	Term
A	amplification
AEL	acute eosinophilic leukemia
AL	acute leukemia
ALCL	anaplastic large-cell lymphoma
ALL	acute lymphocytic leukemia
AML	acute myelogenous leukemia
AML*	acute myelogenous leukemia (primarily treatment associated)
APL	acute promyelocytic leukemia
B-ALL	B-cell acute lymphocyte leukemia
B-CLL	B-cell Lymphocytic leukemia
B-NHL	B-cell Non-Hodgkin Lymphoma
CLL	chronic lymphatic leukemia
CML	chronic myeloid leukemia
CMML	chronic myelomonocytic leukemia
CNS	central nervous system
D	large deletion
DFSP	dermatfibrosarcoma protuberans
DLBL	diffuse large B-cell lymphoma
DLCL	diffuse large-cell lymphoma
Dom	dominant
E	epithelial
F	frames
GIST	gastrointestinal stromal tumour
JMML	juvenile myelomonocytic leukemia
L	leukaemia/lymphoma
M	mesenchymal
MALT	mucosa-associated lymphoid tissue lymphoma
MDS	myelodysplastic syndrome
Mis	Missense
MLCLS	mediastinal large cell lymphoma with sclerosis
MM	multiple myeloma
MPD	Myeloproliferative disorder
N	nonsense
NHL	non-Hodgkin lymphoma
NK/T	natural killer T cell
NSCLC	non small cell lung cancer
O	other
PMBL	primary mediastinal B-cell lymphoma
pre-B All	pre-B-cell acute lymphablastic leukaemia
Rec	reccesive
S	splice site
T	translocation
T-ALL	T-cell acute lymphoblastic leukemia
T-CLL	T-cell chronic lymphocytic leukaemia
TGCT	testicular germ cell tumour
T-PLL	T cell prolymphocytic leukaemia

Table 3: Genes Commonly Mutated in Cancers

Symbol	Locuslink ID	Protein ID*	Chromosome band	Tumour types (somatic)	Tumour types (germline)	Cancer syndrome	Tissue type	Cancer molecular genetics	Mutation type	Translocation partner
ABL1	25	P00519	9q34.1	CML, ALL	–	–	L	Dom	T	*BCR, ETV6*
ABL2	27	P42684	1q24-q25	AML	–	–	L	Dom	T	*ETV6*
AF15Q14	57082	NP_065113	15q14	AML	–	–	L	Dom	T	*MLL*
AF1Q	10962	Q13015	1q21	ALL	–	–	L	Dom	T	*MLL*
AF3p21	51517	Q9NZQ3	3p21	ALL	–	–	L	Dom	T	*MLL*
AF5q31	27125	NP_055238	5q31	ALL	–	–	L	Dom	T	*MLL*
AKT2	208	P31751	19q13.1-q13.2	Ovarian, pancreatic	–	–	E	Dom	A	
ALK	238	Q9UM73	2p23	ALCL	–	–	L	Dom	T	*NPM1, TPM3, TFG, TPM4, ATIC, CLTC, MSN, ALO17*
ALO17	57714	XP_290769	17q25.3	ALCL	–	–	L	Dom	T	*ALK*
APC	324	P25054	5q21	Colorectal, pancreatic, desmoid, hepatoblastoma, glioma, other CNS	Colorectal, pancreatic, desmoid, hepatoblastoma, glioma, other CNS	Adenomatous polyposis coli; Turcot syndrome	E, M, O	Rec	D', Mis, N, F, S	–
ARHGEF1	23365	NP_056128	11q23.3	AML	–	–	L	Dom	T	*MLL*
ARHH	399	Q15669	4p13	NHL	–	–	L	Dom	T	*BCL6*
ARNT	405	P27540	1q21	AML	–	–	L	Dom	T	*ETV6*
ASPSCR1	79058	NP_076988	17q25	Alveolar soft part sarcoma	–	–	M	Dom	T	*TFE3*
ATF1	466	P18846	12q13	Malignant melanoma of soft parts, angiomatoid fibrous histiocytoma	–	–	E, M	Dom	T	*EWSR1*
ATIC	471	P31939	2q35	ALCL	–	–	L	Dom	T	*ALK*
ATM	472	Q13315	11q22.3	T-PLL	Leukaemia, lymphoma, medulloblastoma, glioma	Ataxia telangiectasia	L, O	Rec	D, Mis, N, F, S	–
BCL10	8915	O95999	1p22	MALT	–	–	L	Dom	T	*IGHa*
BCL11A	53335	NP_060484	2p13	B-CLL	–	–	L	Dom	T	*IGHa*

Table 3: Genes Commonly Mutated in Cancers

Symbol	Locuslink ID	Protein ID*	Chromosome band	Tumour types (somatic)	Tumour types (germline)	Cancer syndrome	Tissue type	Cancer molecular genetics	Mutation type	Translocation partner
BCL11B	64919	NP_612808	14q32.1	T-ALL	–	–	L	Dom	T	*TLX3*
BCL2	596	P10415	18q21.3	NHL, CLL	–	–	L	Dom	T	*IGHa*
BCL3	602	P20749	19q13	CLL	–	–	L	Dom	T	*IGHa*
BCL5	603	I52586	17q22	CLL	–	–	L	Dom	T	*MYC*
BCL6	604	P41182	3q27	NHL, CLL	–	–	L	Dom	T, Mis	*IG* loci, *ZNFN1A1, LCP1, PIM1, TFRC, MHC2TA, NACA, HSPCB, HSPCA, HIST1H4I, IL21R, POU2AF1, ARHH, EIF4A2*
BCL7A	605	NP_066273	12q24.1	B-NHL	–	–	L	Dom	T	*MYC*
BCL9	607	O00512	1q21	B-ALL	–	–	L	Dom	T	*IGHa, IGLa*
BCR	613	P11274	22q11.21	CML, ALL	–	–	L	Dom	T	*ABL1, FGFR1*
BHD	201163	NP_659434	17p11.2		Renal, fibrofolliculomas, trichodiscomas	Birt–Hogg–Dube syndrome	E, M	Rec?	Mis, N, F	–
BIRC3	330	Q13489	11q22-q23	MALT	–	–	L	Dom	T	*MALT1*
BLM	641	P54132	15q26.1	–	Leukaemia, lymphoma, skin squamous cell, other cancers	Bloom Syndrome	L, E	Rec	Mis, N, F	–
BMPR1A	657	P36894	10q22.3	–	Gastrointestinal polyps	Juvenile polyposis	E	Rec	Mis, N, F	–
BRAF	673	P15056	7q34	Melanoma, colorectal, papillary thyroid, borderline ovarian, NSCLC, cholangiocarcinoma	–	–	E	Dom	M	–
BRCA1	672	P38398	17q21	Ovarian	Breast, ovarian	Hereditary breast/ovarian	E	Rec	D, Mis, N, F, S	–
BRCA2	675	P51587	13q12	Breast, ovarian, pancreatic	Breast, ovarian, pancreatic, leukaemia (FANCB,	Hereditary breast/	L, E ovarian	Rec	D, Mis, N,	– *F, S*

Table 3: Genes Commonly Mutated in Cancers

Symbol	Locuslink ID	Protein ID*	Chromosome band	Tumour types (somatic)	Tumour types (germline)	Cancer syndrome	Tissue type	Cancer molecular genetics	Mutation type	Translocation partner
BRD4	23476	O60885	19p13.1	Lethal midline carcinoma of young people	FANCD1) –	–	E	Dom	T	*NUT*
BTG1	694	P31607	12q22	BCLL	–	–	L	Dom	T	*MYC*
CBFA2T1	862	Q06455	8q22	AML	–	–	L	Dom	T	*MLL, RUNX1*
CBFA2T3	863	NP_005178	16q24	AML	–	–	L	Dom	T	*RUNX1*
CBFB	865	Q13951	16q22	AML	–	–	L	Dom	T	*MYH11*
CBL	867	P22681	11q23.3	AML	–	–	L	Dom	T	*MLL*
CCND1	595	P24385	11q13	CLL, B-ALL, breast	–	–	L, E	Dom	T	*IGHa, FSTL3*
CDH1	999	P12830	16q22.1	Lobular breast, gastric	Gastric	Familial gastric carcinoma	E	Rec	Mis, N, F, S	–
CDK4	1019	P11802	12q14	–	Melanoma	Familial malignant melanoma	E	Dom	Mis	–
CDKN2A-p14^{ARF}	1029	NP_478102	9p21	Melanoma, multiple other	Melanoma, pancreatic	Familial malignant melanoma	L, E, M, O	Rec	D, S	–
CDKN2A-p16^{INK4A}	1029	P42771	9p21	Melanoma, multiple other	Melanoma, pancreatic	Familial malignant melanoma	L, E, M, O	Rec	D, Mis, N, F, S	–
CDX2	1045	Q99626	13q12.3	AML	–	–	L	Dom	T	*ETV6*
CEBPA	1050	NP_004355	11p15.5	AML, MDS	–	–	L	Dom	Mis, N, F	–
CEP1	11064	NP_008949	9q33	MPD/NHL	–	–	L	Dom	T	*FGFR1*
CHIC2	26511	NP_036242	4q11-q12	AML	–	–	L	Dom	T	*ETV6*
CHN1	1123	P15882	2q31-q32.1	Extraskeletal myxoid chondrosarcoma	–	–	M	Dom	T	*TAF15*
CLTC	1213	Q00610	17q11-qter	ALCL	–	–	L	Dom	T	*ALK*
COL1A1	1277	P02452	17q21.31-q22	Dermatofibrosarcoma protuberans	–	–	M	Dom	T	*PDGFB*
COPEB	1316	Q99612	10p15	Prostatic, glioma	–	–	E, O	Rec	Mis, N	–

Table 3: Genes Commonly Mutated in Cancers

Symbol	Locuslink ID	Protein ID*	Chromosome band	Tumour types (somatic)	Tumour types (germline)	Cancer syndrome	Tissue type	Cancer molecular genetics	Mutation type	Translocation partner
COX6C	1345	P09669	8q22-q23	Uterine leiomyoma	–	–	M	Dom	T	*HMGA2*
CREBBP	1387	Q92793	16p13.3	AL, AML	–	–	L	Dom	T	*MLL, MORF, RUNXBP2*
CTNNB1	1499	P35222	3p22-p21.3	Colorectal, ovarian, hepatoblastoma, others	–	–	E, M, O	Dom	H, Mis	–
CYLD	1540	NP_056062	16q12-q13	Cylindroma	Cylindroma	Familial cylindromatosis	E	Rec	Mis, N, F, S	–
D10S170	8030	NP_005427	10q21	Papillary thyroid, CML	–	–	E	Dom	T	*RET, PDGFRB*
DDB2	1643	Q92466	11p12	–	Skin basal cell, skin squamous cell, melanoma	Xeroderma pigmentosum E	E	Rec	M, N	–
DDIT3	1649	P35638	12q13.1-q13.2	Liposarcoma	–	–	M	Dom	T	*FUS*
DDX10	1662	Q13206	11q22-q23	AML§	–	–	L	Dom	T	*NUP98*
DEK	7913	P35659	6p23	AML	–	–	L	Dom	T	*NUP214*
EGFR	1956	P00533	7p12.3-p12.1	Glioma	–	–	O	Dom	A, O[#]	–
EIF4A2	1974	Q14240	3q27.3	NHL	–	–	L	Dom	T	*BCL6*
ELKS	23085	NP_055879	12p13.3	Papillary thyroid	–	–	E	Dom	T	*RET*
ELL	8178	P55199	19p13.1	AL	–	–	L	Dom	T	*MLL*
EP300	2033	Q09472	22q13	Colorectal, breast, pancreatic, AML	–	–	L, E	Rec	T	*MLL, RUNXBP2*
EP515	2060	P42566	1p32	AL L	–	–	L	Dom	T	*MLL*
ERBB2	2064	P04626	17q21.1	Breast, ovarian, other tumour types	–	–	E	Dom	A	–
ERCC2	2068	P18074	19q13.2-q13.3	–	Skin basal cell, skin squamous cell, melanoma	Xeroderma pigmentosum D	E	Rec	M, N, F, S	–
ERCC3	2071	P19447	2q21	–	Skin basal cell, skin squamous cell, melanoma	Xeroderma pigmentosum B	E	Rec	M, S	–
ERCC4	2072	Q92889	16p13.3-p13.3	–	Skin basal cell, skin squamous cell, melanoma	Xeroderma pigmentosum F	E	Rec	M, N, F	–
ERCC5	2073	P28715	13q33	–	Skin basal cell, skin squamous cell, melanoma	Xeroderma pigmentosum G	E	Rec	M, N, F	–
ERG	2078	P11308	21q22.3	Ewing's sarcoma	–	–	M	Dom	T	*EWSR1*

Table 3: Genes Commonly Mutated in Cancers

Symbol	LocusLink ID	Protein ID*	Chromosome band	Tumour types (somatic)	Tumour types (germline)	Cancer syndrome	Tissue type	Cancer molecular genetics	Mutation type	Translocation partner
ETV1	2115	P50549	7p22	Ewing's sarcoma	–	–	M	Dom	T	*EWSR1*
ETV4	2118	P43268	17q21	Ewing's sarcoma	–	–	M	Dom	T	*EWSR1*
ETV6	2120	P41212	12p13	Congenital fibrosarcoma, multiple leukaemia and lymphoma, secretory breast	–	–	L, E, M	Dom	T	*NTRK3, RUNX1, PDGFRB, ABL1, MN1, ABL2, FACL6, CHIC2, ARNT, JAK2, EVI1, CDX2, STL*
EVI1	2122	Q03112	3q26	AML, CML	–	–	L	Dom	T	*RUNX1, ETV6*
EWSR1	2130	NP_005234	22q12	Ewing's sarcoma, desmoplastic small round cell, ALL	–	–	L, M	Dom	T	*FLI1, ERG, ZNF278, NR4A3, TEC, FEV, ATF1, ETV1, ETV4, WT1, ZNF384*
EXT1	2131	NP_000118	8q24.11-q24.13	–	Exostoses, osteosarcoma	Multiple exostoses type 1	M	Rec	Mis, N, F, S	–
EXT2	2132	Q93063	11p12-p11	–	Exostoses, osteosarcoma	Multiple exostoses type 2	M	Rec	Mis, N, F, S	–
FACL6	23305	NP_056071	5q31	AML, AEL	–	–	L	Dom	T	*ETV6*
FANCA	2175	NP_000126	16q24.3	–	AML, leukaemia	Fanconi anaemia A	L	Rec	D, Mis, N, F, S	–
FANCC	2176	Q00597	9q22.3	–	AML, leukaemia	Fanconi anaemia C	L	Rec	D, Mis, N, F, S	–
FANCD2	2177	NP_149075	3p26	–	AML, leukaemia	Fanconi anaemia D2	L	Rec	D, Mis, N, F	–
FANCE	2178	NP_068741	6p21-p22	–	AML, leukaemia	Fanconi anaemia E	L	Rec	N, F, S	–
FANCF	2188	Q9NPI8	11p15	–	AML, leukaemia	Fanconi anaemia F	L	Rec	N, F	–
FANCG	2189	O15287	9p13	–	AML, leukaemia	Fanconi anaemia G	L	Rec	Mis, N, F, S	–
FEV	54738	NP_059991	2q36	Ewing's sarcoma	–	–	M	Dom	T	*EWSR1*
FGFR1	2260	P11362	8p11.2-p11.1	MPD/NHL	–	–	L	Dom	T	*BCR, FOP, ZNF198, CEP1*
FGFR1OP	11116	NP_008971	6q27	MPD/NHL	–	–	L	Dom	T	*FGFR1*

Table 3: Genes Commonly Mutated in Cancers

Symbol	Locuslink ID	Protein ID*	Chromosome band	Tumour types (somatic)	Tumour types (germline)	Cancer syndrome	Tissue type	Cancer molecular genetics	Mutation type	Translocation partner
FGFR2	2263	P21802	10q26	Gastric	–	–	E	Dom	Mis	–
FGFR3	2261	P22607	4p16.3	Bladder, MM	–	–	L, E	Dom	Mis, T	*IGHα*
FH	2271	P07954	1q42.1	–	Leiomyomatosis, renal	Hereditary leiomyomatosis and renal-cell cancer	E, M	Rec	Mis, N, F	–
FIP1L1	81608	NP_112179	4q12	Idiopathic hypereosinophilic syndrome	–	–	L	Dom	T	*PDGFRA*
FLI	2313	Q01543	11q24	Ewing's sarcoma	–	–	M	Dom	T	*EWSR1*
FLT3	2322	P36888	13q12	AML, ALL	–	–	L	Dom	Mis, O	–
FLT4	2324	P35916	5q35.3	Angiosarcoma	–	–	M	Dom	Mis	–
FNBP1	23048	XP_052666	9q23	AML	–	–	L	Dom	T	*MLL*
FOXO1A	2308	Q12778	13q14.1	Alveolar rhabdomyosarcomas	–	–	M	Dom	T	*PAX3*
FOXO3A	2309	O43524	6q21	AL	–	–	L	Dom	T	*MLL*
FSTL3	10272	O95633	19p13	B-CLL	–	–	L	Dom	T	*CCND1*
FUS	2521	P35637	16p11.2	Liposarcoma	–	–	M	Dom	T	*DDIT3*
GAS7	8522	O60861	17p	AML[y]	–	–	L	Dom	T	*MLL*
GATA1	2623	P15976	Xp11.23	Megakaryoblastic leukaemia of Down syndrome	–	–	L	Dom	Mis, F	–
GMPS	8833	P49915	3q24	AML	–	–	L	Dom	T	*MLL*
GNAS	2778	P04895	20q13.2	Pituitary adenoma	–	–	E	Dom	Mis	–
GOLGA5	9950	NP_005104	14q	Papillary thyroid	–	–	E	Dom	T	*RET*
GPC3	2719	P51654	Xq26.1	–	Wilms' tumour	Simpson–Golabi–Behmel O syndrome	O	Rec	T, D, Mis, N, F, S	–
GPHN	10243	Q9NQX3	14q24	AL	–	–	L	Dom	T	*MLL*
GRAF	23092	NP_055886	5q31	AML, MDS	–	–	L	Dom	T, F, S	*MLL*
HEI10	57820	NP_067001	14q11.1	Uterine leiomyoma	–	–	M	Dom	T	*HMGA2*
HIP1	3092	O00291	7q11.23	CMML	–	–	L	Dom	T	*PDGFRB*
HIST1H4I	8294	NP_003486	6p21.3	NHL	–	–	L	Dom	T	*BCL6*

Table 3: Genes Commonly Mutated in Cancers

Symbol	Locuslink ID	Protein ID*	Chromosome band	Tumour types (somatic)	Tumour types (germline)	Cancer syndrome	Tissue type	Cancer molecular genetics	Mutation type	Translocation partner
HLF	3131	Q16534	17q22	ALL	–	–	L	Dom	T	*TCF3*
HMGA2	8091	P52926	12q15	Lipoma	–	–	M	Dom	T	*LHFP, RAD51L1, LPP, HEI10, COX6C*
HOXA11	3207	P31270	7p15-p14.2	CML	–	–	L	Dom	T	*NUP98*
HOXA13	3209	P31271	7p15-p14.2	AML	–	–	L	Dom	T	*NUP98*
HOXA9	3205	P31269	7p15-p14.2	AML[§]	–	–	L	Dom	T	*NUP98*
HOXC13	3229	P31276	12q13.3	AML	–	–	L	Dom	T	*NUP98*
HOXD11	3237	P31277	2q31-q32	AML	–	–	L	Dom	T	*NUP98*
HOXD13	3239	P35453	2q31-q32	AML[§]	–	–	L	Dom	T	*NUP98*
HRAS	3265	P01112	11p15.5	Infrequent sarcomas, rare other types	–	–	L, M	Dom	Mis	–
HRPT2	3279	NP_01352 2	1q21-q31	Parathyroid adenoma	Parathyroid adenoma, multiple ossifying jaw fibroma	Hyperpara-thyroidism jaw tumour syndrome	E, M	Rec	Mis, N, F	–
HSPCA	3320	P07900	1q21.2-q22	NHL	–	–	L	Dom	T	*BCL6*
HSPCB	3326	P08238	6p12	NHL	–	–	L	Dom	T	*BCL6*
IGHα	3492	–	14q32.33	MM, Burkitt's lymphoma, NHL, CLL, B-ALL, MALT	–	–	L	Dom	T	*MYC, FGFR3, PAX5, IRTA1, IRF4, CCND1, BCL9, BCL6, BCL8, BCL2, BCL3, BCL10, BCL11A, LHX4*
IGKα	50802	–	2p12	Burkitt's lymphoma	–	–	L	Dom	T	*MYC*
IGLα	3535	–	22q11.1-q11.2	Burkitt's lymphoma	–	–	L	Dom	T	*BCL9, MYC*
IL21R	50615	Q9HBE5	16p11	NHL	–	–	L	Dom	T	*BCL6*
IRF4	3662	Q15306	6p25-p23	MM	–	–	L	Dom	T	*IGHα*
IRTA1	83417	NP_11257 2	1q21	B-NHL	–	–	L	Dom	T	*IGHα*
JAK2	3717	O60674	9p24	ALL, AML	–	–	L	Dom	T	*ETV6*
KIT	3815	P10721	4q12	GIST, AML, TGCT	GIST, epithelioma	Familial gastrointestinal stromal	L, M, O	Dom	Mis, O	–
KRAS2	3845	NP_00497 6	12p12.1	Pancreatic, colorectal, thyroid, lung, AML,	–	–	L, E, M, O	Dom	Mis	–

Table 3: Genes Commonly Mutated in Cancers

Symbol	Locuslink ID	Protein ID*	Chromosome band	Tumour types (somatic)	Tumour types (germline)	Cancer syndrome	Tissue type	Cancer molecular genetics	Mutation type	Translocation partner
LAF4	3899	P51826	2q11.2-q12	others	–	–	L	Dom	T	*MLL*
LASP1	3927	Q14847	17q11-q21.3	ALL	–	–	L	Dom	T	*MLL*
LCK	3932	NP_00534 7	1p35-p34.3	AML	–	–	L	Dom	T	*TRBα*
LCP1	3936	P13796	13q14.1-q14.3	T-ALL	–	–	L	Dom	T	*BCL6*
LCX	80312	XP_16761 2	10q21	NHL	–	–	L	Dom	T	*MLL*
LHFP	10186	NP_00577 1	13q12	AML	–	–	M	Dom	T	*HMGA2*
LMO1	4004	P25800	11p15	Lipoma	–	–	L	Dom	T	*TRDα*
LMO2	4005	P25791	11p13	T-ALL	–	–	L	Dom	T	*TRDα*
LPP	4026	NP_00556 9	3q28	T-ALL	–	–	L, M	Dom	T	*HMGA2, MLL*
LYL1	4066	P12980	19p13.2-p13.1	Lipoma, leukaemia	–	–	L	Dom	T	*TRBα*
MADH4	4089	Q13485	18q21.1	T-ALL	–	–	L	Dom	T	*TRBα*
				Colorectal, pancreatic, small intestine	Gastrointestinal polyps	Juvenile polyposis	E	Rec	D, Mis, N, F	–
MALT1	10892	Q9UDY8	18q21	MALT	–	–	L	Dom	T	*BIRC3*
MAML2	84441	XP_04571 6	11q22-q23	Salivary-gland mucoepidermoid	–	–	E	Dom	T	*MECT1*
MAP2K4	6416	P45985	17p11.2	Pancreatic, breast, colorectal	–	–	E	Rec	D, Mis, N	–
MDS1	4197	Q13465	3q26	MDS, AML	–	–	L	Dom	T	*RUNX1*
MECT1	94159	AAK9383	19p13	Salivary-gland mucoepidermoid	–	–	E	Dom	T	*MAML2*
MEN1	4221	O00255	11q13	Parathyroid	Parathyroid adenoma, pituitary adenoma, pancreatic islet cell, carcinoid	Multiple endocrine neoplasia type 1	E	Rec	D, Mis, N, F, S	–
MET	4233	P08581	7q31	Papillary renal, head-neck squamous cell	Papillary renal	Familial papillary renal	E	Dom	Mis	–
MHC2TA	4261	P33076	16p13	NHL	–	–	L	Dom	T	*BCL6*
MLF1	4291	P58340	3q25.1	AML	–	–	L	Dom	T	*NPM1*
MLH1	4292	P40692	3p21.3	Colorectal, endometrial, ovarian, CNS	Colorectal, endometrial, ovarian, CNS	Hereditary non-polyposis colorectal, Turcot syndrome	E, O	Rec	D, Mis, N, F, S	–

Table 3: Genes Commonly Mutated in Cancers

Symbol	Locuslink ID	Protein ID*	Chromosome band	Tumour types (somatic)	Tumour types (germline)	Cancer syndrome	Tissue type	Cancer molecular genetics	Mutation type	Translocation partner
MLL	4297	Q03164	11q23	AML, ALL	–	–	L	Dom	T, O	*MLL, MLLT1, MLLT2, MLLT3, MLLT4, MLLT7, MLLT10, MLLT6, ELL, EPS15, AF1Q, CREBBP, SH3GL1, FNBP1, PNUTL1, MSF, GPHN, GMPS, SSH3BP1, ARHGEF12, GAS7, FOXO3A, LAF4, LCX, SEPT6, LPP, CBFA2T1, GRAF, EP300, PICALM*
MLLT1	4298	Q03111	19p13.3	AL	–	–	L	Dom	T	*MLL*
MLLT10	8028	P55197	10p12	AL	–	–	L	Dom	T	*MLL, PICALM*
MLLT2	4299	P51825	4q21	AL	–	–	L	Dom	T	*MLL*
MLLT3	4300	P42568	9p22	ALL	–	–	L	Dom	T	*MLL*
MLLT4	4301	P55196	6q27	AL	–	–	L	Dom	T	*MLL*
MLLT6	4302	P55198	17q21	AL	–	–	L	Dom	T	*MLL*
MLLT7	4303	NP_00592 9	Xq13.1	AL	–	–	L	Dom	T	*MLL*
MN1	4330	Q10571	22q13	AML_1 meningioma	–	–	L, O	Dom	T	*ETV6*
MSF	10801	NP_00663 1	17q25	AML^3	–	–	L	Dom	T	*MLL*
MSH2	4436	P43246	2p22-p21	Colorectal, endometrial, ovarian	Colorectal, endometrial, ovarian	Hereditary non-polyposis colorectal	E	Rec	D, Mis, N, F, S	–
MSH6	2956	P52701	2p16	Colorectal	Colorectal, endometrial, ovarian	Hereditary non-polyposis colorectal	E	Rec	Mis, N, F, S	–
MSN	4478	P26038	Xq11.2-q12	ALCL	–	–	L	Dom	T	*ALK*
MUTYH	4595	NP_03635 4	1p34.3-1p32.1		Colorectal	–	E	Rec	Mis, N, F, S	–
						Adenomatous polyposis coli				

Table 3: Genes Commonly Mutated in Cancers

Symbol	Locuslink ID	Protein ID*	Chromosome band	Tumour types (somatic)	Tumour types (germline)	Cancer syndrome	Tissue type	Cancer molecular genetics	Mutation type	Translocation partner
MYC	4609	P01106	8q24.12-q24.13	Burkitt's lymphoma, amplified in other cancers, B-CLL	–	–	L, E	Dom	A, T	*IGKα, BCL5, BCL7A, BTG1, TRAα, IGHα*
MYCL1	4610	P12524	1p34.3	Small cell lung	–	–	E	Dom	A	–
MYCN	4613	P04198	2p24.1	Neuroblastoma	–	–	O	Dom	A	–
MYH11	4629	P35749	16p13.13-p13.12	AML	–	–	L	Dom	T	*CBFB*
MYH9	4627	P35579	22q13.1	ALCL	–	–	L	Dom	T	*ALK*
MYST4	23522	NP_036462	10q22	AML	–	–	L	Dom	T	*CREBBP*
NACA	4666	NP_005585	12q23-q24.1	NHL	–	–	L	Dom	T	*BCL6*
NBS1	4683	NP_002476	8q21	–	NHL, glioma, medulloblastoma, rhabdomyosarcoma	Nijmegen breakage syndrome	L, E, M, O	Rec	Mis, N, F	–
NCOA2	10499	Q15596	8q13.1	AML	–	–	L	Dom	T	*RUNXBP2*
NCOA4	8031	Q13772	10q11.2	Papillary thyroid	–	–	E	Dom	T	*RET*
NF1	4763	P21359	17q12	Neurofibroma, glioma	Neurofibroma, glioma	Neurofibromatos is type 1	O	Rec	D, Mis, N, F, S, O	–
NF2	4771	P35240	22q12.2	Meningioma, acoustic neuroma	Meningioma, acoustic neuroma	Neurofibromatos is type 2	O	Rec	D, Mis, N, F, S, O	–
NOTCH1	4851	P46531	9q34.3	T-ALL	–	–	L	Dom	T	*TRBα*
NPM1	4869	P06748	5q35	NHL, APL, AML	–	–	L	Dom	T	*ALK, RARA, MLF1*
NR4A3	8013	Q92570	9q22	Extraskeletal myxoid chondrosarcoma	–	–	M	Dom	T	*EWSR1*
NRAS	4893	P01111	1p13.2	Melanoma, MM, AML, thyroid	–	–	L, E	Dom	Mis	–
NSD1	64324	NP_071900	5q35	AML	–	–	L	Dom	T	*NUP98*
NTRK1	4914	P04629	1q21-q22	Papillary thyroid	–	–	E	Dom	T	*TPM3, TPR, TFG*
NTRK3	4916	Q16288	15q25	Congenital fibrosarcoma, secretory breast	–	–	E, M	Dom	T	*ETV6*
NUMA1	4926	NP_006176	11q13	APL	–	–	L	Dom	T	*RARA*
NUP214	8021	P35658	9q34.1	AML	–	–	L	Dom	T	*DEK, SET*
NUP98	4928	P52948	11p15	AML	–	–	L	Dom	T	*HOXA9, NSD1,*

Table 3: Genes Commonly Mutated in Cancers

Symbol	Locuslink ID	Protein ID*	Chromosome band	Tumour types (somatic)	Tumour types (germline)	Cancer syndrome	Tissue type	Cancer molecular genetics	Mutation type	Translocation partner
NUT	256646	XP_17172 4	15q13	Lethal midline carcinoma of young people	–	–	E	Dom	T	*BRD4*
OLIG2	10215	Q13516	21q22.11	T-ALL	–	–	L	Dom	T	*TRAα*
PAX3	5077	P23760	2q35	Alveolar rhabdomyosarcoma	–	–	M	Dom	T	*FOXO1A*
PAX5	5079	Q02548	9p13	NHL	–	–	L	Dom	T	*IGHα*
PAX7	5081	P23759	1p36.2-p36.12	Alveolar rhabdomyosarcoma	–	–	M	Dom	T	*FOXO1A*
PAX8	7849	Q06710	2q12-q14	Follicular thyroid	–	–	E	Dom	T	*PPARG*
PBX1	5087	NP_00257 6	1q23	Pre-B-ALL	–	–	L	Dom	T	*TCF3*
PCM1	5108	NP_00618 8	8p22-p21.3	Papillary thyroid	–	–	E	Dom	T	*RET*
PDGFB	5155	P01127	22q12.3-q13.1	DFSP	–	–	M	Dom	T	*COL1A1*
PDGFRA	5156	P16234	4q11-q13	GIST	–	–	M, O	Dom	Mis, O	–
PDGFRB	5159	NP_00260 0	5q31-q32	MPD, AML, CMML, CML	–	–	L	Dom	T	*ETV6, TRIP11, HIP1, RAB5EP, H4*
PICALM	8301	Q13492	11q14	T-ALL, AML	–	–	L	Dom	T	*MLLT10, MLL*
PIM1	5292	P11309	6p21.2	NHL	–	–	L	Dom	T	*BCL6*
PML	5371	P29590	15q22	APL	–	–	L	Dom	T	*RARA*
PMS1	5378	P54277	2q31-q33	–	Colorectal, endometrial, ovarian	Hereditary non-polyposis colorectal cancer	E	Rec	Mis, N	–
PMS2	5395	P54278	7p22	–	Colorectal, endometrial, ovarian, medulloblastoma, glioma	Hereditary non-polyposis colorectal cancer, Turcot syndrome	E	Rec	Mis, N, F	–
PMX1	5396	P54821	1q24	AML	–	–	L	Dom	T	*NUP98*
PNUTL1	5413	NP_00267	22q11.2	AML	–	–	L	Dom	T	*MLL*

Table 3: Genes Commonly Mutated in Cancers

Symbol	Locuslink ID	Protein ID*	Chromosome band	Tumour types (somatic)	Tumour types (germline)	Cancer syndrome	Tissue type	Cancer molecular genetics	Mutation type	Translocation partner
POU2AF1	5450	Q16633	11q23.1	NHL	–	–	L	Dom	T	*BCL6*
PPARG	5468	P37231	3p25	Follicular thyroid	–	–	E	Dom	T	*PAX8*
PRCC	5546	Q92733	1q21.1	Papillary renal	–	–	E	Dom	T	*TFE3*
PRKAR1A	5573	P10644	17q23-q24	Papillary thyroid	Myxoma, endocrine, papillary thyroid	Carney complex	E, M	Dom, Rec	T, Mis, N, F, S	*RET*
PRO1073	29005	Q9UHZ2	11q31.1	Renal-cell carcinoma (childhood epithelioid)	–	–	E	Dom	T	*TFEB*
PSIP2	11168	NP_15009 1	9p22.2	AML	–	–	L	Dom	T	*NUP98*
PTCH	5727	Q13635	9q22.3	Skin basal cell, medulloblastoma	Skin basal cell, medulloblastoma	Nevoid basal-cell carcinoma syndrome	E, M	Rec	Mis, N, F, S	–
PTEN	5728	O00633	10q23.3	Glioma, prostatic, endometrial	Harmartoma, glioma, prostatic, endometrial	Cowden syndrome, Bannayan–Riley–Ruvalcaba syndrome	L, E, M, O	Rec	D, Mis, N, F, S	–
PTPN11	5781	Q06124	12q24.1	JMML, AML, MDS	–	–	L	Dom	Mis	–
RABSEP	9135	NP_00469 4	17p13	CMML	–	–	L	Dom	T	*PDGFRB*
RAD51L1	5890	NP_00286 8	14q23-q24.2	Lipoma, uterine leiomyoma	–	–	M	Dom	T	*HMGA2*
RAP1GDS1	5910	P52306	4q21-q25	T-ALL	–	–	L	Dom	T	*NUP98*
RARA	5914	P10276	17q12	APL	–	–	L	Dom	T	*PML, ZNF145, TIF1, NUMA1, NPM1*
RB1	5925	P06400	13q14	Retinoblastoma, sarcoma, breast, small-cell lung	Retinoblastoma, sarcoma, breast, small-cell lung	Familial retinoblastoma	L, E, M, O	Rec	D, Mis, N, F, S	–
RECQL4	9401	O94761	8q24.3	–	Osteosarcoma, skin basal and squamous cell	Rothmund–Thompson syndrome	M	Rec	N, F, S	–
REL	5966	Q04864	2p13-p12	Hodgkin Lymphoma	–	–	L	Dom	A	–
RET	5979	P07949	10q11.2	Medullary thyroid,	Medullary thyroid,	Multiple	E, O	Dom	T, Mis, N,	*H4, PRKAR1A,*

Table 3: Genes Commonly Mutated in Cancers

Symbol	Locuslink ID	Protein ID*	Chromosome band	Tumour types (somatic)	Tumour types (germline)	Cancer syndrome	Tissue type	Cancer molecular genetics	Mutation type	Translocation partner
RPL22	6146	P35268	3q26	AML, CML	–	–	L	Dom	T	*RUNX1*
RUNX1	861	Q01196	21q22.3	AML, pre-B-ALL	–	–	L	Dom	T	*RPL22, MDS1, EVI1, CBFA2T3, CBFA2T1, ETV6*
RUNXBP2	799	NP_006757	8p11	AML	–	–	L	Dom	T	*CREBBP, NCOA2, EP300*
SBDS	51119	Q9Y3A5	7q11	–	AML, MDS	Schwachman–Diamond syndrome	L	Rec	Gene conversion	–
SDHB	6390	P21912	1p36.1-p35	–	Paraganglioma, pheochromocytoma	Familial paraganglioma	O	Rec	Mis, N, F	–
SDHC	6391	O75609	1q21	–	Paraganglioma, pheochromocytoma	Familial paraganglioma	O	Rec	Mis, N, F	–
SDHD	6392	O14521	11q23	–	Paraganglioma, pheochromocytoma	Familial paraganglioma	O	Rec	Mis, N, F, S	–
SEPT6	23157	NP_055944	Xq24	AML	–	–	L	Dom	T	*MLL*
SET	6418	Q01105	9q34	AML	–	–	L	Dom	T	*NUP214*
SFPQ	6421	P23246	1p34.3	Papillary renal cell	–	–	E	Dom	T	*TFE3*
SH3GL1	6455	Q99961	19p13.3	AL	–	–	L	Dom	T	*MLL*
SMARCB1	6598	Q12824	22q11	Malignant rhabdoid	Malignant rhabdoid	Rhabdoid predisposition syndrome	M	Rec	D, N, F, S	–
SMO	6608	Q99835	7q31-q32	Skin basal cell	–	–	E	Dom	Mis	–
SS18	6760	Q15532	18q11.2	Synovial sarcoma	–	–	M	Dom	T	*SSX1, SSX2*
SS18L1	26039	O75177	20q13.3	Synovial sarcoma	–	–	M	Dom	T	*SSX1*
SSH3BP1	10006	NP_005546	10p11.2	AML	–	–	L	Dom	T	*MLL*
SSX1	6756	Q16384	Xp11.23-p11.22	Synovial sarcoma	–	–	M	Dom	T	*SS18*
SSX2	6757	Q16385	Xp11.23-p11.22	Synovial sarcoma	–	–	M	Dom	T	*SS18*
SSX4	6759	O60224	Xp11.23	Synovial sarcoma	–	–	M	Dom	T	*SS18*
STIL	6794	Q15831	19p13.3	NSCLC	Jejunal harmartoma, ovarian, testicular, pancreatic	Peutz–Jeghers syndrome	E, M, O	Rec	D, Mis, N	–
SIL	7955	NOPROT	6q23	B-ALL	–	–	L	Dom	T	*ETV6*

Table 3: Genes Commonly Mutated in Cancers

Symbol	Locuslink ID	Protein ID*	Chromosome band	Tumour types (somatic)	Tumour types (germline)	Cancer syndrome	Tissue type	Cancer molecular genetics	Mutation type	Translocation partner
SUFU	51684	ENSP00000261584 NP_057253	10q24.32	Medulloblastoma	Medulloblastoma	Medulloblastoma predisposition	O	Rec	D, F, S	–
TAF15	8148	Q92804	17q11.1-q11.2	Extraskeletal myxoid chondrosarcomas, ALL	–	–	L, M	Dom	T	*TEC, CHN1, ZNF384*
TAL1	6886	P17542	1p32	Lymphoblastic leukaemia/ biphasic T-ALL	–	–	L	Dom	T	*TRDα*
TAL2	6887	Q16559	9q31	–	–	–	L	Dom	T	*TRBα*
TCF1	6927	P20823	12q24.2	Hepatic adenoma, hepatocellular carcinoma	Hepatic adenoma, hepatocellular carcinoma	Familial hepatic adenoma	E	Rec	Mis, F	–
TCF12	6938	Q99081	15q21	Extraskeletal myxoid chondrosarcoma	–	–	M	Dom	T	*TEC*
TCF3	6929	P15923	19p13.3	pre-B-ALL	–	–	L	Dom	T	*PBX1, HLF, TFPT*
TCL1A	8115	NP_068801	14q32.1	T-CLL	–	–	L	Dom	T	*TRAα*
TEC	7006	P42680	4p12	Extraskeletal myxoid chondrosarcoma	–	–	M	Dom	T	*EWSR1, TAF15, TCF12*
TFE3	7030	P19532	Xp11.22	Papillary renal, alveolar soft part sarcoma	–	–	E	Dom	T	*SFPQ, ASPSCR1, PRCC*
TFEB	7942	P19484	6p21	Renal (childhood epithelioid)	–	–	E, M	Dom	T	*ALPHA*
TFG	10342	NP_006061	3q11-q12	Papillary thyroid, ALCL	–	–	E, L	Dom	T	*NTRK1, ALK*
TFPT	29844	NP_037474	19q13	Pre-B-ALL	–	–	L	Dom	T	*TCF3*
TFRC	7037	P02786	3q29	NHL	–	–	L	Dom	T	*BCL6*
TIF1	8805	O15164	7q32-q34	APL	–	–	L	Dom	T	*RARA*
TLX1	3195	P31314	10q24	T-ALL	–	–	L	Dom	T	*TRBα, TRDα*
TLX3	30012	O43711	5q35.1	T-ALL	–	–	L	Dom	T	*BCL11B*
TNFRSF6	355	P25445	10q24.1	TGCT, nasal NK/T lymphoma, skin squamous-cell carcinoma (burn-	–	–	L, E, O	Rec	Mis	–

Table 3: Genes Commonly Mutated in Cancers

Symbol	Locuslink ID	Protein ID*	Chromosome band	Tumour types (somatic)	Tumour types (germline)	Cancer syndrome	Tissue type	Cancer molecular genetics	Mutation type	Translocation partner
				scar related)						
TOP1	7150	P11387	20q12-q13.1	AML[§]	–	–	L	Dom	T	*NUP98*
TP53	7157	P04637	17p13	Breast, colorectal, lung, sarcoma, adrenocortical, glioma, multiple other types	Breast, sarcoma, adrenocortical carcinoma, glioma, multiple other types	Li–Fraumeni syndrome	L, E, M, O	Rec	Mis, N, F	–
TPM3	7170	P06753	1q22-q23	Papillary thyroid, ALCL	–	–	E, L	Dom	T	*NTRK1, ALK*
TPM4	7171	P07226	19p13.1	ALCL	–	–	L	Dom	T	*ALK*
TPR	7175	P12270	1q25	Papillary thyroid	–	–	E	Dom	T	*NTRK1*
TRA@	6955	–	14q11.2	T-ALL	–	–	L	Dom	T	*ATL, OLIG2, MYC, TCL1A*
TRB@	6957	–	7q35	T-ALL	–	–	L	Dom	T	*HOX11, LCK, NOTCH1, TAL2, LYL1*
TRD@	6964	–	14q11	T-cell leukaemia	–	–	L	Dom	T	*TAL1, HOX11, TLX1, LMO1, LMO2*
TRIM33	51592	Q9UPN9	1p13	Papillary thyroid	–	–	E	Dom	T	*RET*
TRIP11	9321	NP_004230	14q31-q32	AML	–	–	L	Dom	T	*PDGFRB*
TSC1	7248	Q92574	9q34	–	Hamartoma, renal cell	Tuberous sclerosis 1	E, O	Rec	D, Mis, N, F, S	–
TSC2	7249	P49815	16p13.3	–	Hamartoma, renal cell	Tuberous sclerosis 2	E, O	Rec	D, Mis, N, F, S	–
TSHR	7253	P16473	14q31	Toxic thyroid adenoma	Thyroid adenoma	–	E	Dom	Mis	–
VHL	7428	P40337	3p25	Renal, hemangioma, pheochromocytoma	Renal, hemangioma, pheochromocytoma	von Hippel–Lindau syndrome	E, M, O	Rec	D, Mis, N, F, S	–
WAS	7454	P42768	Xp11.23-p11.22	–	Lymphoma	Wiskott–Aldrich syndrome	L	Rec	Mis, N, F, S	–
WHSC1L1	54904	NP_060248	8p12	AML	–	–	L	Dom	T	*NUP98*
WRN	7486	Q14191	8p12-p11.2	–	Osteosarcoma, meningioma, others	Werner syndrome	L, E, M, O	Rec	Mis, N, F, S	–
WT1	7490	NP_000369	11p13	Wilms', desmoplastic small round cell	Wilms'	Denys–Drash syndrome, Frasier	O	Rec	D, Mis, N, F, S	*EWSR1*

Table 3: Genes Commonly Mutated in Cancers

Symbol	Locuslink ID	Protein ID*	Chromosome band	Tumour types (somatic)	Tumour types (germline)	Cancer syndrome	Tissue type	Cancer molecular genetics	Mutation type	Translocation partner
XPA	7507	P23025	9q22.3	..	Skin basal cell, skin squamous cell, melanoma	Xeroderma pigmentosum A	E	Rec	Mis, N, F, S	–
XPC	7508	Q01831	3p25	..	Skin basal cell, skin squamous cell, melanoma	Xeroderma pigmentosum C	E	Rec	Mis, N, F, S	–
ZNF145	7704	Q05516	11q23.1	APL	..	–	L	Dom	T	*RARA*
ZNF198	7750	Q9UBW7	13q11-q12	MPD/NHL	–	..	L	Dom	T	*FGFR1*
ZNF278	23598	NP_055138	22q12-q14	Ewing's sarcoma	–	..	M	Dom	T	*EWSR1*
ZNF384	171017	NP_597733	12p13	ALL	..	–	L	Dom	T	*EWSR1, TAF15*
ZNFN1A1	10320	NP_006051	7p12	ALL, DLBCL	–	–	L	Dom	T	*BCL6*

*From Swiss-Prot/Refseq. ‡D (large deletion) covers the abnormalities that result in allele loss/loss of heterozygosity at many recessive cancer genes. §Refers to cases of acute myeloid leukaemia that are associated with treatment #O (other) in the 'mutation type' column refers primarily to small in-frame deletions/insertions as found in KIT/PDGFRA, and larger duplications/insertions as found in FLT3 and EGFR. Note that where an inversion/large deletion has been shown to result in a fusions protein, these have been listed under translocations. The Wellcome Trust Sanger Institute web version of the cancer-gene set can be found at http://www.sanger.ac.uk/genetics/CPG/Census/. A, amplification; AEL, acute eosinophilic leukaemia; AL, acute leukaemia; ALCL, anaplastic large-cell lymphoma; ALL, acute lymphocytic leukaemia; AML, acute myelogenous leukaemia; APL, acute promyelocytic leukaemia; B-ALL, B-cell acute lymphocytic leukaemia; B-CLL, B-cell lymphocytic leukaemia; B-NHL, B-cell non-Hodgkin's lymphoma; CLL, chronic lymphatic leukaemia; CML, chronic myeloid leukaemia; CMML, chronic myelomonocytic leukaemia; CNS, central nervous system; D, large deletion; DFSP, dermatofibrosarcoma protuberans; DLBCL, diffuse large B-cell lymphoma; Dom, dominant; E, epithelial; F, frameshift; GIST, gastrointestinal stromal tumour; JMML, juvenile myelomonocytic leukaemia; L, leukaemia/lymphoma; M, mesenchymal; MALT, mucosa-associated lymphoid tissue; MDS, myelodysplastic syndrome; MM, multiple myeloma; Mis, missense; N, nonsense; NHL, non-Hodgkin's lymphoma; NK/T, natural killer T cell; NSCLC, non-small-cell lung cancer; O, other; pre-B-ALL, pre-B-cell acute lymphoblastic leukaemia; Rec, recessive; S, splice site; T, translocation; T-ALL, T-cell acute lymphoblastic leukaemia; T-CLL, T-cell chronic lymphocytic leukaemia; TGCT, testicular germ-cell tumour; T-PLL, T-cell prolymphocytic leukaemia.

Table 4: Commonly Upregulated Genes in Cancers

UniGene	Gene symbol	N	Up #	Down #	UniGene	Gene symbol	N	Up #	Down #
Hs. 159430	FNDC3B	11	10	0	Hs. 239388	PAQR8	8	5	1
Hs. 518201	DTX3L	8	7	0	Hs. 592827	RBAK	8	5	1
Hs. 530899	LOC162073	8	7	0	Hs. 525157	TNFSF13B	8	5	1
Hs. 15159	CKLF	11	9	1	Hs. 126774	DTL	13	8	0
Hs. 474150	BID	16	13	0	Hs. 385913	ANP32E	13	8	1
Hs. 7753	CALU	15	12	0	Hs. 532968	DKFP762E1312	13	8	1
Hs. 418795	GLT2SD1	10	8	0	Hs. 372429	PDIA6	13	8	1
Hs 435556	BFAR	12	9	0	Hs. 233952	PSMA7	13	8	1
Hs. 459362	PACI	12	9	1	Hs. 533770	SLC38A1	13	8	1
Hs. 521800	Cborf76	8	6	0	Hs. 489284	ARPC18	18	11	0
Hs. 209561	KIAA1715	8	6	0	Hs. 497788	EPRS	18	11	0
Hs. 585011	Clorf96	8	6	1	Hs. 79110	NCL	18	11	0
Hs. .403933	FBX032	8	6	1	Hs. 251531	PSMA4	18	11	0
Hs. 368853	AYTL2	15	11	1	Hs. 429180	Elf2S2	18	11	1
Hs. 511093	NUSAP1	11	8	0	Hs. 46S885	ILF3	18	11	1
Hs. 370895	RPN2	14	10	0	Hs. 169840	TTK	18	11	1
Hs. 180062	PSMBB	17	12	0	Hs. 489365	APIST	15	9	1
Hs. 444600	BOLAZ	10	7	0	Hs. 256639	PPIH	15	9	1
Hs. 44589O	CHIH4	13	9	0	Hs. 14559	CEP55	10	6	1
Hs. 534392	KDELR3	13	9	0	Hs. 308613	MTERFD1	10	6	1
Hs. 632 191	XTP3TPA	13	9	0	Hs. 21331	ZWILCH	10	6	1
Hs. 387567	ACLV	19	13	1	Hs. 524S99	NAPIL1	17	10	1
Hs. 533282	NONO	18	12	0	Hs. 78171	PGKI	17	10	2
Hs. 83753	SNRPB	18	12	0	Hs. 512380	PLEKHB2	12	7	1
Hs. 471441	PSMBZ	18	12	1	Hs. 352018	TAP1	19	11	1
Hs. 482497	TNPOI	18	12	1	Hs. 194698	CCNB2	14	8	1
Hs. 370937	TAPBP	15	10	0	Hs. 153357	PLOD3	14	8	1
Hs .126941	FAM49B	12	8	0	Hs. 471200	NRP2	14	8	2
Hs. 408629	KDELCI	12	8	0	Hs. 250822	AURKA	16	9	1
Hs. 497384	IPO9	12	8	1	Hs. 75528	GNI2	16	9	1
Hs. 8752	TMEM4	12	8	1	Hs. 1197	HSPEI	16	9	1
Hs. 195642	C17orf27	9	6	0	Hs. 202672	DNMTI	18	10	1
Hs. 358997	TTL	9	6	0	Hs. 433670	FTL	18	10	1
Hs. l600	CCT5	20	13	0	Hs. 519972	HLA-F	18	10	1
Hs. 269408	E2F3	17	11	0	HS. 520210	KDELR2	18	10	1
Hs. 234027	ZBTB12	17	11	1	Hs. 40515.1	CARD-4	11	6	1
Hs. 520205	EIF2AK1	14	9	0	Hs. 477700	DBRI	11	6	1
Hs. 89545	PSMB4	14	9	0	Hs. I4468	FLJ11286	11	6	1
Hs. 449415	EIF2C2	14	9	1	Hs. 516077	FLJ14668	11	6	1
Hs. 409065	FEN1	14	9	1	HS. 494337	GOLPH2	11	6	1
Hs. 313	SPP1	14	9	2	Hs.. 371036	NOX4	11	6	1
Hs. .525135	FARP1	14	9	2	Hs. .438683	SLAMF8	11	6	1
Hs. 524390	K-ALPHA-1	11	7	0	Hs. 520714	SNXIO	11	6	1
Hs. .432360	SCNM1	11	7	0	Hs. 159428	BAX	13	7	1
Hs. 172028	ADAM10	19	12	0	Hs. .311609	DDX39	13	7	1
Hs. 381189	CBX3	19	12	0	Hs. 463035	FKBP10	13	7	1
Hs. 522257	HNRPK	19	12	0	Hs. 438695	FKBP11	13	7	1
Hs. 470943	STATI	19	12	0	Hs. 515255	LSM4	13	7	1
Hs. 118638	NME1	19	12	1	Hs. 55285	MORC2	13	7	1
Hs. 519452	NPM1	19	12	1	Hs. 43666	PTP4A3	13	7	1
Hs. 506748	HDGF	16	10	0	Hs. 369440	SFXN1	13	7	1
Hs. 386283	ADAM12	16	10	2	Hs. 517155	TMEPAI	13	7	1
Hs. 474740	APOL2	8	5	0	Hs. 631580	UBA2	13	7	1
Hs. 552608	C1orf58	8	5	0	Hs. 46346S	UTP16	13	7	1
Hs. 470654	CDCA7	8	5	0	Hs. 492974	WISP1	13	7	1
Hs. 179'B8	FMNL3	8	5	0	Hs. 113876	WHSC1	13	7	1
Hs. 143618	GEMIN6	8	5	0	Hs. 494614	BAT2D1	15	8	2
Hs. 6459	GPRI72A	8	5	0	Hs. 166463	HNRPU	19	10	2
Hs. 133294	IQGAP3	8	5	0					

No number of studies (types of cancer) which have available expression data on a test gene.
Up # or down # number of cancer types whose expression of the tested gene is up or down –regulated.
All these genes are significantly consistently up-regulated (P<10) in a large majority of cancer types.
doi: 10.137/journal pone. 0001149.001

Atty. Docket No. 030258-61532

Table 5: Commonly Downregulated Genes in Cancers

UniGene	Gene symbol	N	Up #	Down #	UniGene	Gene symbol	N	Up #	Down #
Hs. 401835	TCEAL2	10	0	8	Hs. 306083	LOC91689	8	0	5
Hs. 58351	ABCA8	13	0	10	Hs. 160953	PS3AIP1	8	0	5
Hs. 525205	NDRG2	12	0	9	Hs. 2112252	SLC24A3	8	0	5
Hs. 524085	USP2	12	0	9	Hs. 163079	TUBAL3	8	0	5
Hs. 172755	BRP44L	11	0	8	Hs. 389171	PINK1	13	0	8
Hs. 22242	ECHDC3	11	0	8	Hs. 470887	GULP1	13	1	8
Hs. 196952	HLF	19	1	13	Hs. 490981	MSRA	13	1	8
Hs. 496587	CHRDL1	12	0	8	Hs. 476092	CLEC3B	18	0	11
Hs. 476319	ECHDC2	12	0	8	Hs. 386502	FMO4	18	0	11
Hs. 409352	FLJ20701	12	0	8	Hs. 137367	ANK2	18	1	11
Hs. 103253	PLIN	12	0	8	Hs. 212088	EPHX2	18	1	11
Hs. 293970	ALDH6A1	18	1	12	Hs. 157818	KCNAB1	18	1	11
Hs. 390729	ERBB4	17	0	11	Hs. 163924	NR3C2	18	1	11
Hs. 553502	RORA	17	0	11	Hs. 269128	PPP2R1B	18	1	11
Hs. 388918	RECK	14	0	9	Hs. 40582	CDC14B	15	1	9
Hs. 216226	SYNGR1	14	0	9	Hs. 438867	FLJ20489	10	1	6
Hs. 506357	fam107a	14	1	9	Hs. 224008	FEZ1	17	1	10
Hs. 476454	ABHD6	11	0	7	Hs. 443789	C6orf60	12	1	7
Hs. 519694	C5orf4	11	0	7	Hs. 475319	LRRFIP2	12	1	7
Hs. 528385	DHRS4	11	0	7	Hs. 514713	MPPE1	12	1	7
Hs. 477288	TRPM3	1	0	7	Hs. 183153	ARL4D	19	1	11
Hs. 420830	HIF3A	11	1	7	Hs. 642660	C10orf116	19	1	11
Hs. 511265	SEMA6D	11	1	7	Hs. 495912	DMD	19	1	11
Hs. 436657	CLU	19	1	12	Hs. 503126	SHANK2	14	1	8
Hs. 78482	PALM	16	0	10	Hs. 481342	SORBS2	14	1	8
Hs. 82318	WASF3	16	0	10	Hs. 169441	MAGI1	16	1	9
Hs. 268869	ADHFE1	8	0	5	Hs. 75652	GSTM5	18	1	10
Hs. 34494	AGXT2	8	0	5	Hs. 405156	PPAP2B	18	1	10
Hs. 249129	CIDEA	8	0	5	Hs. 271771	SNCA	18	1	10
Hs. 302754	EFCBP1	8	0	5	Hs. 181855	CASC5	9	1	5
Hs. 521953	EFHC2	8	0	5	Hs. 506458	ANKS1B	11	1	6
Hs. 200100	Ells1	8	0	5	Hs. 445885	KIAA1217	11	1	6
Hs. 479703	FLJ21511	8	0	5	Hs. 643583	DKFZp667G2110	13	1	7
Hs. 500750	HPSE2	8	0	5	Hs. 406787	FBXO3	13	1	7
Hs. 380929	LDHD	8	0	5	Hs. 431498	FOXP1	13	1	7

All these genes are significantly consistently down-regulated ($P < 10^{-5}$) in a large majority of cancer types.
doi:10.1371/journal.pone.0001149.t002

Table 6: Commonly Upregulated Genes in Pancreatic Cancer

Accession	Gene Symbol	Gene Name	FC
NM_006475	POSTN	periostin, osteoblast specific factor	13.28
NM_005980	S100P	S100 calcium binding protein P	12.36
NM 004385	CSPG2	chondroitin sulfate proteoglycan 2 (versican)	10.57
NM 003118	SPARC	secreted protein, acidic cysteine-rich (osteonectin)	10.46
NM 003225	TFF1	trefoil factor 1 (breast cancer, estrogen-inducible sequence expressed in)	8.13
NM 002026	FN1	fibronectin 1	7.93
NM 006142	SFN	stratifin	7.81
NM 000393	COL5A2	collagen, type V, alpha 2	7.22
NM 005940	MMP11	matrix metalloproteinase 11 (stromelysin 3)	7.17
NM 000088	COL1A1	collagen, type I, alpha 1	6.50
NM 000930	PLAT	plasminogen activator, tissue	6.46
NM 003064	SLPI	secretory leukocyte protease inhibitor (antileukoproteinase)	6.01
NM 006516	SLC2A1	solute carrier family 2 (facilitated glucose transporter), member 1	5.39
NM 003226	TFF3	trefoil factor 3 (intestinal)	5.28
NM 004460	FAP	fibroblast activation protein alpha	5.20
NM 003467	CXCR4	chemokine (C-X-C motif) receptor 4	5.18
NM 003247	THBS2	thrombospondin 2	5.04
NM 012101	TRIM29	tripartite motif-containing	4.91
NM 033664	CDH11	cadherin 11, type 2, OB-cadherin (osteoblast)	4.52
NM 006169	NNMT	nicotinamide N-methyltransferase	4.51
NM 004425	ECM1	extracellular matrix protein 1	4.39
NM 003358	UGCG	UDP-glucose ceramide glucosyltransferase	4.36
NM 000700	ANXA1	annexin A1	4.31
NM 004772	C5orf13	chromosome 5 open reading frame 13	4.29
NM 182470	PKM2	pyruvate kinase, muscle	4.28
NM 004994	MMP9	matrix metalloproteinase 9 (gelatinase B, 92kDa gelatinase, 92kDa type IV collagenase)	4.19
NM 006868	RAB31	RAB31, member RAS oncogene family	4.18
NM 001932	MPP3	membrane protein, palmitoylated 3 (MAGUK p55 subfamily member 3)	4.16
AF200348	D2S448	Melanoma associated gene	4.14
NM 000574	DAF	decay accelerating factor for complement (CD55, Cromer blood group system)	4.11
NM 000213	ITGB4	integrin beta	4.11
NM 001645	APOC1	apolipoprotein C-I	3.86
NM 198129	LAMA3	laminin, alpha 3	3.86
NM 002997	SDC1	syndecan 1	3.80
NM 001769	CD9	CD9 antigen (p24)	3.78
BC004376	ANXA8	annexim A8	3.74
NM 005620	S100A11	S100 calcium binding protein A11 (calgizzarin)	3.72
NM 002659	PLAUR	plasminogen activator urokinase receptor	3.70
NM 002966	S100A10	S100 calcium binding protein A10 (annexin II ligand, calpactin I, light polypeptide (p11))	3.67
NM 004898	CLOCK	clock homolog (mouse)	3.65
NM 002345	LUM	lumican	3.59
NM 006097	MYL9	myosin light polypeptide 9, regulatory	3.44
NM 004120	GBP2	guanylate binding protein 2, interferon-inducible	3.44
AK056875	LOC91316	similar to bK246H3.1 (immunoglobulin lambda-like polypeptide 1, pre-B-cell specific)	3.40
NM 001827	CKS2	CDC28 protein kinase regulatory subunit 2	3.36
NM 002203	ITGA2	integrin alpha 2 (CD49B, alpha 2 subunit of VLA-2 receptor)	3.35
NM 000599	IGFBP5	insulin-like growth factor binding protein 5	3.33
NM 004530	MMP2	matrix metalloproteinase 2 (gelatinase A, 72kDa gelatinase, 72kDa type IV collagenase)	3.33
NM 004335	BST2	bone marrow stromal cell antigen	3.30
NM 000593	TAP1	transporter 1, ATP-binding cassette, sub-family B (MDR/TAP)	3.29
NM 004915	ABCG1	ATP-binding cassette sub-family G (WHITE), member	3.27
NM 001235	SERPINH 1	serine (or cysteine) proteinase inhibitor, clade H (heat shock protein 47), member 1 (collagen binding protein 1)	3.25
NM 001165	BIRC3	baculoviral IAP repeat-containing 3	3.23
NM 002658	PLAU	plasminogen activator, urokinase	3.20
NM 021103	TMSB10	thymosin, beta 10	3.18
NM 000304	PMP22	peripheral myelin protein 22	3.15
XM 371541	KIAA1641	KIAA1641 protein	3.11
NM 012329	MMD	monocyte to macrophage differentiation-associated	3.07
NM 182744	NBL1	neuroblastoma suppression of tumorigenicity 1	3.06
NM 002245	KCNK1	potassium channel, subfamily K, member 1	3.03
NM 000627	LTBP1	latent transforming growth factor beta binding protein 1	3.02
NM 000063	C2	complement component 2	3.01
NM 000100	CSTB	cystatin B (stefin B)	2.99
NM 000396	CTSK	cathepsin K (pycnodysostosis)	2.98

NM 016816	OAS1	2' 5'-oligoadenylate synthetase 1, 40/46kDa	2.98
NM 004240	TRIP10	thyroid hormone receptor interactor 10	2.95
NM 000138	FBN1	fibrillin 1 (Marfan syndrome)	2.94
NM 002318	LOXL2	lysyl oxidase-like 2	2.92
NM 002053	GBP1	guanylate binding protein 1 interferon-inducible, lysyl 67kDa	2.90
NM 005564	LCN2	lipocalin 2 (oncogene 24p3)	2.88
NM 153490	KRT13	keratin 13	2.85
NM 004723	ARHGEF 2	rho/rac guanine nucleotide exchange factor (GEF) 2	2.80
NM 004146	NDUFB7	NADH dehydrozenase (ubiquinone) 1 beta subcomplex, 7, 18kDa	2.79
NM 003937	KYNU	kynureninase (L-kynurenine hydrolase)	2.77
NM 002574	PRDX1	Peroxiredoxin 1	2.77
NM 002444	MSN	moesin	2.73
NM 002901	RCN1	reticulocalbin 1, EF-hand calcium binding domain	2.73
NM 005165	ALDOC	aldolase C, fructose-bisphosphate	2.72
NM 002204	ITGA3	integrin, alpha 3 (antigen CD49C, alpha 3 subunit of VLA-3 receptor)	2.72
NM 033138	CALD1	caldesmon 1	2.71
NM 003816	ADAM9	a disintegrin and metalloproteinase domain 9 (meltrin gamma)	2.69
NM 173843	IL1RN	interleukin 1 receptor antagonist	2.66
NM 000602	SERPINE 1	serine (or cysteine) proteinase inhibitor, clade E (nexin, plasminggen activator inhibitor type 1), member 1	2.65
NM 002213	ITGB5	integrin, beta 5	2.64
NM 004447	EPS8	epidermal growth factor receptor pathway substrate 8	2.64
NM 002928	RGS16	regulator of G-protein singalling 16	2.62
NM 001288	CLIC1	chloride intracellular channel 1	2.61
NM 015996	TAGLN	transgelin	2.57
NM 002087	GRN	granulin	2.55
NM 001183	ATP6AP1	ATPase, H+ transporting, lysosomal accessory protein 1	2.54
NM 001730	KLF5	Kruppel-like factor 5 (intestinal)	2.51
NM 003516	HIST2H2 AA	histone 2, H2aa	2.50
NM 014736	KIAA0101	KIAA0101 gene product	2.49
NM 002290	LAMA4	laminin, alpha 4	2.49
NM 001826	CKS1B	CDC28 protein kinase reaulatory subunit 1B	2.48
NM 001814	CTSC	cathepsin C	2.45
NM 176825	SULT1C1	sulfotransferase family cytosolic, 1C, member 1	2.43
NM 002862	PYGB	phosphorylase, glycogen; brain	2.41
NM 000917	P4HA1	procollagen-proline, 2-oxoglutarate 4-dioxygenase (proline 4- hydroxylase), alpha polypeptideI	2.41
NM 001428	EN01	enolase 1 (alpha)	2.40
NM 001425	EMP3	epithelial membrane protein 3	2.40
NM 019111	HLA-DRA	maior histocompatibility complex, class II, DR alpha	2.38
NM 001387	DPYSL3	dihydropyrimidinase-like 3	2.36
NM 006471	MRCL3	myosin regulatory light chain MRCL3	2.34
NM 006332	IFI30	interferon gamma-inducible protein 30	2.34
NM 001312	CRIP2	cysteine-rich protein 2	2.33
NM 002224	ITPR3	inositol 1 4 5-triphosphate receptor tvpe 3	2.31
NM 053025	MYLK	myosin light peptide kinase	2.29
NM 002785	PSG11	pregnancy specific beta-1-glycoprotein 11	2.27
NM 000900	MGP	matrix Gla protein	2.26
NM 000962	PTGS1	prostaglandin-endoperoxide synthase 1 (prostaglandin G/H synthase and cyclooxyenase)	2.25
NM 005915	MCM6	minichromosome maintenance deficient 6 (MIS5 homolog, S. pombe) (S. cerevisiae)	2.24
NM 001067	TOP2A	topoisomerase (DNA) II alpha 170kDa	2.23
NM 001878	CRABP2	cellular retinoic acid binding protein 2	2.23
NM 006745	SC4MOL	sterol-C4-methyl oxidase-like	2.22
NM 003528	HIST2H2	histone 2, H2be	2.22
BF347579		Transcribed sequence with strong similarity to protein pir:I38500 (H.sapiens) I38500 interferon gamma receptor accessory factor-1 precursor - human	2.21
NM 005261	GEM	GTP binding protein overexpressed in skeletal muscle	2.19
NM 021874	CDC25B	cell division cycle 25B	2.18
NM 022550	XRCC4	X-ray repair complementing defective repair in Chinese hamster cells 4	2.17
NM 020250	GSN	gelsolin (amyloidosis, Finnish type)	2.17
NM 002916	RFC4	replication factor C (activator 1) 4, 37kDa	2.16
NM 005606	LGMN	legumain	2.14
NM 006762	LAPTM5	Lysosomal-associated multispanning membrane protein-5	2.14
NM 002727	PRG1	proteoglycan 1, secretory granule	2.14
NM 002609	PDGFRB	platelet-derived growth factor receptor, beta polypeptide	2.14
NM 001424	EMP2	epithelial membrane protein 2	2.12
NM_005022	PFN1	profilin 1	2.12
NM_001657	AREG	amphiregulin amphireaulin (schwannoma-derived growth factor)	2.11
NM_005100	AKAP12	A kinase (PRKA) anchor protein (gravin) 12	2.11
NM_000860	HPGD	hydroxyprostaglandin dehydrogenase 15 (NAD)	2.10

NM_007115	TNFAIP6	tumor necrosis factor alpha-induced protein 6	2.09
NM_021638	AFAP	actin filament associated protein	2.08
NM_001946	DUSP6	dual specificity phosphatase 6	2.05
NM_181802	UBE2C	ubiquitin-conjugating enzyme E2C	2.04
NM_002593	PCOLCE	procollagen C-endopeptidase enhancer	2.02
NM_033292	CASP1	caspase 1, apoptosis-related cysteine protease (interleukin 1, beta, convertase)	2.02
NM_003870	IQGAP1	IQ motif containing GTPase activating protein 1	2.02
NM_005563	STMN1	stathmin 1/oncoprotein 18	2.01
NM_005558	LAD1	ladinin 1	2.01
NM_001776	ENTPD1	ectonucleoside triphosphate diphosphohydrolase 1	2.00
NM_001299	CNN1	calponin 1, basic, smooth muscle	2.00
AK055128	PSMD14	proteasome (prosome, macropain) 26S subunit, non-ATPase, 14	2.00
NM_006304	SHFM1	split hand/foot malformation (ectrodactyly) type 1	1.98
NM_004024	ATF3	activating transcription factor 3	1.98
NM_000291	PGK1	phosphoglycerate kinase 1	1.98
NM_006520	TCTE1L	t-complex-associated-testis-expressed 1-like	1.97
NM_201380	PLEC1	plectin 1 intermediate filament binding protein 500kDa	1.97
NM_002838	PTPRC	protein tyrosine phosphatase, receptor type, C	1.97
NM_000211	ITGB2	integrin, beta 2 (antigen CD18 (p95), lymphocyte function-associated antigen 1; macrophage antigen 1 (mac-1) beta subunit)	1.97
NM_002577	PAK2	p21 (CDKN1A)-activated kinase 2	1.96
NM_000295	SERPINA 1	serine (or cysteine) proteinase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 1	1.96
NM_183001	SHC1	SHC (Src homology 2 domain containing) transforming protein 1	1.96
NM_005019	PDE1A	phosphodiesterase 1A, calmodulin-dependent	1.95
NM_002298	LCP1	lymphocyte cytosolic protein 1 (L-plastin)	1.95
NM_006769	LMO4	LIM domain only 4	1.94
NM_001465	FYB	FYN binding protein (FYB-120/130)	1.93
NM_183422	TSC22	transforming growth factor beta-stimulated protein TSC-22	1.92
NM_001777	CD47	CD47 antigen (Rh-related antigen, integrin-associated signal transducer)	1.92
NM_001755	CBFB	core-binding factor, beta subunit	1.90
NM_005544	IRS1	insulin receptor substrate 1	1.88
NM_000698	ALOX5	arachidonate 5-lipoxygenase	1.88
NM_006096	NDRG1	N-myc downstream regulated gene 1	1.88
NM_001105	ACVR1	activin A receptor, type 1	1.87
NM_003105	SORL1	sortilin-related receptor, L(DLR class) A repeats-containing	1.85
NM_001998	FBLN2	fibulin 2	1.85
NM_014791	MELK	maternal embryonic leucine zipper kinase	1.85
NM_003092	SNRPB2	small nuclear ribonucleoprotein polypeptide B	1.84
NM_001120	TETRAN	tetracycline transporter-like protein	1.84
NM_182943	PLOD2	procollagen-lysine, 2-oxoglutarate 5-dioxygenase (lysine hydroxylase) 2	1.83
NM_181862	BACH	brain acyl-CoA hydrolase	1.82
NM_021102	SPINT2	serine protease inhibitor, Kunitz type, 2	1.82
NM_004419	DUSP5	dual specificity phosphatase 5	1.81
NM_006482	DYRK2	dual specificity tyrosine-(Y)-phosphorylation regulated kinase 2	1.81
NM_145690	YWHAZ	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, zeta polypeptide	1.81
NM_000714	BZRP	benzodiazapine receptor (peripheral)	1.81
NM_013995	LAMP2	lysosomal-associated membrane protein 2	1.80
CA450153	ACYP1	acylphosphatase 1, erythrocyte (common) type	1.80
NM_000405	GM2A	GM2 ganglioside activator protein	1.79
NM_139275	AKAP1	A kinase (PRKA) anchor protein 1	1.79
NM_001679	ATP1B3	ATPase, Na+/K+ transporting, beta 3 polypeptide	1.79
NM_016343	CENPF	centromere protein F, 350/400ka (mitosin)	1.79
NM_002201	ISG20	interferon stimulated gene 20kDa	1.79
NM_002463	MX2	myxovirus (influenza virus) resistance 2 (mouse)	1.79
NM_006820	C1orf29	chromosome 1 open reading frame 29	1.79
NM_201397	GPX1	glutathione peroxidase 1	1.78
NM_005738	ARL4	ADP-ribosylation factor-like 4	1.78
NM_001038	SCNN1A	sodium channel nonvoltage-gated 1 alpha	1.78
NM_002863	PYGL	phosphorylase, glycogen; liver (Hers disease, glycogen storage disease type VI)	1.77
NM_001281	CKAP1	cytoskeleton associated protein 1	1.76
NM_003879	CFLAR	CASP8 and FADD-like apoptosis regulator	1.75
NM_182948	PRKACB	protein kinase, cAMP-dependent catalytic, beta	1.75
NM_006009	TUBA3	tubulin, alpha 3	1.74
NM_201444	DGKA	diacylglycerol kinase, alpha 80kDa	1.74
NM_005471	GNPDA1	glucosamine-6-phosphate deaminase 1	1.74
NM_001451	FOXF1	forkhead box F1	1.73
NM_001988	EVPL	envoplakin	1.73
NM_021724	NR1D1	nuclear receptor subfamily 1, group D member 1	1.72
NM_006364	SEC23A	Sec23 homolog A (S. cerevisiae)	1.72
NM_002129	HMGB2	high-mobility group box 2	1.72

NM_004172	SLC1A3	solute carrier family 1 (glial high affinity glutamate transporter), member 3	1.71
NM_001421	ELF4	E74-like factor 4 (ets domain transcription factor)	1.71
NM_005566	LDHA	lactate dehydrogenase A	1.70
NM_000270	NP	nucleoside phosphorylase	1.69
NM_153425	TRADD	TNFRSF1A-associated via death domain	1.67
NM_004762	PSCD1	pleckstrin homology, Sec7 and coiled-coil domains (cytohesin 1)	1.67
NM_001985	ETFB	electron-transfer-flavoprotein, beta polypeptide	1.67
NM_016587	CBX3	chromobox homolog 3 (HP1 gamma homolog, Drosophila)	1.66
NM_002085	GPX4	glutathione peroxidase 4 (phospholipid hydroperoxidase)	1.66
NM_002795	PSMB3	proteasome (prosome, macropain) subunit, beta type, 3	1.65
NM_000963	PTGS2	prostaglandin-endoperoxide synthase 2 (prostaglandin G/H synthase and cyclopxyoenase)	1.65
NM_001642	APLP2	amyloid beta (A4) precursor-like protein 2	1.65
NM_000569	FCGR3A	Fc fragment of IgG low affinity iiia receptor for (CD16)	1.64
NM_000362	TIMP3	tissue inhibitor of metalloproteinase 3 (Sorsby fundus dystrophy, pseudoinflammatory)	1.63
NM_002417	MKI67	antigen identified by monoclonal antibody Ki-67	1.63
NM_000175	GPI	glucose phosophate isomerase	1.63
AF179995	SEPT8	septin 8	1.63
NM_004121	GGTLA1	gamma-glutamyltransferase-like activity 1	1.62
NM_002690	POLB	polymerase (DNA directed), beta	1.62
NM_004334	BST1	bone marrow stromal cell antigen 1	1.62
NM_001892	CSNK1A1	casein kinase 1, alpha 1	1.61
NM_014670	BZW1	basic leucine zipper and W2 domains 1	1.61
NM_001110	ADAM10	a disintegrin and metalloproteinase domain 10	1.60
NM_005792	MPHOSP H6	M-phase phosphoprotein 6	1.60
NM_001126	ADSS	adenylosuccinate synthase	1.60
XM 376059	SERTAD2	SERTA domain containing 2	1.59
NM_001664	ARHA	ras homolog gene family, member A	1.59
NM_002475	MLC1SA	myosin light chain 1 slow a	1.59
NM_014498	GOLPH4	golgi phosphoprotein 4	1.59
NM_005964	MYH10	myosin heavy polypeptide 10 non-muscle	1.59
NM_003330	TXNRD1	thioredoxin reductase 1	1.59
NM_001757	CBR1	carbonyl reductase 1	1.58
NM_003130	SRI	sorcin	1.57
NM_006765	TUSC3	tumor suppressor candidate 3	1.57
NM_183047	PRKCBP 1	protein kinase C binding protein 1	1.57
NM_005333	HCCS	holocytochrome c synthase (cytochrome c heme-lyase)	1.57
NM_001444	FABP5	fatty acid binding protein 5 (psoriasis-associated)	1.57
NM_001799	CDK7	cyclin-dependent kinase 7 (M015 homolog, Xenopus laevis, cdk-activating kinase)	1.57
NM_001539	DNAJA1	DnaJ (Hsp40) homolog subfamily A member 1	1.57
NM_004475	FLOT2	flotillin 2	1.57
NM_004308	ARHGAP 1	Rho GTPase activating protein 1	1.56
NM_002388	MCM3	MCM3 minichromosome maintenance deficient 3 (S. cerevisiae)	1.56
NM_006435	IFITM2	interferon induced transmembrane protein 2 (1-8D)	1.56
NM_000454	SOD1	superoxide dismutase 1, soluble (amyotrophic lateral sclerosis 1 (adult))	1.56
NM_015161	ARL6IP	ADP-ribosylation factor-like 6 interacting protein	1.56
NM_078480	SIAHBP1	fuse-binding protein-interacting repressor	1.56
NM_025207	PP591	FAD-synthetase	1.56
NM_002833	PTPN9	protein tyrosine phosphatase non-receptor type 9	1.55
NM_001753	CAV1	caveolin 1 caveolae protein 22kDa	1.55
NM_003286	TOP1	topoisomerase (DNA) I	1.55
BU739663		Transcribed sequence with moderate similarity to protein sp:P13196 (H.sapiens) HEM1_HUMAN 5-aminolevulinic acid synthase, nonspecific mitochondrial precursor	1.55
NM_006788	RALBP1	ralA binding protein 1	1.54
NM_000944	PPP3CA	protein phosphatase 3 (formerly 2B), catalytic subunit, alpha isoform (calcineurin A alpha)	1.54
NM_003374	VDAC1	voltage-dependent anion channel 1	1.54
NM_000560	CD53	CD53 antigen	1.54
NM_002037	FYN	FYN oncogene related to SRC FGR, YES	1.54
NM_002885	RAP1GA1	RAP1 GTPase activating protein 1	1.53
NM_018979	PRKWNK 1	lprotein kinase, lysine deficient 1	1.53
NM_002835	PTPN12	protein tyrosine phosphatase, non-receptor type 12	1.53
NM_007315	STAT1	signal transducer and activator of transcription 1, 91kDa	1.52
NM_014846	KIAA0196	KIAA0196 gene product	1.52
NM_001237	CCNA2	cyclin A2	1.52
NM_004596	SNRPA	small nuclear ribonucleoprotein polypeptide A	1.52
NM_002790	PSMA5	proteasome (prosome, macropoain) subunit, alpha type, 5	1.52
NM_015361	R3HDM	R3H domain (binds single-stranded nucleic acids) containing	1.52
NM_001665	ARHG	ras homolog gene family, member G (rho G)	1.51
NM_002788	PSMA3	proteasome (prosome macropain) subunit, alpha type, 3	1.50
NM_006904	PRKDC	protein kinase, DNA-activated, catalytic polypeptide	1.50
NM_003400	XPO1	exportin 1 (CRM1 homolog, yeast)	1.50

NM_178014	OK/SW-cl.56	beta 5-tubulin	1.50
NM_002634	PHB	prohibitin	1.49
NM_004792	PPIG	peptidyl-prolyl isomerase G (cyclophilin G)	1.49
NM_002508	NID	nidogen (enactin)	1.49
NM_001765	CD1C	CD1C antigen, c polypeptide	1.48
NM_000311	PRNP	prion protein (p27-30) (Creutzfeld-Jakob disease, Gerstmann-Strausler-Scheinker syndrome, fatal familial insomnia)	1.48
NM_006437	ADPRTL1	ADP-ribosyltransferase (NAD+; poly (ADP-ribose) polymerase)-like 1	1.48
NM_002759	PRKR	protein kinase, interferon-inducible double stranded RNA dependent	1.48
NM_014669	KIAA0095	KIAA0095 gene product	1.47
NM_003391	WNT2	wingless-type MMTV integration site family member 2	1.47
NM_004309	ARHGDIA	Rho GDP dissociation inhibitor (GDI) alpha	1.47
NM_000418	IL4R	interleukin 4 receptor	1.46
NM_003352	UBL1	ubiquitin-like 1 (sentrin)	1.46
NM_006290	TNFAIP3	tumor necrosis factor alpha-induced protein 3	1.45
NM_004763	ITGB1BP1	integrin beta 1 binding protein 1	1.45
NM_005754	G3BP	Ras-GTPase-activating protein SH3-domain-binding protein	1.45
NM_021990	GABRE	gamma-aminobutyric acid (GABA) A receptor, epsilon	1.44
NM_001379	DNMT1	DNA (cytosine-5-)-methyltransferase 1	1.44
NM_001154	ANXA5	annexin A5	1.44
NM_004354	CCNG2	cyclin G2	1.44
NM_005002	NDUFA9	NADH dehydroaenase (ubiquinone) 1 alpha subcomplex, 9, 39kDa	1.43
NM_001931	DLAT	dihydrolipoamide S-acetyltransferase (E2 component of pyruvate dehydroaenase complex)	1.43
NM_005902	MADH3	MAD mothers against decapentaplegic homolog 3 (Drosophila)	1.43
NM_000110	DPYD	dihydropyrimidine dehydrogenase	1.43
NM_001316	CSE1L	CSE1 chromosome segregation 1-like (yeast)	1.43
NM_000167	GK	glycerol kinase	1.43
NM_001924	GADD45 A	growth arrest and DNA-damage-inducible, alpha	1.42
NM_014225	PPP2R1A	protein phosphatase 2 (formerly 2A), regulatory subunit A (PR 65), alpha isoform	1.42
NM_001233	CAV2	caveolin 2	1.42
NM_176863	PSME3	proteasome (prosome, macropain) activator subunit 3 (PA28 gamma; Ki)	1.42
NM_001905	CTPS	CTP synthase	1.41
NM_005653	TFCP2	transcription factor CP2	1.41
NM_003405	YWHAH	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, eta polypeptide	1.41
NM_003392	WNT5A	wingless-type MMTV integration site family, member 5A	1.40
NM_002375	MAP4	microtubule-associated protein 4	1.40
NM_006353	HMGN4	high mobility group nucleosomal binding domain 4	1.39
NM_006527	SLBP	stem-loop (histone) binding protein	1.39
NM_000517	HBA2	hemoglobin alpha 2	1.38
NM_002661	PLCG2	phospholipase C, gamma 2 (phosphatidylinositol-specific)	1.38
NM_001493	GDI1	GDP dissociation inhibitor 1	1.38
NM_181430	FOXK2	forkhead box K2	1.38
NM_002086	GRB2	growth factor receptor-bound protein 2	1.38
NM_002868	RAB5B	RAB5B, member RAS oncogene family	1.37
NM_002768	PCOLN3	procollagen (type III) N-endopeptidase	1.37
NM_014742	TM9SF4	transmembrane 9 superfamily protein member 4	1.37
NM_004344	CETN2	centrin, EF-hand protein, 2	1.37
NM_002881	RALB	v-ral simian leukemia viral oncogene homolog B (ras related; GTP binding protein)	1.36
NM_004099	STOM	stomatin	1.36
NM_031844	HNRPU	heterogeneous nuclear ribonucleoprotein U (scaffold attachment factor A)	1.36
NM_000480	AMPD3	adenosine monophosphate deaminase (isoform E)	1.35
NM_006561	CUGBP2	CUG triplet repeat RNA binding protein 2	1.35
NM_152879	DGKD	diacylglycerol kinase delta 130kDa	1.35
NM_138558	PPP1R8	protein phosphatase 1 reQulatory (inhibitor) subunit 8	1.35
NM_004941	DHX8	DEAH (Asp-Glu-Ala-His) box polypeptide 8	1.34
NM_021079	NMT1	N-myristoyltransferase 1	1.33
NM_004622	TSN	translin	1.33
NM_002473	MYH9	myosin, heavy polypeptide 9, non-muscle	1.33
NM_006889	CD86	CD86 antigen (CD28 antigen ligand 2, B7-2 antigen)	1.33
NM_004383	CSK	c-src tyrosine kinase	1.33
NM_004317	ASNA1	arsA arsenite transoorter ATP-binding homolog 1 (bacterial)	1.33
NM_024298	LENG4	leukocyte receptor cluster (LRC) member 4	1.32
NM_001912	CTSL	cathepsin L	1.32
NM_001357	DHX9	DEAH (Asp-Glu-Ala-His) box polypeptide 9	1.32
NM_006849	PDIP	protein disulfide isomerase, pancreatic	1.32
NM_018457	DKFZP564J157	DKFZ, 0564J157 protein	1.31
NM_024880	TCF7L2	transcription factor 7-like 2 (T-cell specific, HMG-box)	1.31
NM_002081	GPC1	glypican 1	1.31
NM_004235	KLF4	Kruppel-like factor 4 (gut)	1.31

NM_005565	LCP2	lymphocyte cytosolic protein 2 (SH2 domain containing leukocyte protein of 76kDa)	1.30
NM_002667	PLN	phospholamban	1.30
NM_004946	DOCK2	dedicator of cytokinesis 2	1.30
NM_002035	FVT1	follicular lymphoma variant translocation 1	1.29
NM_002865	RAB2	RAB2 member RAS oncogene family	1.29
NM_002806	PSMC6	proteasome (prosome macropain) 26S subunit ATPase 6	1.29
NM_004240	TRIP10	thyroid hormone receptor interactor 10	1.28
NM_003760	EIF4G3	eukaryotic translation initiation factor 4 gamma, 3	1.28
NM_005151	USP14	ubiquitin specific protease 14 (tRNA guanine transglycosylase)	1.28
NM_015922	H105E3	NAD(P) deoendent steroid dehydropenase-like	1.27
NM_033306	CASP4	caspase 4 apoptosis-related cysteine protease	1.27
NM_198189	COPS8	COP9 constitutive photomorphogenic homolog subunit 8 (Arabidopsis)	1.27
NM_001933	DLST	dihydrolipoamide S-succinyltransferase (E2 component of 2-oxo-glutarate complex)	1.27
NM_015004	KIAA0116	KIAA0116 protein	1.27
NM_033362	MRPS12	mitochondrial ribosomal protein S12	1.27
NM_004180	TANK	TRAF family member-associated NFKB activator	1.26
NM_014734	KIAA0247	KIAA0247	1.26
NM_005271	GLUD1	glutamate dehydropenase 1	1.25
NM_003009	SEPW1	selenoprotein W, 1	1.25
NM_182641	FALZ	fetal Alzheimer antigen	1.24
NM_007362	NCBP2	nuclear cap binding protein subunit 2 20kDa	1.24
NM_004292	RIN1	Ras and Rab interactor 1	1.24
NM_014608	CYFIP1	cytoplasmic FMR1 interacting protein 1	1.23
NM_022333	TIAL1	TIA1 cytotoxic oranule-associated RNA binding protein-like 1	1.23
NM_003126	SPTA1	spectrin alpha erythrocytic 1 (elliptocytosis 2)	1.22
NM_014602	PIK3R4	phosphoinositide-3-kinase regulatory subunit 4, p150	1.18
NM_002194	INPP1	inositol polyphosphate-1-phosphatase	1.16

Note: Accession IDs "NM_XXXX" are uniquely assigned to each gene by National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/sites/entrez?db=nuccore).

Table 7: Commonly Downregulated Genes in Pancreatic Cancer			
Accession	Gene Symbol	Gene Name	FC
NM_006499	LGALS8	galoctosite-binding, soluble, 8 (galectin 8)	0.87
NM_000466	PEX1	peroxisome biogenesis factor 1	0.81
NM_002766	PRPSAP1	phosphoribosyl pyrophosphate synthetase-associated protein 1	0.81
NM_147131	GALT	galactose-1-phosphate uridylyltransferase	0.80
NM_002101	GYPC	glycophorin C (Gerbich blood group)	0.80
NM_002880	RAF1	v-raf-1 murine leukemia viral oncogene homolog 1	0.80
NM_004649	C218rf33	chromosome 21 open reading frame 33	0.80
NM_003262	TLOC1	translocation protein 1	0.79
NM_147223	NCOA1	nuclear receptor coactivator 1	0.79
NM_007062	PWP1	nuclear phosphoprotein similar to S. cerevisiae PWP1	0.79
NM_005561	LAMP1	lysosomal-associated membrane protein 1	0.79
NM_006810	PDIR	for protein disulfide isomerase-related	0.78
NM_033360	KRAS2	v-Ki-ras2 Kirsten rat sarcoma 2 viral oncogene homolog	0.77
NM_001513	GSTZ1	glutathione transferase zeta 1 (maleylacetoacetate isomerase)	0.77
NM_006184	NUCB1	nucleobindin 1	0.77
NM_001634	AMD1	adenosylmethionine decarboxylase 1	0.76
NM_006749	SLC20A2	solute carrier family 20 (phosphate transporter), member 2	0.76
NM_003144	SSR1	signal sequence receptor alpha (translocon-associated protein alpha)	0.76
NM_004606	TAF1	TAF1 RNA polymerase II, TATA box binding protein (TBP)-associated factor 250kDa	0.75
BX648788		MRNA; cDNA DKFZP686M12165 (from clone DKFZP686M12165)	0.75
NM_004035	ACOX1	acyl-Coenzyme A oxidase 1 palmitoyl	0.74
NM_000287	PEX6	peroxisomal biogenesis factor 6	0.73
NM_003884	PCAF	p300/CBP-associated factor	0.73
NM_006870	DSTN	destrin (actin depolymerizing factor)	0.73
NM_001604	PAX6	paired box gene 6 (aniridia keratitis)	0.72
NM_000722	CACNA2 D1	calcium channel voltage-dependent alpha 2/delta subunit 1	0.72
NM_033022	RPS24	ribosomal protein S24	0.72
NM_004563	PCK2	phosphoenolpyruvate carboxykinase 2 (mitochondrial)	0.72
NM_002602	PDE6G	phosphodiesterase 6G cGMP-specific, rod, gamma	0.72
NM_001889	CRYZ	crystalline, zeta (quinone reductase)	0.72
NM_002339	LSP1	lymphocyte-specific protein 1	0.71
NM_016848	SHC3	src homology 2 domain containing transforming protein C3	0.71
NM_002906	RDX	radixin	0.71
NM_007014	WWP2	Nedd-4-like ubiquitin-protein ligase	0.71
NM_000414	HSD17B4	hydroxysteroid (17-beta) dehydrogenase 4	0.71
NM_001127	AP1B1	adaptor-related protein complex 1, beta 1 subunit	0.70
NM_002402	MEST	mesoderm specific transcript homolog (mouse)	0.70
NM_033251	RPL13	ribosomal protein L13	0.70
NM_139069	MAPK9	mitogen-activated protein kinase 9	0.70
NM_002913	RFC1	replication factor C (activator 1) 1, 145kDa	0.70
NM_000487	ARSA	arylsulfatase A	0.70
NM_006973	ZNF32	zinc finger protein 32 (KOX 30)	0.70
NM_005310	GRB7	growth factor receptor-bound protein 7	0.69
NM_005962	MXI1	MAX interacting protein 1	0.69
NM_005359	MADH4	MAD, mothers against decapentaplegic homolog 4 (Drosophila)	0.69
NM_002340	LSS	lanosterol synthase (2 3-oxidosqualene-lanosterol cyclase)	0.68
NM_003684	MKNK1	MAP kinase-interacting serine/threonine kinase 1	0.68
NM_005671	D8S2298 E	reproduction 8	0.68
NM_000309	PPOX	protoporphyrinogen oxidase	0.68
NM_000994	RPL32	ribosomal protein L32	0.68
NM_000972	RPL7A	ribosomal protein L7a	0.67
NM_005101	G1P2	interferon, alpha-inducible protein (clone IFI-15K)	0.67
NM_001129	AEBP1	AE binding protein 1	0.67
NM_001011	RPS7	ribosomal protein S7	0.67
NM_001153	ANXA4	annexin A4	0.66
NM_012335	MY01F	myosin IF	0.66
NM_005007	NFKBIL1	nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor-like 1	0.66
NM_001870	CPA3	carboxypeptidase A3 (mast cell)	0.66
NM_181826	NF2	neurofibromin 2 (bilateral acoustic neuroma)	0.66
NM_000285	PEPD	peptidase D	0.66
NM_006180	NTRK2	neurotrophic tyrosine kinase, receptor type 2	0.66
NM_000543	SMPD1	sphingomyelin phosphodiesterase 1, acid lysosomal (acid sphinagmyelinase)	0.65
NM_001459	FLT3LG	fms-related tyrosine kinase 3 ligand	0.65
NM_003750	EIF3S10	eukaryotic translation initiation factor 3, subunit 10 theta, 150/170kDa	0.65

NM_005570	LMAN1	lectin mannose-binding, 1	0.65
NM_004409	DMPK	dystrophia myotonica-protein kinase	0.65
NM_172159	KCNAB1	potassium voltage-gated channel, shaker-related subfamily, beta member 1	0.65
XM 352750	COL14A1	collagen, type XIV, alpha 1 (undulin)	0.65
NM_001731	BTG1	B-cell translocation gene 1, anti-proliferative	0.65
NM_000884	IMPDH2	IMP (inosine monophosphate) dehydrogenase 2	0.64
NM_001885	CRYAB	crystallin, alpha B	0.64
NM_000240	MAOA	monoamine oxidase A	0.64
NM_003136	SRP54	signal recognition particle 54kDa	0.63
NM_000281	PCBD	6-pyruvoyl-tetrahydropterin synthase/dimerization cofactor of hepatocyte nuclear factor 1 alpha (TCF1)	0.63
NM_005729	PPIF	peptidylprolpyl isomerase F (cyclophilin F)	0.63
NM_006481	TCF2	transcription factor 2, hepatic; LF-B3' variant hepatic nuclear factor	0.63
NM_002089	CXCL2	chemokine (C-X-C motif) ligand 2	0.63
NM_001961	EEF2	eukaryotic translation elongation factor 2	0.63
NM_001801	CDO1	cysteine dioxygenase type I	0.63
NM_006389	HYOU1	hypoxia up-regulated 1	0.63
XM 167711	ITGA8	integrin, alpha 8	0.62
NM_014765	TOMM20	translocase of outer mitochondrial membrane 20 homolog (yeast)	0.62
NM_006714	SMPDL3 A	sphingomyelin phosphodiesterase, acid-like 3A	0.62
NM_000016	ACAOM	acyl-Coenzyme A dehydrogenase C-4 to C-12 straight chain	0.62
NM_003924	PHOX2B	paired-like homeobox 2b	0.62
NM_002078	GOLGA4	golgi autoantigen, golgin subfamily a 4	0.62
NM_002736	PRKAR2 B	protein kinase cAMP-dependent, regulatory, type II beta	0.62
BQ217469	KIAA0114	KIAA0114 gene product	0.61
NM_006307	SRPX	sushi-repeat-containing protein X-linked	0.61
NM_002184	IL6ST	interleukin 6 signal transducer (gp130 oncostatin M receptor)	0.61
NM_153186	ANKR015	ankyrin repeat domain 15	0.61
NM_003038	SIC1A4	solute carrier family 1 (glutamate/neutral amino acid transporter), member 4	0.60
NM_006195	PBX3	pre-B-cell leukemia transcription factor 3	0.60
NM_000327	ROM1	retinal outer segment membrane protein 1	0.60
NM_003463	PTP4A1	protein tyrosine phosphatase type IVA, member 1	0.60
NM_001520	GTF3C1	general transcription factor iiiC polypeptide 1 alpha 220kDa	0.60
NM_006277	ITSN2	intersectin 2	0.59
NM_000985	RPL17	ribosomal protein L17	0.59
NM_000909	NPY1R	neuropeptide Y receptor Y1	0.59
NM_001014	RPS10	ribosomal protein S10	0.59
NM_022307	ICA1	islet cell autoantigen 1 69kDa	0.58
NM_002567	PBP	prostatic binding protein	0.58
NM_012324	MAPK8IP 2	mitogen-activated protein kinase 8 interacting protein 2	0.58
NM_004490	GRB14	growth factor receptor-bound protein 14	0.58
NM_004733	SLC33A1	solute carrier family 33 (acetyl-CoA transporter), member 1	0.57
NM_002197	AC01	aconitase 1, soluble	0.57
NM_000505	F12	coagulation factor Xii (Hageman factor)	0.57
NM_005010	NRCAM	neuronal cell adhesion molecule	0.56
NM_006963	ZNF22	zinc finger protein 22 (KOX 15)	0.56
NM_006827	TMP21	transmembrane trafficking protein	0.55
NM_004394	DAP	death-associated protein	0.54
NM_001089	ABCA3	ATP-binding cassette, sub-family A (ABC), member 3	0.54
NM_004470	FKBP2	FK506 binding protein 2, 13kDa	0.53
NM_005749	TOB1	transducer of ERBB2, 1	0.53
NM_001355	DDT	D-dopachrome tautomerase	0.53
NM_002111	HD	huntington (Huntington disease)	0.53
NM_002635	SIC25A3	solute carrier family 25 (mitochondrial carrier; phosphate carrier), member 3	0.53
NM_005596	NFIB	nuclear factor I/B	0.53
NM_006273	CCL7	chemokine (C-C motif) ligand 7	0.53
NM_001013	RPS9	ribosomal protein S9	0.52
NM_001551	IGBP1	immunoglobulin (CD79A) binding protein 1	0.52
NM_004498	ONECUT 1	one cut domain, family member 1	0.52
NM_004484	GPC3	glypican 3	0.52
NM_130797	DPP6	dipeptidylpeptidase 6	0.52
NM_000746	CHRNA7	cholinergic receptor, nicotinic, alpha polypeptide 7	0.51
NM_001756	SERPINA 6	serine (or cysteine) proteinase inhibitor, clade A (alpha-1 antiproteinase antitrypsin), member 6	0.51
NM_001327	CTAG1	cancer/testis antigen 1	0.51
NM_003651	CSDA	cold shock domain protein A	0.50
NM_005848	IRLB	c-myc promoter-binding protein	0.50
BC040073	H19	H19, imprinted maternally expressed untranslated mRNA	0.50
NM_002228	JUN	v-jun sarcoma virus 17 oncogene homolog (avian)	0.49
NM_000795	DRD2	dopamine receptor D2	0.48
NM_002084	GPX3	glutathione peroxidase 3 (plasma)	0.48

NM_002716	PPP2R1B	protein phosphatase 2 (formerly 2A), regulatory subunit A (PR 65), beta isoform	0.48
NM_005166	APLP1	amyloid beta (A4) precursor-like protein 1	0.48
NM_005911	MAT2A	methionine adenosyltransferase II, alpha	0.47
NM_000208	INSR	insulin receptor	0.47
NM_170736	KCNJ15	potassium inwardly-rectifying channel, subfamily J, member 15	0.47
NM_001190	BCAT2	branched chain aminotransferase 2, mitochondrial	0.47
NM_005336	HDLBP	high density lipoprotein binding protein (vigilin)	0.46
NM_001076	UGT2B15	UDP glycosyltransferase 2 family, polypeptide B15	0.46
NM_001152	SLC25A5	solute carrier family 25 (mitochondrial carrier; adenine nucleotide translocator, member 5	0.46
NM_002729	HHEX	hematopoietically expressed homeobox	0.46
NM_002847	PTPRN2	protein tyrosine phosphatase, receptor type, N polypeptide 2	0.44
NM_000447	PSEN2	presenilin 2 (Alzheimer disease 4)	0.44
NM_152868	KCNJ4	potassium inwardly-rectifying channel, subfamily J, member 4	0.44
NM_001759	CCND2	cyclin D2	0.44
NM_000316	PTHR1	parathyroid hormone receptor 1	0.44
NM_001612	ACRV1	acrosomal vesicle protein 1	0.43
NM_002467	MYC	v-myc myelocytomatosis viral oncogene homolog (avian)	0.43
NM_004454	ETV5	ets variant gene 5 (ets-related molecule)	0.43
NM_002846	PTPRN	protein tyrosine phosphatase, receptor type N	0.43
NM_005622	SAH	SA hypertension-associated homolog (rat)	0.42
NM_001989	EVX1	eve, even-skipped homeo box homolog 1 (Drosophila)	0.42
NM_000166	GJB1	gap junction protein, beta 1, 32kDa (connexin 32, Charcot-Marie-Tooth neuropathy, X-linked)	0.42
NM_014685	HERPUD 1	homocysteine-inducible, endoplasmic reticulum stress-inducible, ubiquitin-like domain member 1	0.42
NM 001735	C5	complement component 5	0.41
NM 005504	BCAT1	branched chain aminotransferase 1, ctyosolic	0.41
NM 006808	SEC61B	Sec61 beta subunit	0.40
NM 006751	SSFA02	sperm specific antigen 2	0.39
NM 005947	MT1B	metallothionein 1B (functional)	0.38
NM 005576	LOXL1	lysyl oxidase-like 1	0.37
NM 005627	SGK	serum/glucocorticoid regulated kinase	0.36
NM 004683	RGN	regucalcin (senescence marker protein-30)	0.36
NM 00918	P4HB	procollagen-proline, 2-oxoglutarate 4-dioxygenase (proline 4-hydroxylase), beta polypeptide (protein disulfide isomerase; thyroid hormone binding protein p55)	0.36
BC044862		Macrophage stimulating 1 (hepatocyte growth factor-like), mRNA (cDNA clone IMAGE:4821945), with apparent retained intron	0.35
NM 005952	MT1X	metallothionein 1X	0.35
NM 000429	MAT1A	methionine adenosyltransferase I, alpha	0.35
NM 004010	DMD	dystrophin (muscular dystrophy, Duchenne and Becker types)	0.34
NM 000689	ALDH1A1	aldehyde dehydrogenase 1 family, member A1	0.34
NM 002889	RARRES2	retinoic acid receptor responder (tazarotene induced) 2	0.33
NM 006280	SSRA	signal sequence receptor, delta (translocon-associated protein delta)	0.33
NM 003819	PABPC4	poly(A) binding protein, cytoplasmic 4 (inducible form)	0.32
NM 000755	CRAT	carnitine aceltyltransferase	0.32
NM 015684	ATP5S	ATP synthase, H+ transporting, mitochondrial F0 complex, subunit s (factor B)	.030
NM 033200	BC002942	hypothetical protein BC002942	0.30
BCG986717		Transcribed sequences	0.29
NM 148923	CYB5	cytochrome b-5	0.29
NM 000609	CXCL12	chemokine (C-X-C motif) ligand 12 (stromal cell-derived factor 1)	0.29
NM 001979	EPHX2	epoxide hydrolase 2, cytoplasmic	0.28
NM 001332	CTNND2	catenin (caherin-associated protein), delta 2 (neural plakophilin-related arm-repeat protein)	0.27
NM 001831	CLU	clusterin (complement lysis inhibitor, SP-40, 40, sulfated glycoprotein 2, testosterone-repressed prostate message 2, apolipoprotein J)	0.27
NM 005080	XBP1	X-box binding protein 1	0.27
NM 000156	GAMT	guanidinoacetate N-methyltransferase	0.27
NM 182848	CLDN10	claudin 10	0.26
NM 000065	C6	complement component 6	0.26
NM 000128	F11	coagulation factor XI (plasma thromboplasin antecedent)	0.24
NM 003822	MR5A2	nuclear receptor subfamily 5, group A, member 2	0.24
NM 006406	PRDX4	peroxiredoxin 4	0.21
BM799844	BNIP3	BCL2/adenovirus E1B 19kDa interacting protein 3	0.21
NM 018646	TRPV6	transient receptor potential cation channel, subfamily V, member 6	0.21
NM 005013	NUCB2	nucleobindin 2	0.21
NM 000624	SERPINA 3	serine (or cysteine) proteinase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 3	0.19
NM 005065	SEL 1L	sel-1 suppressor of lin-12-like (C. elegans)	0.18
NM 198235	RNASE 1	ribonuclease, RNase A family, 1 (pancreatic)	0.17
NM 006498	LGALS2	lectin, galactoside-binding, soluble, 2 (galectin 2)	0.16
NM 002899	RBP1	retinol binding protein 1, cellular	0.12
NM 004413	DPEP1	dipeptidase 1 (renal)	0.12
NM 021603	FXYD2	FXYD domain contaming ion transport regulator 2	0.09
NM 138938	PAP	pancreatitis-associated protein	0.08

NM 201553	FGL	fibrinogen-like 1	0.07
NM 001482	GATM	glycerine amidinotransferase (L-arrginine: glycine amidinotransferase)	0.04
NM 033240	ELA2A	elastase 2[a]	0.02
NM 000101	CYBA	cytochrome b-245, alpha polypeptide	0.02

Note: Accession IDs "NM_XXXX" are uniquely assigned to each gene by National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/sites/entrez?db=nuccore).

Table 8. microRNAs that are up-regulated in glioblastoma cells.

Fold change	microRNA
Up 10X	miR-10b, miR-10a, miR-96
Up 2-10X	miR-182, miR-199b, miR-21, miR124, miR-199a, miR-199-s, miR-199a, miR-106b, miR-15b, miR-188, miR-148a, miR-104, miR-224, miR-368, miR-23a, miR-210N, miR-183, miR-25, miR-200cN, miR-373, miR-17-5p, let-7a, miR-16, miR-19b, miR-26a, miR-27a, miR-92, miR-93, miR-320 and miR-20
Up 1-2 X	miR-143, miR-186. miR-337, miR-30a-3p, miR-355, miR-324-3p etc.

Table 9. microRNAs that are down-regulated in glioblastoma cells.

Fold change	microRNA
Down 10X	miR-218, miR-124a, miR-124b, miR-137, miR-184, miR-129, miR-33, miR-139, miR-128b, miR-128a, miR-330, miR-133a, miR-203, miR-153, miR-326, miR-105, miR-338, miR-133b, miR-132, miR-154, miR-29bN
Down 2-10X	miR-7N, miR-323, miR-219, miR-328, miR-149, miR-122a, miR-321, miR-107, miR-190, miR-29cN, miR-95, miR-154, miR-221, miR-299, miR-31, miR-370, miR-331, miR-342, miR-340

Table 10. MMP genes contained within microvesicles isolated from glioblastoma cell line.

Gene Symbol	Accession ID	Gene Description
MMP1	AK097805	Homo sapiens cDNA FLJ40486 fis, clone TESTI2043866. [AK097805]
MMP8	NM_002424	Homo sapiens matrix metallopeptidase 8 (neutrophil collagenase) (MMP8), mRNA [NM_002424]
MMP12	NM_002426	Homo sapiens matrix metallopeptidase 12 (macrophage elastase) (MMP12), mRNA [NM_002426]
MMP15	NM_002428	Homo sapiens matrix metallopeptidase 15 (membrane-inserted) (MMP15), mRNA [NM_002428]
MMP20	NM_004771	Homo sapiens matrix metallopeptidase 20 (enamelysin) (MMP20), mRNA [NM_004771]
MMP21	NM_147191	Homo sapiens matrix metallopeptidase 21 (MMP21), mRNA [NM_147191]
MMP24	NM_006690	Homo sapiens matrix metallopeptidase 24 (membrane-inserted) (MMP24), mRNA [NM_006690]
MMP26	NM_021801	Homo sapiens matrix metallopeptidase 26 (MMP26), mRNA [NM_021801]
MMP27	NM_022122	Homo sapiens matrix metallopeptidase 27 (MMP27), mRNA [NM_022122]

Note: Gene symbols are standard symbols assigned by Entrz Gene (http://www.ncbi.nlm.nih.gov/sites/entrez?db=gene). Accession IDs are uniquely assigned to each gene by National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/sites/entrez?db=nuccore).

Table 11. Genes containing somatic mutations in glioblastoma adapted from the result of TCGA project (McLendon et al., 2008).

Hugo Gene Symbol	Entrez_Gene_Id
A2M	2
A2M	2
A2M	2
ABCA3	21
ABCC4	10257
ABCC4	10257
ABCC4	10257
ADAM12	8038
ADAM15	8751
ADAMTSL3	57188
ADAMTSL3	57188
ADM	133
AIFM1	9131
AKAP2	11217
AKAP2	11217
ALK	238
ANK2	287
ANK2	287
ANK2	287
ANK2	287
ANK2	287
ANXA1	301
ANXA7	310
AOC3	8639
AOC3	8639
APBB1IP	54518
APC	324
ARNT	405
ASPM	259266
ASPM	259266
ASXL1	171023
ASXL1	171023
ATM	472
ATM	472
ATM	472
ATP6V1E1	529
ATR	545
AVIL	10677
AXL	558
BAI3	577
BAI3	577
BAI3	577
BAMBI	25805
BCAR1	9564
BCAR1	9564
BCL11A	53335
BCL11A	53335
BCL11A	53335
BCL11A	53335
BCL11A	53335
BCL2L13	23786
BCR	613
BMPR1A	657
BRCA1	672
BRCA2	675
BRCA2	675
BRCA2	675
BTK	695
C18orf25	147339
C20orf160	140706
C20orf160	140706
C22orf24	25775
C6orf60	79632
C6orf60	79632
C9orf72	203228
CAND1	55832
CASP9	842
CAST	831
CAST	831
CAST	831
CBL	867
CBL	867
CCR5	1234
CD46	4179
CDC123	8872
CDKL5	6792
CDKN2A	1029
CDKN2A	1029
CDKN2A	1029
CENPF	1063
CENPF	1063
CENTG1	116986
CENTG1	116986
CES3	23491
CES3	23491
CHAT	1103
CHAT	1103
CHD5	26038
CHEK1	1111
CHEK1	1111
CHEK1	1111
CHEK1	1111
CHEK2	11200
CHEK2	11200
CHEK2	11200
CHEK2	11200
CHEK2	11200
CHEK2	11200
CHEK2	11200
CHEK2	11200
CHEK2	11200
CHEK2	11200
CHEK2	11200
CHEK2	11200
CHEK2	11200
CHI3L2	1117
CHIC2	26511
CHL1	10752
CHL1	10752
CMTM3	123920
CNTFR	1271
COL11A1	1301
COL1A1	1277
COL1A1	1277
COL1A1	1277
COL1A1	1277
COL1A2	1278
COL1A2	1278
COL3A1	1281
COL3A1	1281
COL3A1	1281
COL3A1	1281
COL5A1	1289
COL6A2	1292
COL6A2	1292
COL6A2	1292
CRLF1	9244
CSF3R	1441
CSF3R	1441
CSMD3	114788
CSMD3	114788
CSNK1E	1454
CTNNB1	1499
CTSH	1512
CTSH	1512
CYLD	1540
CYP27B1	1594
CYP27B1	1594
CYP3A4	1576
DCX	1641
DDIT3	1649
DDR2	4921
DDR2	4921
DDR2	4921
DES	1674
DES	1674

Hugo Gene Symbol	Entrez_Gene_Id
DGKD	8527
DGKG	1608
DHTKD1	55526
DMBT1	1755
DMRT3	58524
DOCK1	1793
DOCK1	1793
DOCK1	1793
DOCK8	81704
DOCK8	81704
DPYSL4	10570
DPYSL4	10570
DST	667
DST	667
DST	667
DST	667
DST	667
DST	667
DST	667
DST	667
DTX3	196403
EGFR	1956
EGFR	1956
EGFR	1956
EGFR	1956
EGFR	1956
EGFR	1956
EGFR	1956
EGFR	1956
EGFR	1956
EGFR	1956
EGFR	1956
EGFR	1956
EGFR	1956
EGFR	1956
EGFR	1956
EGFR	1956
EGFR	1956
EGFR	1956
EGFR	1956
EGFR	1956
EGFR	1956
ELAVL2	1993
EP300	2033
EP300	2033
EP400	57634
EP400	57634
EP400	56734
EPHA2	1969
EPHA3	2042
EPHA3	2042

Hugo Gene Symbol	Entrez_Gene_Id
EPHA4	2043
EPHA4	2043
EPHA6	285220
EPHA7	2045
EPHA7	2045
EPHA8	2046
EPHA8	2046
EPHB1	2047
ERBB2	2064
ERBB2	2064
ERBB2	2064
ERBB2	2064
ERBB2	2064
ERBB2	2064
ERBB2	2064
ERBB2	2064
ERBB2	2064
ERBB2	2064
ERBB2	2064
ERBB2	2064
ERBB3	2065
ESR1	2099
ETNK2	55224
EYA1	2138
EYA1	2138
F13A1	2162
FBXW7	55294
FBXW7	55294
FGFR1	2260
FGFR1	2260
FGFR2	2263
FGFR3	2261
FKBP9	11328
FKBP9	11328
FKBP9	11328
FKBP9	11328
FKBP9	11328
FKBP9	11328
FKBP9	11328
FKBP9	11328
FKBP9	11328
FKBP9	11328
FKBP9	11328
FKBP9	11328
FLI1	2313
FLI1	2313
FLT1	2321
FLT4	2324
FN1	2335
FN1	2335
FN1	2335
FN1	2335
FN1	2335

Hugo Gene Symbol	Entrez_Gene_Id
FN1	2335
FOXO3	2309
FOXO3	2309
FOXO3	2309
FRAP1	2475
FURIN	5045
FURIN	5045
FURIN	5045
GARNL3	84253
GATA3	2625
GATA3	2625
GCLC	2729
GDF10	2662
GLI1	2735
GLI3	2737
GLTSCR2	29997
GNAI1	2770
GNAS	2778
GNAS	2778
GPR78	27201
GRIA2	2891
GRLF1	2909
GRN	2896
GRN	2896
GSTM5	2949
GSTM5	2949
GSTM5	2949
GSTM5	2949
GSTM5	2949
GSTM5	2949
GSTM5	2949
GSTM5	2949
GSTM5	2949
GYPC	2995
HCK	3055
HCK	3055
HELB	92797
HLA-E	3133
HLA-E	3133
HLA-E	3133
HLA-E	3133
HS3ST3A1	9955
HSP90AA1	3320
HSP90AA1	3320
HSPA8	3312
HSPA8	3312
HSPA8	3312
HSPA8	3312
HSPA8	3312
HSPA8	3312
HSPA8	3312
ID3	3399
IFITM3	10410

Hugo Gene Symbol	Entrez_Gene_Id
IFITM3	10410
IFITM3	10410
IFITM3	10410
IFITM3	10410
IFITM3	10410
IFITM3	10410
IL1RL1	9173
IL31	386653
ILK	3611
ING4	51147
ING4	51147
ING4	51147
INHBE	83729
IQGAP1	8826
IRAK3	11213
IRS1	3667
IRS1	3667
ISL1	3670
ITGAL	3683
ITGB2	3689
ITGB2	3689
ITGB2	3689
ITGB3	3690
ITGB3	3690
ITGB3	3690
ITGB3	3690
ITGB3	3690
JAG1	182
KIAA1632	57724
KIF3B	9371
KIT	3815
KIT	3815
KIT	3815
KLF4	9314
KLF4	9314
KLF6	1316
KLF6	1316
KLK8	11202
KPNA2	3838
KPNA2	3838
KRAS	3845
KSR2	283455
KSR2	283455
KTN1	3895
LAMP1	3916
LAMP1	3916
LAX1	54900
LCK	3932
LDHA	3939
LDHA	3939
LGALS3BP	3959
LGALS3BP	3959
LGALS3BP	3959

Hugo Gene Symbol	Entrez_Gene_Id
LRRN2	10446
LTF	4057
LTF	4057
LYN	4067
MAG	4099
MAP3K6	9064
MAPK13	5603
MAPK7	5598
MAPK8IP2	23542
MAPK8IP3	23162
MAPK9	5601
MAPK9	5601
MARK1	4139
MARK1	4139
MDM2	4193
MDM4	4194
MEOX2	4223
MET	4233
MET	4233
MET	4233
MLH1	4292
MLH1	4292
MLH1	4292
MLL4	9757
MLL4	9757
MLL4	9757
MLLT7	4303
MMD2	221938
MN1	4330
MSH2	4436
MSH2	4436
MSH6	2956
MSH6	2956
MSH6	2956
MSH6	2956
MSI1	4440
MSI1	4440
MTAP	4507
MUSK	4593
MYCN	4613
MYCN	4613
MYLK2	85366
MYO3A	53904
MYST4	23522
MYST4	23522
MYST4	23522
MYST4	23522
NBN	4683
NDUFA10	4705
NEK10	152110
NELL2	4753
NF1	4763
NF1	4763

Hugo Gene Symbol	Entrez_Gene_Id
NF1	476
NF1	476
NF1	476
NF1	476
NF1	476
NF1	476
NF1	476
NF1	476
NF1	476
NF1	476
NF1	476
NF1	476
NF1	476
NF1	476
NMBR	482
NMBR	482
NOS3	484
NOS3	484
NOTCH1	485
NOTCH1	485
NRXN3	936
NTRK3	491
NUMA1	492
NUP214	802
ONECUT2	948
OR5P2	12006
PAX5	507
PDGFRA	515
PDGFRA	515
PDGFRA	515
PDGFRB	515
PDGFRB	515
PDK2	516
PDPK1	517
PDZD2	2303
PDZD2	2303
PHLPP	2323
PI15	5105
PI15	5105
PIK3C2A	528
PIK3C2B	528
PIK3C2G	528
PIK3C2G	528
PIK3C2G	528
PIK3C2G	528
PIK3C2G	528
PIK3CA	529
PIK3CA	529
PIK3CA	529
PIK3CA	529
PIK3CA	529
PIK3R1	529
PIK3R1	529

Hugo Gene Symbol	Entrez_Gene_Id	Hugo Gene Symbol	Entrez_Gene_Id	Hugo Gene Symbol	Entrez_Gene_Id
PIK3R1	5295	PTEN	5728	SLIT2	9353
PIK3R1	5295	PTEN	5728	SMAD2	4087
PIK3R1	5295	PTEN	5728	SMAD4	4089
PIK3R1	5295	PTEN	5728	SNF1LK2	23235
PIM1	5292	PTEN	5728	SNF1LK2	23235
PLAG1	5324	PTEN	5728	SNX13	23161
PML	5371	PTEN	5728	SOCS1	8651
PMS2	5395	PTEN	5728	SOX11	6664
POU2F1	5451	PTEN	5728	SOX11	6664
PPP2R5D	5528	PTEN	5728	SPARC	6678
PRKCA	5578	PTEN	5728	SPDEF	25803
PRKCA	5578	PTEN	5728	SPN	6693
PRKCB1	5579	PTEN	5728	SPRED3	399473
PRKCB1	5579	PTK2B	2185	SRPK2	6733
PRKCD	5580	PTPN11	5781	ST7	7982
PRKCD	5580	PTPN11	5781	STAT1	6772
PRKCD	5580	RADIL	55698	STAT3	6774
PRKCD	5580	RADIL	55698	STK32B	55351
PRKCD	5580	RB1	5925	STK36	27148
PRKCD	5580	RB1	5925	SYP	6855
PRKCZ	5590	RB1	5925	TAF1	6872
PRKCZ	5590	RB1	5925	TAF1	6872
PRKD2	25865	RB1	5925	TAOK3	51347
PRKD2	25865	RB1	5925	TAS1R1	80835
PRKDC	5591	RB1	5925	TBK1	29110
PRKDC	5591	RB1	5925	TBK1	29110
PRKDC	5591	RB1	5925	TCF12	6938
PROX1	5629	RINT1	60561	TCF12	6938
PSMD13	5719	RIPK4	54101	TCF12	6938
PSMD13	5719	RNF38	152006	TERT	7015
PSMD13	5719	ROR2	4920	TERT	7015
PTCH1	5727	ROR2	4920	TGFBR2	7048
PTCH1	5727	ROS1	6098	TIMP2	7077
PTEN	5728	ROS1	6098	TNC	3371
PTEN	5728	RPN1	6184	TNC	3371
PTEN	5728	RPS6KA3	6197	TNC	3371
PTEN	5728	RTN1	6252	TNFRSF11B	4982
PTEN	5728	RUNX1T1	862	TNK2	10188
PTEN	5728	RYR3	6263	TNK2	10188
PTEN	5728	RYR3	6263	TNK2	10188
PTEN	5728	SAC	55811	TNK2	10188
PTEN	5728	SAC	55811	TOP1	7150
PTEN	5728	SEMA3B	7869	TP53	7157
PTEN	5728	SERPINA3	12	TP53	7157
PTEN	5728	SERPINE1	5054	TP53	7157
PTEN	5728	SHH	6469	TP53	7157
PTEN	5728	SLC12A6	9990	TP53	7157
PTEN	5728	SLC12A6	9990	TP53	7157
PTEN	5728	SLC25A13	10165	TP53	7157
PTEN	5728	SLC25A13	10165	TP53	7157
PTEN	5728	SLC2A2	6514	TP53	7157
PTEN	5728	SLIT2	9353	TP53	7157
PTEN	5728	SLIT2	9353	TP53	7157

Hugo Gene Symbol	Entrez_Gene_Id
TP53	7157
TP53	7157
TP53	7157
TP53	7157
TP53	7157
TP53	7157
TP53	7157
TP53	7157
TP53	7157
TP53	7157
TP53	7157
TP53	7157
TP53	7157
TP53	7157
TP53	7157
TP53	7157
TP53	7157
TP53	7157
TP53	7157
TP53	7157
TP53	7157
TP53	7157
TP53	7157
TP53	7157
TP53	7157
TP53	7157
TP53	7157
TP53	7157
TP53	7157
TP53	7157
TP53	7157
TP53	7157
TP53	7157
TP53	7157
TP53	7157
TP53	7157
TP53	7157
TP53	7157
TP53	7157
TP53	7157
TP53	7157
TP53	7157
TP53	7157
TP53	7157
TP53	7157
TPBG	7162
TRIM24	8805
TRIM3	10612
TRIM33	51592
TRIP6	7205
TRRAP	8295
TRRAP	8295
TSC1	7248
TSC2	7249
TSC2	7249
TSC2	7249
UNG	7374
UPF2	26019
UPF2	26019
VAV2	7410
VLDLR	7436
WNT2	7472
ZEB1	6935
ZEB1	6935
ZNF384	171017
ZNF384	171017

Note: Hugo Gene Symbols are assigned to individual genes by HUGO Gene Nomenclature Committee (http://www.genenames.org/). Entrez_Gene_Ids are assigned to individual genes by Entrz Gene (http://www.ncbi.nlm.nih.gov/sites/entrez?db=gene).

Table 12. Genes containing somatic mutations in glioblastoma adapted from the paper by Parsons et. al. (Parsons et al., 2008)

Gene symbol	Accession ID
A2M	NM_000014
A4GALT	CCDS14041.1
A4GNT	CCDS3097.1
AACS	CCDS9263.1
ABCA10	CCDS11684.1
ABCA12	NM_015657
ABCA13	NM_152701
ABCA4	CCDS747.1
ABCA5	CCDS11685.1
ABCA7	CCDS12055.1
ABCA9	CCDS11681.1
ABCB1	CCDS5608.1
ABCB6	CCDS2436.1
ABCC10	CCDS4896.1
ABCC11	CCDS10732.1
ABCC3	NM_003786
ABCC5	NM_005688
ABCD2	CCDS8734.1
ABCF2	CCDS5922.1
ABCG2	CCDS3628.1
ABHD3	NM_138340
ABHD4	CCDS9572.1
ABHD7	CCDS736.1
ABL2	NM_007314
ABTB2	CCDS7890.1
ACAD9	CCDS3053.1
ACADS	CCDS9207.1
ACADSB	CCDS7634.1
ACAT2	CCDS5268.1
ACCN1	CCDS11276.1
ACCN3	CCDS5914.1
ACF	CCDS7241.1
ACLY	CCDS11412.1
ACOX3	CCDS3401.1
ACP5	CCDS12265.1
ACRBP	CCDS8554.1
ACTG1	CCDS11782.1
ACTN1	CCDS9792.1
ACTR10	NM_018477
ACTR1A	CCDS7536.1
ACTR8	CCDS2875.1
ACTRT1	CCDS14611.1
ADAM12	CCDS7653.1
ADAM15	CCDS1084.1
ADAM18	CCDS6113.1
ADAM28	NM_014265
ADAM29	CCDS3823.1
ADAMTS1	NM_006988
ADAMTS13	CCDS6970.1
ADAMTS17	CCDS10383.1
ADAMTS20	NM_175851
ADAMTS4	CCDS1223.1
ADAMTS8	NM_007037
ADAR	CCDS1071.1
ADARB2	CCDS7058.1
ADCY1	NM_021116
ADCY8	CCDS6363.1
ADRBK2	CCDS13832.1
AGC1	NM_001135
AGL	CCDS759.1
AGPAT1	CCDS4744.1
AGPS	CCDS2275.1
AGRN	NM_198576
AHDC1	NM_001029882
AHI1	NM_017651
AIM1L	NM_017977
AKAP11	CCDS9383.1
AKAP13	NM_007200
AKAP4	CCDS14329.1
AKAP9	CCDS5622.1
AKNA	CCDS6805.1
AKR7A2	CCDS194.1
ALDH18A1	CCDS7443.1
ALDH1A2	CCDS10163.1
ALDH1L1	CCDS3034.1
ALDH2	CCDS9155.1
ALLC	NM_018436
ALOX12	CCDS11084.1
ALOXE3	CCDS11130.1
ALPI	CCDS2492.1
ALPK2	CCDS11966.1
ALPK3	CCDS10333.1
ALPL	CCDS217.1
ALS2CL	CCDS2743.1
ALS2CR12	CCDS2346.1
AMACO	CCDS7589.1
AMID	CCDS7297.1
ANK2	CCDS3702.1
ANK3	CCDS7258.1
ANKMY1	CCDS2536.1
ANKRD10	CCDS9520.1
ANKRD11	NM_013275
ANKRD12	CCDS11843.1
ANKRD15	CCDS6441.1
ANKRD28	NM_015199
ANP32D	NM_012404
AP3B1	CCDS4041.1
APG7L	CCDS2605.1
API5	NM_006595
APOB	CCDS1703.1
APOBEC3G	CCDS13984.1
APRG1	NM_178339
AQP10	CCDS1065.1
AR	CCDS14387.1
ARD1B	ENST00000286794
ARHGAP4	CCDS14736.1
ARHGAP5	NM_001173
ARHGAP8	CCDS14058.1
ARHGDIG	CCDS10404.1
ARHGEF9	NM_015185
ARID1A	CCDS285.1
ARL1	NM_001177
ARNT2	NM_014862
ARP10	CCDS13985.1
ARSE	CCDS14122.1
ASB4	CCDS5641.1
ASCL4	NM_203436
ASCL5	ENST00000344317
ASGR1	CCDS11089.1
ASH1L	CCDS1113.1
ASIP	CCDS13232.1
ASTN	CCDS1319.1
ATAD2B	ENST00000295142
ATP10B	ENST00000327245
ATP12A	NM_001676
ATP13A1	NM_020410
ATP13A2	CCDS175.1
ATP1A2	CCDS1196.1
ATP2A1	CCDS10643.1
ATP2A3	CCDS11041.1
ATP2B1	CCDS9035.1
ATP2B2	CCDS2601.1
ATP6V1G3	CCDS1396.1
ATP7B	NM_000053
ATP8A1	CCDS3466.1
ATP8B1	CCDS11965.1
ATRNL1	CCDS7592.1
ATXN1	NM_000332
AUTS2	CCDS5539.1
AXIN2	CCDS11662.1
AZI1	NM_001009811
B3Gn-T6	NM_138706
BAD	CCDS8065.1
BAI2	CCDS346.1
BAMBI	CCDS7162.1
BAT2D1	CCDS1296.1
BAZ1A	CCDS9651.1
BCAR3	CCDS745.1
BCL2L1	CCDS13188.1
BCL2L12	CCDS12776.1
BCL2L2	CCDS9591.1
BCL6	CCDS3289.1

Gene symbol	Accession ID
BCOR	CCDS14250.1
BFSP1	CCDS13126.1
BIN1	CCDS2137.1
BIRC1	CCDS4009.1
BIRC6	NM_016252
BMP3	CCDS3588.1
BMPER	CCDS5442.1
BNC2	CCDS6482.1
BOC	CCDS2971.1
BPY2IP1	NM_018174
BRAF	CCDS5863.1
BRF1	CCDS10001.1
BRP44L	CCDS5293.1
BRPF1	CCDS2575.1
BSN	CCDS2800.1
BST1	CCDS3416.1
BTAF1	CCDS7419.1
BTBD1	CCDS10322.1
BTBD3	CCDS13113.1
BTC	CCDS3566.1
BTK	CCDS14482.1
BTNL2	CCDS4749.1
BTNL9	CCDS4460.1
BUCS1	CCDS10587.1
C10orf18	ENST00000263123
C10orf26	CCDS7540.1
C10orf33	CCDS7474.1
C10orf47	CCDS7085.1
C10orf64	ENST00000265453
C10orf71	ENST00000323868
C10orf80	NM_001008723
C10orf81	CCDS7583.1
C11orf11	NM_006133
C11ORF4	CCDS8066.1
C12orf11	CCDS8708.1
C12orf42	NM_198521
C14orf115	CCDS9830.1
C14orf131	NM_018335
C14orf133	CCDS9862.1
C14orf145	NM_152446
C14orf155	CCDS9679.1
C14orf159	NM_024952
C14orf31	CCDS9704.1
C14orf43	CCDS9819.1
C14orf49	CCDS9935.1
C15orf2	CCDS10015.1
C15orf42	ENST00000268138
C16orf9	CCDS10402.1
C17orf27	NM_020914
C17orf31	CCDS11016.1
C18orf25	NM_001008239
C18orf4	CCDS11995.1
C19orf29	ENST00000221899

Gene symbol	Accession ID
C1orf147	NM_001025592
C1orf151	NM_001032363
C1orf16	CCDS1355.1
C1orf173	NM_001002912
C1orf84	NM_015284
C1QDC1	CCDS8720.1
C20orf10	CCDS13352.1
C20orf102	CCDS13299.1
C20orf114	CCDS13218.1
C20orf23	CCDS13122.1
C20orf78	ENST00000278779
C21orf29	CCDS13712.1
C21orf5	CCDS13643.1
C21orf69	NM_058189
C2orf17	CCDS2434.1
C2orf29	CCDS2050.1
C2orf3	CCDS1961.1
C3orf14	CCDS2896.1
C4orf7	CCDS3537.1
C5AR1	NM_001736
C6	CCDS3936.1
C6orf103	ENST00000326929
C6orf150	CCDS4978.1
C6orf163	NM_001010868
C6orf165	CCDS5009.1
C6orf168	NM_032511
C6orf170	NM_152730
C6orf21	NM_001003693
C6orf213	NM_001010852
C6orf29	CCDS4724.1
C6orf4	CCDS5092.1
C6orf68	CCDS5118.1
C7orf16	CCDS5436.1
C8A	CCDS606.1
C8B	NM_000066
C8orf77	NM_001039382
C8ORFK23	NM_001039112
C9orf126	NM_173690
C9orf19	CCDS6598.1
C9orf5	NM_032012
C9orf50	NM_199350
CA2	CCDS6239.1
CAB39	CCDS2478.1
CABIN1	CCDS13823.1
CABP1	CCDS9204.1
CACNA1A	NM_000068
CACNA1C	NM_000719
CACNA1E	NM_000721
CACNA1H	NM_021098
CACNA1I	NM_001003406
CACNA1S	CCDS1407.1
CACNA2D3	NM_018398
CACNB2	CCDS7125.1

Gene symbol	Accession ID
CACNG4	CCDS11667.1
CADPS	CCDS2898.1
CADPS2	NM_017954
CALM1	CCDS9892.1
CAMSAP1	NM_015447
CAPN12	CCDS12519.1
CAPN3	CCDS10084.1
CAPN3	CCDS10084.1
CAPZA3	CCDS8681.1
CARD11	CCDS5336.1
CART1	CCDS9028.1
CASC5	NM_170589
CASQ1	CCDS1198.1
CCDC15	NM_025004
CCNF	CCDS10467.1
CCNL2	ENST00000321423
CCNYL1	ENST00000339882
CD19	CCDS10644.1
CD84	CCDS1206.1
CD96	CCDS2958.1
CDA08	CCDS10728.1
CDC2L6	CCDS5085.1
CDC7	CCDS734.1
CDCA8	CCDS424.1
CDH23	NM_022124
CDH24	CCDS9585.1
CDH26	CCDS13485.1
CDH5	CCDS10804.1
CDK5	NM_004935
CDK6	CCDS5628.1
CDT1	NM_030928
CDX1	CCDS4304.1
CDYL2	NM_152342
CEACAM1	CCDS12609.1
CELSR3	CCDS2775.1
CENPF	NM_016343
CENTG3	NM_031946
CEP135	NM_025009
Cep164	NM_014956
CEP2	CCDS13255.1
CETP	CCDS10772.1
CFTR	CCDS5773.1
CGI-38	CCDS10835.1
CGI-96	CCDS14036.1
CGNL1	CCDS10161.1
CHAD	CCDS11568.1
CHD4	CCDS8552.1
CHD5	CCDS57.1
CHD6	CCDS13317.1
CHD9	NM_025134
CHDH	CCDS2873.1
CHEK1	CCDS8459.1
ChGn	CCDS6010.1

Gene symbol	Accession ID
CHKA	CCDS8178.1
CHL1	CCDS2556.1
CHRM2	CCDS5843.1
CHRM5	CCDS10031.1
CHRNA3	CCDS10305.1
CHRNA4	CCDS13517.1
CHRNA9	CCDS3459.1
CHST13	CCDS3039.1
CIDEA	CCDS11856.1
CIDEC	CCDS2587.1
CIZ1	CCDS6894.1
CKLFSF5	CCDS9599.1
CLASP1	NM_015282
CLASP2	NM_015097
CLCN1	CCDS5881.1
CLCN5	CCDS14328.1
CLDN11	CCDS3213.1
CLEC1A	CCDS8612.1
CLEC4E	CCDS8594.1
CLEC7A	CCDS8613.1
CLIC6	CCDS13638.1
CLN8	CCDS5956.1
CLSPN	CCDS396.1
CLSTN2	CCDS3112.1
CLTA	CCDS6600.1
CMIP	NM_198390
CMYA1	CCDS2683.1
CMYA4	CCDS11292.1
CNNM2	CCDS7543.1
CNOT1	CCDS10799.1
CNOT10	CCDS2655.1
CNOT7	CCDS6000.1
CNR2	CCDS245.1
CNTN4	CCDS2558.1
CNTNAP2	CCDS5889.1
COCH	CCDS9640.1
COG5	CCDS5742.1
COG5	CCDS5742.1
COH1	CCDS6280.1
COL14A1	NM_021110
COL18A1	NM_030582
COL23A1	CCDS4436.1
COL24A1	NM_152890
COL3A1	CCDS2297.1
COL4A2	NM_001846
COL4A4	NM_000092
COL4A5	CCDS14543.1
COL5A3	CCDS12222.1
COL6A3	NM_004369
COL6A3	NM_057167
COL8A2	CCDS403.1
COPB	CCDS7815.1
COQ2	NM_015697

Gene symbol	Accession ID
CPB1	NM_001871
CPN1	CCDS7486.1
CPNE2	CCDS10774.1
CPNE4	CCDS3072.1
CPS1	CCDS2393.1
CPSF4	CCDS5664.1
CPT1B	CCDS14098.1
CPT1C	CCDS12779.1
CRA	CCDS942.1
CRAT	CCDS6919.1
CREB1	CCDS2374.1
CRIM2	ENST00000257704
CRISPLD1	CCDS6219.1
CRR9	CCDS3862.1
CRX	CCDS12706.1
CRY2	CCDS7915.1
CRYAA	CCDS13695.1
CSK	CCDS10269.1
CSMD1	NM_033225
CSN3	CCDS3538.1
CSNK2A2	CCDS10794.1
CSPG2	CCDS4060.1
CSPG5	CCDS2757.1
CSPG6	NM_005445
CSTF1	CCDS13452.1
CTEN	CCDS11368.1
CTNNA2	NM_004389
CTNNA3	CCDS7269.1
CTSW	CCDS8117.1
CUBN	CCDS7113.1
CUGBP1	CCDS7938.1
CUGBP1	CCDS7939.1
CUL4B	NM_003588
CUTL1	CCDS5721.1
CX40.1	CCDS7191.1
CXCR3	CCDS14416.1
CXorf17	CCDS14356.1
CXorf20	CCDS14184.1
CXorf27	ENST00000341016
CXorf37	CCDS14322.1
CXXC5	NM_016463
CYBB	CCDS14242.1
CYP26C1	CCDS7425.1
CYP2C19	CCDS7436.1
CYP2R1	CCDS7818.1
CYP4F12	NM_023944
DAB2IP	CCDS6832.1
DCBLD2	NM_080927
DCC	CCDS11952.1
DCT	CCDS9470.1
DCTN4	CCDS4310.1
DDB1	NM_001923
DDR1	CCDS4690.1

Gene symbol	Accession ID
DDX1	CCDS1686.1
DDX31	CCDS6951.1
DDX54	NM_024072
DEFB112	NM_001037498
DEFB125	CCDS12989.1
DELGEF	CCDS7828.1
DEPDC5	NM_014662
DFNB31	CCDS6806.1
DGCR6	CCDS13753.1
DGKD	CCDS2504.1
DHPS	CCDS12276.1
DHX29	NM_019030
DIO3	NM_001362
DKFZp434I099	CCDS10787.1
DKFZp547A023	CCDS845.1
DKFZp547B1713	CCDS1591.1
DKFZP564B1023	CCDS1403.1
DKFZp564I1922	CCDS14124.1
DKFZp761L1417	CCDS5658.1
DKFZp761N1114	CCDS1455.1
DLD	CCDS5749.1
DLEC1	ENST00000337335
DLGAP2	NM_004745
DMN	NM_015286
DMTF1	CCDS5601.1
DNAH1	NM_015512
DNAH10	CCDS9255.1
DNAH11	NM_003777
DNAH3	CCDS10594.1
DNAH5	CCDS3882.1
DNAH8	CCDS4838.1
DNAH9	CCDS11160.1
DNAI2	CCDS11697.1
DNCH1	CCDS9966.1
DNCLI2	CCDS10818.1
DNHD3	NM_020877
DNTTIP1	CCDS13369.1
DOCK4	NM_014705
DOCK8	CCDS6440.1
DOCK9	NM_015296
DOK6	NM_152721
DONSON	CCDS13632.1
DRCTNNB1A	CCDS5377.1
DRD3	CCDS2978.1
DRG1	CCDS13897.1
DSG1	CCDS11896.1
DSG2	NM_001943
DSG3	CCDS11898.1
DSG4	CCDS11897.1
DSPP	NM_014208
DST	CCDS4959.1
DTX1	CCDS9164.1
DTX4	ENST00000227451

Gene symbol	Accession ID
DULLARD	CCDS11093.1
DUSP22	CCDS4468.1
DUSP3	CCDS11469.1
DYRK3	NM_001004023
DZIP3	CCDS2952.1
E2F4	NM_001950
EAF1	CCDS2626.1
EBF	CCDS4343.1
EBF3	NM_001005463
ECEL1	CCDS2493.1
ECHDC2	CCDS571.1
ECOP	NM_030796
EDD1	NM_015902
EDG3	CCDS6680.1
EDG8	CCDS12240.1
EEF1A1	ENST00000331523
EFCBP1	NM_022351
EFHC2	NM_025184
EGF	CCDS3689.1
EGFR	CCDS5514.1
EHBP1L1	ENST00000309295
EIF2A	NM_032025
EIF3S12	CCDS12517.1
EIF4G1	CCDS3259.1
EIF4G2	NM_001418
EME2	NM_001010865
EML4	CCDS1807.1
EMR4	ENST00000359590
EN2	CCDS5940.1
ENO1	CCDS97.1
ENPP2	CCDS6329.1
ENPP6	CCDS3834.1
ENPP7	CCDS11763.1
ENSA	CCDS958.1
ENST00000294635	ENST00000294635
ENST00000310882	ENST00000310882
ENST00000326382	ENST00000326382
ENST00000328067	ENST00000328067
ENST00000331583	ENST00000331583
ENST00000334627	ENST00000334627
ENST00000336168	ENST00000336168
ENST00000355177	ENST00000355177
ENST00000355324	ENST00000355324
ENST00000355607	ENST00000355607
ENST00000357689	ENST00000357689
ENST00000358347	ENST00000358347
ENST00000359736	ENST00000359736
EPB41L2	CCDS5141.1
EPB41L4B	NM_019114
EPB49	CCDS6020.1
EPC1	CCDS7172.1
EPHA2	CCDS169.1
EPHA5	CCDS3513.1

Gene symbol	Accession ID
EPHA6	ENST00000334709
EPHA8	CCDS225.1
EPO	CCDS5705.1
ERCC5	NM_000123
ERF	CCDS12600.1
ERN1	NM_001433
ESCO2	NM_001017420
ESPNP	ENST00000270691
ESR1	CCDS5234.1
ESR2	CCDS9762.1
ETV1	NM_004956
EVI1	CCDS3205.1
EVPL	CCDS11737.1
EXOC6B	ENST00000272427
EXTL1	CCDS271.1
F13B	CCDS1388.1
F2RL1	CCDS4033.1
F3	CCDS750.1
F5	CCDS1281.1
FAD158	CCDS725.1
FADS1	CCDS8011.1
FAM43A	NM_153690
FAM46B	CCDS294.1
FAM47A	NM_203408
FAM48A	ENST00000360252
FAM63B	NM_019092
FAM78B	NM_001017961
FAM92B	NM_198491
FANCA	NM_000135
FANCD2	CCDS2595.1
FASN	CCDS11801.1
FAT	NM_005245
FBN3	CCDS12196.1
FBXO40	NM_016298
FBXW7	CCDS3777.1
FCGBP	CCDS12546.1
FCHSD1	NM_033449
FECH	CCDS11964.1
FEZ1	NM_005103
FGD1	CCDS14359.1
FGD4	CCDS8727.1
FGF2	NM_002006
FGFR3	CCDS3353.1
FGIF	CCDS8300.1
FIGF	CCDS14166.1
FLII	CCDS11192.1
FLJ10276	CCDS363.1
FLJ10514	CCDS1311.1
FLJ11088	CCDS8716.1
FLJ11535	CCDS12043.1
FLJ12529	CCDS8006.1
FLJ12644	CCDS12843.1
FLJ12671	CCDS1153.1

Gene symbol	Accession ID
FLJ12700	CCDS5898.1
FLJ13273	CCDS3672.1
FLJ13576	CCDS5757.1
FLJ13725	CCDS10840.1
FLJ13841	CCDS11819.1
FLJ13941	CCDS40.1
FLJ14397	CCDS1945.1
FLJ16165	NM_001004318
FLJ16331	NM_001004326
FLJ16478	NM_001004341
FLJ20035	NM_017631
FLJ20097	ENST00000317751
FLJ20186	CCDS10989.1
FLJ20232	CCDS13995.1
FLJ20272	NM_017735
FLJ20294	NM_017749
FLJ20298	CCDS14522.1
FLJ21159	CCDS3792.1
FLJ21963	CCDS9022.1
FLJ22709	CCDS12351.1
FLJ23049	CCDS3199.1
FLJ23447	CCDS12300.1
FLJ23577	ENST00000303168
FLJ23577	CCDS3910.1
FLJ23790	CCDS6346.1
FLJ25715	NM_182570
FLJ25801	CCDS3850.1
FLJ27465	NM_001039843
FLJ30525	CCDS787.1
FLJ30655	CCDS3740.1
FLJ30707	CCDS9427.1
FLJ31438	NM_152385
FLJ32796	CCDS1507.1
FLJ32934	CCDS1082.1
FLJ33167	CCDS3837.1
FLJ33387	CCDS9783.1
FLJ34512	CCDS10424.1
FLJ34658	CCDS3913.1
FLJ35709	CCDS7767.1
FLJ35728	CCDS1537.1
FLJ36004	CCDS8704.1
FLJ36208	NM_145270
FLJ36601	CCDS14238.1
FLJ37440	CCDS2095.1
FLJ38964	NM_173527
FLJ38973	NM_153689
FLJ39058	CCDS8489.1
FLJ39198	NM_001039769
FLJ39873	CCDS2980.1
FLJ40243	NM_173489
FLJ40342	CCDS11512.1
FLJ40869	CCDS1691.1
FLJ41170	NM_001004332

Gene symbol	Accession ID
FLJ41766	ENST00000338573
FLJ43706	NM_001039774
FLJ44186	CCDS5854.1
FLJ44861	CCDS11778.1
FLJ45300	NM_001001681
FLJ45744	CCDS12424.1
FLJ45964	CCDS2530.1
FLJ45974	NM_001001707
FLJ46072	CCDS6410.1
FLJ90650	CCDS4124.1
FLT1	CCDS9330.1
FMN2	NM_020066
FMNL2	NM_001004417
FN1	CCDS2399.1
FNBP1	NM_015033
FNDC1	NM_032532
FOXA2	CCDS13147.1
FOXB1	NM_012182
FOXI1	CCDS4372.1
FOXM1	CCDS8515.1
FOXR2	NM_198451
FRAS1	NM_025074
FREM2	NM_207361
FRMD3	NM_174938
FRMD4B	ENST00000264546
FRMPD1	CCDS6612.1
FRMPD4	NM_014728
FSD2	NM_001007122
FSTL1	CCDS2998.1
FSTL4	NM_015082
FSTL5	CCDS3802.1
FUBP1	CCDS683.1
FUT2	NM_000511
FXYD6	CCDS8387.1
FYCO1	CCDS2734.1
FZD10	CCDS9267.1
FZD3	CCDS6069.1
FZD6	CCDS6298.1
FZD9	CCDS5548.1
G3BP2	CCDS3571.1
GABPA	CCDS13575.1
GABRA6	CCDS4356.1
GABRD	CCDS36.1
GAD2	CCDS7149.1
GALNT13	CCDS2199.1
GALNT3	CCDS2226.1
GALNT7	CCDS3815.1
GALNTL1	NM_020692
GANAB	CCDS8026.1
GAPVD1	NM_015635
GAS6	CCDS9540.1
GATA4	CCDS5983.1
GATA6	CCDS11872.1

Gene symbol	Accession ID
GBF1	CCDS7533.1
GCGR	NM_000160
GCM1	CCDS4950.1
GCM2	CCDS4517.1
GCNT3	CCDS10172.1
GDF3	CCDS8581.1
GEFT	CCDS8947.1
GFI1B	CCDS6957.1
GFM1	NM_024996
GGA2	CCDS10611.1
GGPS1	CCDS1604.1
GHSR	CCDS3218.1
GIMAP1	CCDS5906.1
GIMAP5	CCDS5907.1
GIMAP8	NM_175571
GIT2	CCDS9138.1
GJA4	NM_002060
GJB4	CCDS383.1
GK	CCDS14225.1
GLRA1	CCDS4320.1
GMCL1L	CCDS4433.1
GMDS	CCDS4474.1
GML	CCDS6391.1
GNAI2	CCDS2813.1
GNAT1	CCDS2812.1
GNL2	CCDS421.1
GNPTG	CCDS10436.1
GNS	CCDS8970.1
GOLGA3	CCDS9281.1
GOLGA4	CCDS2666.1
GORASP2	NM_015530
GOT2	CCDS10801.1
GP6	NM_016363
GPBP1	NM_022913
GPI7	CCDS3336.1
GPR114	CCDS10785.1
GPR116	CCDS4919.1
GPR132	CCDS9997.1
GPR142	CCDS11698.1
GPR144	NM_182611
GPR145	CCDS5044.1
GPR174	CCDS14443.1
GPR37	CCDS5792.1
GPR37L1	CCDS1420.1
GPR40	CCDS12458.1
GPR43	CCDS12461.1
GPR61	CCDS801.1
GPR73L1	CCDS13089.1
GPR74	CCDS3551.1
GPR78	CCDS3403.1
GPR83	CCDS8297.1
GPR85	CCDS5758.1
GPRC5C	CCDS11699.1

Gene symbol	Accession ID
GPS1	CCDS11800.1
GPS2	NM_032442
GPSM2	CCDS792.1
GPT	CCDS6430.1
GRAP2	CCDS13999.1
GRASP	CCDS8817.1
GRCA	CCDS8563.1
GREB1	NM_014668
GRIA4	CCDS8333.1
GRIK4	CCDS8433.1
GRIN2B	CCDS8662.1
GRIN3A	CCDS6758.1
GRINA	NM_001009184
GRM1	CCDS5209.1
GRM3	CCDS5600.1
GSR	NM_000637
GSTO2	CCDS7556.1
GTF2A2	CCDS10173.1
GTF2H4	NM_020442
GTF3C4	CCDS6953.1
GUCY1A3	NM_000856
GUCY1B2	CCDS9426.1
GZMH	CCDS9632.1
HAMP	CCDS12454.1
HBB	NM_000519
HBXAP	CCDS8253.1
HCFC2	CCDS9097.1
HDAC2	NM_001527
HDAC9	NM_178425
HDC	CCDS10134.1
HECW2	NM_020760
HERC1	NM_003922
HERC2	CCDS10021.1
HGSNAT	ENST00000332689
HHIP	CCDS3762.1
HIF3A	CCDS12681.1
HIP1	NM_005338
HIVEP1	NM_002114
HIVEP2	NM_006734
HIVEP3	CCDS463.1
HMG20A	CCDS10295.1
HMGCL	CCDS243.1
HMP19	CCDS4391.1
HNT	CCDS8491.1
HORMAD1	CCDS967.1
HOXA6	CCDS5407.1
HP	NM_005143
HP1BP3	NM_016287
HPCAL4	CCDS441.1
HRB	CCDS2467.1
HRBL	CCDS5697.1
HRG	CCDS3280.1
HS2ST1	CCDS712.1

Gene symbol	Accession ID
HS2ST1	CCDS711.1
HSA9761	CCDS3981.1
HSD17B2	CCDS10936.1
HSD17B8	CCDS4769.1
HSPA4L	CCDS3734.1
HSPC111	NM_016391
HSPG2	NM_005529
HTR3C	CCDS3250.1
HTR3E	CCDS3251.1
HXMA	CCDS10586.1
HYPB	CCDS2749.1
IBTK	NM_015525
ICAM3	CCDS12235.1
ICEBERG	NM_021571
IDE	CCDS7421.1
IDH1	CCDS2381.1
IFI44	CCDS688.1
IFIT3	CCDS7402.1
IFNAR1	CCDS13624.1
IFRD1	NM_001007245
IGF1	CCDS9091.1
IGF2	CCDS7728.1
IGFBP7	CCDS3512.1
IGSF1	CCDS14629.1
IGSF10	CCDS3160.1
IGSF9	CCDS1190.1
IKBKE	NM_014002
IL12RB2	CCDS638.1
IL17B	CCDS4297.1
IL17RE	CCDS2589.1
IL1F9	CCDS2108.1
IL1RL1	CCDS2057.1
IL3	CCDS4149.1
ILT7	CCDS12890.1
IMP4	CCDS2160.1
IMPDH1	NM_183243
INDO	NM_002164
INSIG2	CCDS2122.1
IPO13	CCDS503.1
IPO8	CCDS8719.1
IQGAP2	NM_006633
IQWD1	CCDS1267.1
IRS1	CCDS2463.1
IRTA2	CCDS1165.1
IRX6	NM_024335
ISL1	NM_002202
ITGA4	NM_000885
ITGA7	CCDS8888.1
ITGAL	NM_002209
ITGAX	CCDS10711.1
ITIH5	NM_032817
ITLN1	CCDS1211.1
ITPKB	CCDS1555.1

Gene symbol	Accession ID
ITPR3	CCDS4783.1
IVNS1ABP	CCDS1368.1
JMJD1A	CCDS1990.1
JMJD1B	NM_016604
JUNB	CCDS12280.1
K0574_HUMAN	ENST00000261275
KATNAL2	NM_031303
KBTBD3	CCDS8334.1
KBTBD4	CCDS7940.1
KCNA4	NM_002233
KCNA7	CCDS12755.1
KCNB2	CCDS6209.1
KCNC4	CCDS821.1
KCND2	CCDS5776.1
KCNG3	CCDS1809.1
KCNH1	CCDS1496.1
KCNH5	CCDS9756.1
KCNJ15	CCDS13656.1
KCNK1	CCDS1599.1
KCNK5	CCDS4841.1
KCNN1	NM_002248
KCNQ3	NM_004519
KCNQ4	CCDS456.1
KCTD7	CCDS5534.1
KCTD8	CCDS3467.1
KDELR2	CCDS5351.1
KDR	CCDS3497.1
KEL	NM_000420
KIAA0082	CCDS4835.1
KIAA0101	CCDS10193.1
KIAA0103	CCDS6309.1
KIAA0133	NM_014777
KIAA0143	NM_015137
KIAA0153	CCDS14047.1
KIAA0317	NM_001039479
KIAA0329	NM_014844
KIAA0350	NM_015226
KIAA0367	NM_015225
KIAA0404	NM_015104
KIAA0406	CCDS13300.1
KIAA0528	NM_014802
KIAA0649	CCDS6988.1
KIAA0652	CCDS7921.1
KIAA0664	NM_015229
KIAA0672	NM_014859
KIAA0690	CCDS7457.1
KIAA0701	NM_001006947
KIAA0703	NM_014861
KIAA0748	ENST00000316577
KIAA0759	CCDS9852.1
KIAA0774	NM_001033602
KIAA0802	CCDS11841.1
KIAA0831	NM_014924

Gene symbol	Accession ID
KIAA0863	NM_014913
KIAA0980	NM_025176
KIAA1024	NM_015206
KIAA1033	NM_015275
KIAA1086	ENST00000262961
KIAA1109	ENST00000264501
KIAA1223	NM_020337
KIAA1274	NM_014431
KIAA1328	NM_020776
KIAA1377	NM_020802
KIAA1411	NM_020819
KIAA1441	CCDS992.1
KIAA1467	NM_020853
KIAA1505	NM_020879
KIAA1524	NM_020890
KIAA1576	NM_020927
KIAA1618	CCDS11772.1
KIAA1754L	NM_178495
KIAA1804	CCDS1598.1
KIAA1862	NM_032534
KIAA1909	NM_052909
KIAA1946	NM_177454
KIAA1967	NM_021174
KIAA2022	NM_001008537
KIAA2026	NM_001017969
KIDINS220	NM_020738
KIFC2	CCDS6427.1
KIFC3	CCDS10789.1
KIRREL2	CCDS12479.1
KIRREL3	NM_032531
KLHDC5	NM_020782
KLHL10	NM_152467
KLHL4	CCDS14456.1
KLK9	CCDS12816.1
KLP1	CCDS12926.1
KLRG1	CCDS8599.1
KNTC1	NM_014708
KREMEN2	CCDS10484.1
KREMEN2	CCDS10483.1
KRT9	NM_000226
KRTAP12-3	NM_198697
KRTAP20-2	CCDS13604.1
KRTHA4	CCDS11390.1
KSR1	NM_014238
L1CAM	CCDS14733.1
L3MBTL2	CCDS14011.1
LACE1	CCDS5067.1
LACRT	CCDS8883.1
LAMA1	NM_005559
LAMA3	CCDS11880.1
LAMA4	NM_002290
LAMB3	CCDS1487.1
LAMP3	CCDS3242.1

Gene symbol	Accession ID
LAP1B	CCDS1335.1
LARGE	CCDS13912.1
LARP5	NM_015155
LATS1	NM_004690
LATS2	CCDS9294.1
LAX	CCDS1441.1
LBP	CCDS13304.1
LCA10	NM_001039768
LCT	CCDS2178.1
LDLRAD3	NM_174902
LEMD2	CCDS4785.1
LENG8	CCDS12894.1
LETM1	CCDS3355.1
LETMD1	CCDS8806.1
LIP8	CCDS11126.1
LIPM	ENST00000282673
LMNB1	CCDS4140.1
LMX1A	CCDS1247.1
LNX	CCDS3492.1
LNX2	CCDS9323.1
LOC113655	CCDS6431.1
LOC124842	CCDS11283.1
LOC126248	CCDS12429.1
LOC131368	CCDS2947.1
LOC131873	ENST00000358511
LOC134145	NM_199133
LOC146562	CCDS10521.1
LOC158830	NM_001025265
LOC200312	NM_001017981
LOC221955	CCDS5350.1
LOC257106	CCDS1215.1
LOC283537	CCDS9332.1
LOC284912	CCDS13918.1
LOC284948	CCDS1976.1
LOC339977	NM_001024611
LOC374768	NM_199339
LOC387755	NM_001031853
LOC387856	NM_001013635
LOC388595	NM_001013641
LOC388969	NM_001013649
LOC391123	NM_001013661
LOC392617	ENST00000333066
LOC400707	NM_001013673
LOC441136	NM_001013719
LOC441233	NM_001013724
LOC442213	NM_001013732
LOC494115	NM_001008662
LOC51058	CCDS476.1
LOC54103	NM_017439
LOC54499	CCDS1251.1
LOC550631	NM_001017437
LOC63928	CCDS10617.1
LOC643866	NM_001039771

Gene symbol	Accession ID
LOC648272	ENST00000343945
LOC651746	ENST00000296657
LOC651863	ENST00000333744
LOC90379	NM_138353
LOC90826	CCDS3771.1
LOC92154	NM_138383
LOC93349	NM_138402
LPAL2	ENST00000342479
LPHN1	CCDS12307.1
LPHN2	CCDS689.1
LPHN3	NM_015236
LPIN3	NM_022896
LPL	CCDS6012.1
LRAT	CCDS3789.1
LRCH1	NM_015116
LRFN5	CCDS9678.1
LRP1	CCDS8932.1
LRP10	CCDS9578.1
LRP1B	CCDS2182.1
LRP2	CCDS2232.1
LRRC16	NM_017640
LRRC4	CCDS5799.1
LRRC4B	ENST00000253728
LRRC7	CCDS645.1
LRRIQ1	NM_032165
LRRK1	NM_024652
LRRN1	NM_020873
LRRN3	CCDS5754.1
LRRN5	CCDS1448.1
LTB4R2	CCDS9624.1
LTBP1	NM_000627
LTBP3	CCDS8103.1
LTBP4	NM_003573
LTK	CCDS10077.1
LUC7L	CCDS10401.1
LY6K	CCDS6385.1
LYNX1	ENST00000317543
LYPLA1	CCDS6157.1
LYRIC	CCDS6274.1
LYST	NM_000081
LYZL4	CCDS2697.1
LZTR2	NM_033127
M160	CCDS8577.1
MACF1	CCDS435.1
MAEA	NM_001017405
MAGEA4	CCDS14702.1
MAGEB10	NM_182506
MAGEC1	NM_005462
MAGEH1	CCDS14369.1
MAGI-3	CCDS859.1
MAK10	CCDS6673.1
MALT1	CCDS11967.1
MAMDC2	CCDS6631.1

Gene symbol	Accession ID
MAN1B1	CCDS7029.1
MAN2A1	NM_002372
MAN2B1	NM_000528
MAP1B	CCDS4012.1
MAP3K11	CCDS8107.1
MAP3K14	NM_003954
MAP3K8	CCDS7166.1
MAP3K9	NM_033141
MAP4K4	NM_004834
MAP7D3	ENST00000218318
MARCO	CCDS2124.1
MARK3	NM_002376
MARS	CCDS8942.1
MARS2	NM_138395
MASS1	NM_032119
MAST4	ENST00000261569
MATN1	CCDS336.1
MBD1	CCDS11941.1
MBNL1	CCDS3163.1
MCCC1	CCDS3241.1
MCF2L	ENST00000261963
MCFD2	NM_139279
MCM10	CCDS7095.1
MCPH1	NM_024596
MDGA1	NM_153487
MDH2	CCDS5581.1
MEA	CCDS4879.1
MED12	NM_005120
MEFV	CCDS10498.1
MEN1	CCDS8083.1
METTL5	NM_014168
MGAM	NM_004668
MGC16635	CCDS14097.1
MGC19764	NM_144975
MGC20419	CCDS562.1
MGC20741	CCDS4861.1
MGC21830	CCDS10463.1
MGC24039	NM_144973
MGC2655	CCDS10491.1
MGC26598	CCDS9036.1
MGC26818	CCDS44.1
MGC27016	CCDS3790.1
MGC29814	CCDS11742.1
MGC29875	CCDS1493.1
MGC33367	CCDS10738.1
MGC33414	CCDS279.1
MGC33486	CCDS8133.1
MGC33889	CCDS14216.1
MGC34647	CCDS10895.1
MGC35118	CCDS10046.1
MGC35194	CCDS147.1
MGC35366	CCDS9057.1
MGC39581	CCDS12149.1

Gene symbol	Accession ID
MGC42174	NM_152383
MGC4251	CCDS11474.1
MGC4268	CCDS2152.1
MGC45562	CCDS11371.1
MGC45780	CCDS6064.1
MGC47869	CCDS8667.1
MHC2TA	CCDS10544.1
MIA3	ENST00000320831
MICAL-L2	CCDS5324.1
MINK1	NM_170663
MIPEP	CCDS9303.1
MIR16	CCDS10578.1
MKI67	CCDS7659.1
MLL	NM_005933
MLL3	CCDS5931.1
MLL4	NM_014727
MLLT4	CCDS5303.1
MLLT7	NM_005938
MME	CCDS3172.1
MMP10	CCDS8321.1
MMP16	CCDS6246.1
MOCS1	CCDS4845.1
MON2	NM_015026
MPDU1	CCDS11115.1
MPDZ	NM_003829
MPP1	CCDS14762.1
MPZ	CCDS1229.1
MRC2	CCDS11634.1
MRGX1	CCDS7846.1
MRPL13	CCDS6332.1
MRPL16	CCDS7976.1
MRPL37	ENST00000329505
MRPL44	CCDS2459.1
MRPL46	CCDS10341.1
MRPL55	CCDS1567.1
MRPS5	CCDS2010.1
MRPS7	CCDS11718.1
MRVI1	NM_006069
MS4A7	CCDS7985.1
MSI2	CCDS11596.1
MSL2L1	NM_018133
MSRB3	CCDS8973.1
MTA1	NM_004689
MTHFD2L	NM_001004346
MTNR1B	CCDS8290.1
MTP	CCDS3651.1
MTR	CCDS1614.1
MTX2	CCDS2272.1
MUC15	CCDS7859.1
MUC16	NM_024690
MUC5AC	ENST00000349637
MUC7	CCDS3541.1
MVP	CCDS10656.1

Gene symbol	Accession ID
MYBPC3	NM_000256
MYBPHL	NM_001010985
MYF6	CCDS9019.1
MYH14	NM_024729
MYH15	ENST00000273353
MYH3	CCDS11157.1
MYH4	CCDS11154.1
MYO15A	NM_016239
MYO18B	NM_032608
MYO1B	CCDS2311.1
MYO1D	NM_015194
MYO1E	NM_004998
MYO3A	CCDS7148.1
MYO3B	NM_138995
MYO5A	NM_000259
MYO5C	NM_018728
MYO9B	NM_004145
MYOCD	CCDS11163.1
MYOM1	NM_003803
MYOM2	CCDS5957.1
MYR8	NM_015011
MYRIP	CCDS2689.1
MYST3	CCDS6124.1
MYT1L	NM_015025
NAGA	CCDS14030.1
NALP1	NM_014922
NALP11	CCDS12935.1
NALP7	CCDS12912.1
NAPSB	ENST00000253720
NARG1L	CCDS9379.1
NAV1	CCDS1414.1
NCBP1	CCDS6728.1
NCKAP1L	NM_005337
NCOA5	CCDS13392.1
NCOA6	CCDS13241.1
NDUFA11	CCDS12155.1
NDUFB2	CCDS5862.1
NDUFS6	CCDS3866.1
NEB	NM_004543
NEIL3	CCDS3828.1
NEUROG2	CCDS3698.1
NF1	CCDS11264.1
NFATC3	CCDS10862.1
NFATC4	CCDS9629.1
NGEF	CCDS2500.1
NHS	CCDS14181.1
NIF3L1BP1	CCDS2900.1
NIN	NM_182944
NISCH	NM_007184
NKG7	CCDS12830.1
NKRF	NM_017544
NKX2-5	CCDS4387.1
NLGN1	CCDS3222.1

Gene symbol	Accession ID
NLGN2	CCDS11103.1
NLN	CCDS3989.1
NM_001080470.1	ENST00000271263
NMBR	CCDS5196.1
NMUR1	CCDS2486.1
NNT	CCDS3949.1
NOD3	NM_178844
NOR1	CCDS409.1
NOS3	CCDS5912.1
NOTCH1	NM_017617
NOTCH2	CCDS908.1
NOTCH3	CCDS12326.1
NOTCH4	NM_004557
NOX4	CCDS8285.1
NP_001073909.1	ENST00000327928
NP_001073931.1	ENST00000341689
NP_001073940.1	ENST00000292357
NP_001073948.1	ENST00000296794
NP_001073961.1	ENST00000219301
NP_001073971.1	ENST00000266524
NP_001074294.1	ENST00000342607
NPC1L1	CCDS5491.1
NPL	CCDS1350.1
NPLOC4	NM_017921
NPPA	CCDS139.1
NPR3	NM_000908
NPTXR	NM_014293
NR_002781.1	ENST00000246203
NR2E1	CCDS5063.1
NRAP	CCDS7578.1
NRBP2	NM_178564
NRK	NM_198465
NRP1	CCDS7177.1
NRP2	CCDS2364.1
NRXN2	CCDS8077.1
NS3TP2	CCDS4136.1
NT5E	CCDS5002.1
NTN2L	CCDS10469.1
NTRK3	CCDS10340.1
NUAK1	NM_014840
NUP160	NM_015231
NUP188	NM_015354
NUP205	NM_015135
NUP210L	NM_207308
NUP98	CCDS7746.1
NURIT	CCDS9399.1
NXF3	CCDS14503.1
NXF5	CCDS14491.1
NXPH1	NM_152745
OAS3	NM_006187
OBSCN	CCDS1570.1
ODZ2	ENST00000314238
OLIG2	CCDS13620.1

Gene symbol	Accession ID
OPRD1	CCDS329.1
OPRL1	CCDS13556.1
OR10G3	NM_001005465
OR10G4	NM_001004462
OR10H2	CCDS12333.1
OR10P1	NM_206899
OR10T2	NM_001004475
OR13J1	NM_001004487
OR1L8	NM_001004454
OR2A12	NM_001004135
OR2AG1	NM_001004489
OR2AG2	NM_001004490
OR2D2	NM_003700
OR2G3	NM_001001914
OR2L13	CCDS1637.1
OR2L2	NM_001004686
OR2S2	CCDS6596.1
OR2T4	NM_001004696
OR2V2	CCDS4461.1
OR2Y1	NM_001001657
OR2Z1	NM_001004699
OR3A1	CCDS11023.1
OR4A5	NM_001005272
OR4L1	NM_001004717
OR4N2	NM_001004723
OR4P4	NM_001004124
OR52A5	NM_001005160
OR52B2	NM_001004052
OR52D1	NM_001005163
OR52E6	NM_001005167
OR52I1	NM_001005169
OR52N4	NM_001005175
OR56A4	NM_001005179
OR56B1	NM_001005180
OR56B4	NM_001005181
OR5A1	NM_001004728
OR5AP2	NM_001002925
OR5AU1	NM_001004731
OR5B17	ENST00000357377
OR5BF1	NM_001001918
OR5D14	NM_001004735
OR5K4	NM_001005517
OR5M1	ENST00000303005
OR5M8	NM_001005282
OR5M9	NM_001004743
OR6C74	NM_001005490
OR6K3	NM_001005327
OR6W1P	ENST00000340373
OR7A5	CCDS12318.1
OR7D4	NM_001005191
OR8D2	NM_001002918
OR8K3	NM_001005202
OR9K2	NM_001005243
OR9Q2	NM_001005283
OSAP	NM_032623
OSBPL2	CCDS13494.1
OSBPL5	NM_145638
OSBPL9	CCDS558.1
OSR2	NM_053001
OSTM1	CCDS5062.1
OTOF	CCDS1725.1
OTOG	ENST00000342528
OTOR	CCDS13124.1
OTUD1	ENST00000298035
OVCH1	NM_183378
OVOL1	CCDS8112.1
OXA1L	CCDS9573.1
p44S10	CCDS2901.1
PADI2	CCDS177.1
PAPLN	NM_173462
PAPOLG	CCDS1863.1
PAPPA2	NM_020318
PARC	CCDS4890.1
PARP11	CCDS8523.1
PAX9	CCDS9662.1
PCAF	CCDS2634.1
PCDH11X	CCDS14463.1
PCDHA10	NM_031859
PCDHA13	CCDS4240.1
PCDHB7	CCDS4249.1
PCDHGA4	NM_032053
PCDHGA9	NM_032089
PCDHGB7	NM_032101
PCDHGC4	CCDS4260.1
PCDHGC4	CCDS4261.1
PCDHGC4	CCDS4263.1
PCGF2	NM_007144
PCNXL2	ENST00000344698
PCSK2	CCDS13125.1
PCYOX1	CCDS1902.1
PDCD10	CCDS3202.1
PDCD11	NM_014976
PDE1C	CCDS5437.1
PDE4A	CCDS12238.1
PDE4B	CCDS632.1
PDE4C	CCDS12373.1
PDE4D	NM_006203
PDGFB	CCDS13987.1
PDGFRA	CCDS3495.1
PDGFRB	CCDS4303.1
PDHA2	CCDS3644.1
PDHB	CCDS2890.1
PDIA2	NM_006849
PDK1	CCDS2250.1
PDLIM4	CCDS4152.1
PDZD2	NM_178140
PDZD7	NM_024895
PEG10	ENST00000362013
PELP1	NM_014389
PENK	CCDS6168.1
PERQ1	NM_022574
PEX1	CCDS5627.1
PEX10	CCDS41.1
PFAS	CCDS11136.1
PFKFB3	CCDS7078.1
PGAP1	CCDS2318.1
PGBD5	CCDS1583.1
PHC3	NM_024947
PHEMX	CCDS7733.1
PHF2	ENST00000298216
PHF21A	NM_016621
PHIP	CCDS4987.1
PHKA2	CCDS14190.1
PHLPP	NM_194449
PHLPPL	NM_015020
PHOX2B	CCDS3463.1
PIGN	NM_176787
PIGQ	CCDS10411.1
PIGR	CCDS1474.1
PIK3C2G	NM_004570
PIK3CA	NM_006218
PIK3CG	CCDS5739.1
PIK3R1	CCDS3993.1
PIK3R4	CCDS3067.1
PIK3R5	CCDS11147.1
PIP5K1A	CCDS990.1
PIP5K3	CCDS2382.1
PISD	CCDS13899.1
PITPNM1	NM_004910
PITPNM2	CCDS9242.1
PITPNM3	CCDS11076.1
PIWIL3	NM_001008496
PKD1	NM_000296
PKD1L2	NM_182740
PKHD1	CCDS4935.1
PKHD1L1	NM_177531
PKIA	CCDS6222.1
PLA1A	CCDS2991.1
PLCH2	NM_014638
PLCXD3	NM_001005473
PLD2	CCDS11057.1
PLEC1	NM_201378
PLEKHA4	CCDS12737.1
PLEKHH2	CCDS1812.1
PLIN	CCDS10353.1
PLSCR3	NM_020360
PLXDC2	CCDS7132.1
PLXNA3	CCDS14752.1
PLXNB2	ENST00000359337

Gene symbol	Accession ID
PLXNC1	CCDS9049.1
PMS1	CCDS2302.1
PMS2L4	ENST00000275546
PNLIP	CCDS7594.1
PNOC	CCDS6066.1
PODXL2	CCDS3044.1
POLD1	CCDS12795.1
POLE	CCDS9278.1
POLG2	NM_007215
POLM	NM_013284
POLR3B	CCDS9105.1
POLR3E	CCDS10605.1
POPDC2	CCDS2992.1
POR	CCDS5579.1
PORCN	CCDS14296.1
POT1	CCDS5793.1
POU1F1	CCDS2919.1
POU2F1	CCDS1259.1
POU6F2	NM_007252
PPAP2C	CCDS12023.1
PPARA	NM_001001930
PPBP	CCDS3563.1
PPEF2	NM_006239
PPIG	CCDS2235.1
PPL	CCDS10526.1
PPM2C	CCDS6259.1
PPP1CC	CCDS9150.1
PPP1R12A	NM_002480
PPP1R12C	CCDS12916.1
PPP2CZ	CCDS855.1
PPP2R2C	CCDS3387.1
PPRC1	CCDS7529.1
PRCC	CCDS1157.1
PRDM16	NM_199454
PRDM5	CCDS3716.1
PRELP	CCDS1438.1
PRIC285	CCDS13527.1
PRKCBP1	CCDS13404.1
PRKCZ	CCDS37.1
PRKDC	NM_006904
PRKG2	CCDS3589.1
PRKRA	CCDS2279.1
PRO1853	CCDS1788.1
PRO1855	CCDS11566.1
PROM1	NM_006017
PROSC	CCDS6096.1
PRPF18	CCDS7100.1
PRR12	ENST00000246798
PRSS16	CCDS4623.1
PRSS22	CCDS10481.1
PSF1	NM_021067
PSIP1	CCDS6479.1
PSMD8	CCDS12515.1
PSRC2	NM_144982
PTAR1	ENST00000340434
PTCH2	CCDS516.1
PTEN	NM_000314
PTGDR	CCDS9707.1
PTGFR	CCDS686.1
PTGS2	CCDS1371.1
PTPLA	CCDS7121.1
PTPN23	CCDS2754.1
PTPRF	CCDS489.1
PTPRK	CCDS5137.1
PTPRM	CCDS11840.1
PTPRS	CCDS12139.1
PTPRU	CCDS334.1
PTX3	CCDS3180.1
PUM1	CCDS338.1
PYGB	CCDS13171.1
Q13034_HUMAN	ENST00000225928
Q4VXG5_HUMAN	ENST00000327794
Q4VXG5_HUMAN	ENST00000331811
Q5JX50_HUMAN	ENST00000325076
Q5JYU7_HUMAN	ENST00000333418
Q5T740_HUMAN	ENST00000343319
Q5W0A0_HUMAN	ENST00000298738
Q68CJ6_HUMAN	ENST00000341513
Q6IEE8_HUMAN	ENST00000354872
Q6PK04_HUMAN	ENST00000329214
Q6RGF6_HUMAN	ENST00000359144
Q6ZRB0_HUMAN	ENST00000297487
Q6ZSY1_HUMAN	ENST00000320930
Q6ZT40_HUMAN	ENST00000296564
Q6ZUG5_HUMAN	ENST00000344062
Q6ZV46_HUMAN	ENST00000341696
Q76B61_HUMAN	ENST00000360022
Q86U37_HUMAN	ENST00000335192
Q86XQ1_HUMAN	ENST00000261673
Q86YU6_HUMAN	ENST00000330768
Q8IUR1_HUMAN	ENST00000327506
Q8N1R6_HUMAN	ENST00000331014
Q8N646_HUMAN	ENST00000359720
Q8N800_HUMAN	ENST00000322516
Q8N822_HUMAN	ENST00000317280
Q8N8C3_HUMAN	ENST00000319889
Q8N8K0_HUMAN	ENST00000301807
Q8N9H1_HUMAN	ENST00000359503
Q8NBE0_HUMAN	ENST00000297801
Q8NDH2_HUMAN	ENST00000322527
Q8NGK8_HUMAN	ENST00000334020
Q8NGL5_HUMAN	ENST00000328673
Q8NH06_HUMAN	ENST00000324144
Q8NHB0_HUMAN	ENST00000315712
Q8TBR1_HUMAN	ENST00000354206
Q96CH6_HUMAN	ENST00000329920
Q96CK5_HUMAN	ENST00000273582
Q96DR3_HUMAN	ENST00000324748
Q96FF7_HUMAN	ENST00000269720
Q96NE0_HUMAN	ENST00000329922
Q96NL2_HUMAN	ENST00000272907
Q96PS2_HUMAN	ENST00000326978
Q9H030_HUMAN	ENST00000237449
Q9H6A9_HUMAN	ENST00000309024
Q9H800_HUMAN	ENST00000357106
Q9H8D1_HUMAN	ENST00000360549
Q9HAC4_HUMAN	ENST00000206466
Q9P1M5_HUMAN	ENST00000303007
Q9ULE4_HUMAN	ENST00000265018
Q9Y6V0-3	ENST00000333891
QPCT	CCDS1790.1
QRICH2	NM_032134
QSCN6	CCDS1337.1
QSER1	NM_024774
QTRTD1	NM_024638
RAB36	CCDS13805.1
RAB3C	CCDS3976.1
RAB3GAP2	NM_012414
RAB3IL1	CCDS8014.1
RAC2	CCDS13945.1
RAD23A	CCDS12289.1
RAD51L3	CCDS11287.1
RAD52	CCDS8507.1
RAFTLIN	NM_015150
RAI1	CCDS11188.1
RALBP1	CCDS11845.1
RANBP17	NM_022897
RANP1	ENST00000333828
RAP140	CCDS2877.1
RAPGEF4	NM_007023
RAPGEF6	NM_016340
RAPGEFL1	CCDS11363.1
RAPH1	CCDS2359.1
RARSL	CCDS5011.1
RASGRF1	CCDS10309.1
RASGRF2	CCDS4052.1
RASL11B	CCDS3490.1
RAX	CCDS11972.1
RB1	NM_000321
RBM14	CCDS8147.1
RBM19	CCDS9172.1
RBM21	CCDS8021.1
RBM25	NM_021239
RBM27	ENST00000265271
RBM34	ENST00000362051
RBMS3	NM_001003792
RBP3	CCDS7218.1
RBPSUH	CCDS3436.1
RC74	NM_018250

Gene symbol	Accession ID
RCD-8	CCDS10849.1
RDHE2	CCDS6167.1
RDS	CCDS4871.1
REG1B	CCDS1963.1
REN	NM_000537
REPS2	CCDS14180.1
RET	CCDS7200.1
RFC2	CCDS5567.1
RFNG	NM_002917
RFX3	CCDS6450.1
RGS22	NM_015668
RGSL1	CCDS1346.1
RHOT1	NM_001033568
RICTOR	NM_152756
RIMBP2	NM_015347
RIMS2	NM_014677
RIMS4	CCDS13338.1
RIPK4	CCDS13675.1
RLBP1	NM_000326
RLTPR	NM_001013838
RNASEH2A	CCDS12282.1
RNF103	NM_005667
RNF127	CCDS14575.1
RNF128	CCDS14521.1
RNF19	CCDS6286.1
RNF25	CCDS2420.1
RNF40	CCDS10691.1
RNPC2	CCDS13265.1
ROBO3	NM_022370
ROCK1	CCDS11870.1
ROM1	CCDS8024.1
ROS1	CCDS5116.1
RoXaN	CCDS14013.1
RP1L1	NM_178857
RPL11	CCDS238.1
RPS14	CCDS4307.1
RPS6KA2	CCDS5294.1
RPS6KB2	NM_003952
RPUSD3	CCDS2586.1
RRAGD	CCDS5022.1
RSHL1	CCDS12675.1
RSU1	CCDS7112.1
RTN1	CCDS9740.1
RTTN	NM_173630
RUNX1	CCDS13639.1
RUNX1T1	CCDS6256.1
RWDD1	NM_001007464
RYR2	NM_001035
RYR3	NM_001036
SALL3	CCDS12013.1
SAMD11	ENST00000294573
SAMD9	NM_017654
SAPS2	NM_014678
SARG	CCDS1475.1
SARS	CCDS795.1
SASH1	CCDS5212.1
SCHIP1	CCDS3186.1
SCN1B	CCDS12441.1
SCN3A	NM_006922
SCN3B	CCDS8442.1
SCN5A	NM_000335
SCN9A	NM_002977
SCRIB	CCDS6411.1
SCUBE1	CCDS14048.1
SDC3	NM_014654
SDR-O	CCDS8926.1
SEC24C	CCDS7332.1
SELO	NM_031454
SEMA5A	CCDS3875.1
SEMA5B	CCDS3019.1
SEMA7A	CCDS10262.1
SEN2L	CCDS2611.1
SENP3	NM_015670
SEPT2	CCDS2548.1
SERPINA12	CCDS9926.1
SERPINA9	NM_175739
SERPINB3	CCDS11987.1
SERPINB7	CCDS11988.1
SERPINE2	CCDS2460.1
SERPING1	CCDS7962.1
SET7	CCDS3748.1
SETDB2	CCDS9417.1
SEZ6	NM_178860
SEZ6L	CCDS13833.1
SFI1	NM_001007467
SFMBT2	NM_001029880
SFRP2	NM_003013
SFTPB	CCDS1983.1
SG223_HUMAN	ENST00000330777
SGCZ	CCDS5992.1
SGK2	CCDS13320.1
SGPP1	CCDS9760.1
SGPP2	CCDS2453.1
SGSH	CCDS11770.1
SH3BP1	CCDS13952.1
SH3BP2	NM_003023
SH3GL3	CCDS10325.1
SHANK2	CCDS8198.1
SHANK3	ENST00000262795
SHB	NM_003028
SHE	NM_001010846
SHMT2	CCDS8934.1
SIGLEC11	CCDS12790.1
SIGLEC5	NM_003830
SIGLEC8	NM_014442
SIM2	CCDS13646.1
SIPA1L2	NM_020808
SIPA1L3	NM_015073
SKIV2L	CCDS4731.1
SKP2	CCDS3915.1
SLC10A4	CCDS3482.1
SLC11A1	CCDS2415.1
SLC12A1	CCDS10129.1
SLC12A5	CCDS13391.1
SLC14A1	CCDS11925.1
SLC14A2	CCDS11924.1
SLC16A5	CCDS11713.1
SLC1A2	NM_004171
SLC22A11	CCDS8074.1
SLC22A18	CCDS7740.1
SLC22A3	CCDS5277.1
SLC24A6	NM_024959
SLC25A13	CCDS5645.1
SLC26A4	CCDS5746.1
SLC2A1	CCDS477.1
SLC30A1	CCDS1499.1
SLC30A5	CCDS3996.1
SLC30A9	CCDS3465.1
SLC35B2	NM_178148
SLC35D3	NM_001008783
SLC35F2	NM_017515
SLC38A1	NM_030674
SLC38A4	CCDS8750.1
SLC38A6	CCDS9751.1
SLC39A2	CCDS9563.1
SLC43A3	CCDS7956.1
SLC4A1	CCDS11481.1
SLC4A5	CCDS1936.1
SLC4A7	NM_003615
SLC5A5	CCDS12368.1
SLC5A7	CCDS2074.1
SLC7A10	CCDS12431.1
SLC7A13	NM_138817
SLC7A14	NM_020949
SLC7A6	NM_003983
SLC8A1	CCDS1806.1
SLC9A1	CCDS295.1
SLC9A2	CCDS2062.1
SLC9A3R2	NM_004785
SLC9A4	NM_001011552
SLCO1B1	CCDS8685.1
SLCO2A1	CCDS3084.1
SLCO4C1	NM_180991
SLCO6A1	NM_173488
SLIT2	CCDS3426.1
SLITRK1	CCDS9464.1
SLITRK5	CCDS9465.1
SLITRK6	ENST00000313206
SMARCA2	NM_003070

Gene symbol	Accession ID
SMARCA4	CCDS12253.1
SMARCC2	CCDS8907.1
SMC5L1	CCDS6632.1
SMCR8	CCDS11195.1
SMF_HUMAN	ENST00000261804
SN	CCDS13060.1
SNED1	ENST00000310397
SNRPA	CCDS12565.1
SNX13	NM_015132
SNX27	CCDS1001.1
SNX4	CCDS3032.1
SOCS5	CCDS1830.1
SOHLH1	NM_001012415
SORCS2	NM_020777
SORCS3	CCDS7558.1
SORL1	CCDS8436.1
SOS1	CCDS1802.1
SOSTDC1	CCDS5360.1
SOX13	NM_005686
SOX30	CCDS4339.1
SOX8	CCDS10428.1
SP100	CCDS2477.1
SPACA4	CCDS12725.1
SPAG1	NM_003114
SPAG5	NM_006461
SPAG7	NM_004890
SPATA1	CCDS697.1
SPATA2	CCDS13422.1
SPATC1	CCDS6413.1
Spc25	CCDS2229.1
SPEG	ENST00000265327
SPEN	CCDS164.1
SPG3A	CCDS9700.1
SPI1	CCDS7933.1
SPIN3	NM_001010862
SPIRE2	NM_032451
SPN	CCDS10650.1
SPOCK3	NM_016950
SPON2	CCDS3347.1
SPRED2	NM_181784
SPTB	NM_001024858
SPTBN1	NM_178313
SPTBN2	CCDS8150.1
SPTBN4	CCDS12559.1
SPTBN5	NM_016642
SREBF2	CCDS14023.1
SRGAP1	CCDS8967.1
SRPK2	CCDS5735.1
SRRM2	NM_016333
SSFA2	CCDS2284.1
ST14	CCDS8487.1
ST8SIA4	CCDS4091.1
STAB1	NM_015136

Gene symbol	Accession ID
STAP2	CCDS12128.1
STIM2	CCDS3440.1
STK33	CCDS7789.1
STK39	NM_013233
STRA6	CCDS10261.1
STS	CCDS14127.1
STS-1	NM_032873
STX11	CCDS5205.1
STX12	CCDS310.1
STXBP2	CCDS12181.1
STXBP3	CCDS790.1
STYK1	CCDS8629.1
SUCLA2	CCDS9406.1
SUCLG2	NM_003848
SULT6B1	NM_001032377
SUNC1	NM_152782
SUSD5	ENST00000309558
SV2B	CCDS10370.1
SWAP70	NM_015055
SYDE2	ENST00000234668
SYN2	NM_133625
SYNE1	CCDS5236.1
SYNE1	CCDS5237.1
SYNE2	CCDS9761.1
SYT15	NM_181519
SYT16	NM_031914
SYT6	CCDS871.1
TAAR9	ENST00000340640
TACC2	CCDS7626.1
TACC3	CCDS3352.1
TAF1L	NM_153809
TAF4B	ENST00000269142
TAF6	CCDS5686.1
TANC1	NM_033394
TAOK1	NM_020791
TARBP2	CCDS8861.1
TAS1R2	CCDS187.1
TAS2R3	CCDS5867.1
TBC1D20	CCDS13002.1
TBC1D4	NM_014832
TBCD	NM_001033052
TBX20	CCDS5445.1
TBX22	CCDS14445.1
TCF7L1	CCDS1971.1
TCF8	CCDS7169.1
TCHH	ENST00000290632
TCN2	CCDS13881.1
TDRD5	CCDS1332.1
TDRD9	CCDS9987.1
TEAD2	CCDS12761.1
TEPP	CCDS10790.1
TERF2IP	NM_018975
TFE3	CCDS14315.1

Gene symbol	Accession ID
TGFBRAP1	CCDS2067.1
TGM1	CCDS9622.1
TGM5	NM_004245
THAP9	CCDS3598.1
THBS1	NM_003246
THEA	CCDS592.1
THOP1	CCDS12095.1
THRAP3	ENST00000354618
THSD7B	ENST00000272643
TIMP2	CCDS11758.1
TINAG	CCDS4955.1
TJP3	NM_014428
TLL1	CCDS3811.1
TLN1	NM_006289
TLX3	NM_021025
TM4SF14	CCDS7369.1
TM4SF3	CCDS8999.1
TM9SF4	CCDS13196.1
TMED1	CCDS12249.1
TMEM131	ENST00000186436
TMEM132C	ENST00000315208
TMEM16B	NM_020373
TMEM16C	NM_031418
TMEM16E	NM_213599
TMEM16G	NM_001001891
TMEM16J	NM_001012302
TMEM38A	CCDS12349.1
TMEM46	NM_001007538
TMEM63B	NM_018426
TMEM8	CCDS10407.1
TMPRSS2	NM_005656
TMPRSS4	NM_019894
TNC	CCDS6811.1
TNFAIP2	CCDS9979.1
TNFSF18	CCDS1305.1
TNFSF4	CCDS1306.1
TNFSF9	CCDS12169.1
TNIP1	NM_006058
TNIP2	CCDS3362.1
TNK1	NM_003985
TNMD	CCDS14469.1
TNN	NM_022093
TNPO1	CCDS4016.1
TNR	CCDS1318.1
TNRC15	NM_015575
TNRC4	CCDS1002.1
TNRC6C	NM_018996
TOE1	CCDS521.1
TOP2A	NM_001067
TOR1A	CCDS6930.1
TOSO	CCDS1473.1
TP53	CCDS11118.1
TPH2	NM_173353

Gene symbol	Accession ID
TPR	NM_003292
TPST2	CCDS13839.1
TRAM1L1	CCDS3707.1
TRAPPC3	CCDS404.1
TREML2	CCDS4853.1
TREML3	ENST00000332842
TRIM14	CCDS6734.1
TRIM42	CCDS3113.1
TRIM45	CCDS893.1
TRIM46	CCDS1097.1
TRIM55	CCDS6186.1
TRIM56	NM_030961
TRIM58	CCDS1636.1
TRIO	CCDS3883.1
TRIOBP	NM_007032
TRIP12	NM_004238
TRIP6	CCDS5708.1
TRMT5	NM_020810
TRPC4AP	CCDS13246.1
TRPC6	CCDS8311.1
TRPM2	CCDS13710.1
TRPM3	CCDS6634.1
TRPM4	NM_017636
TRPM5	NM_014555
TRPM6	CCDS6647.1
TRPM7	NM_017672
TRPV5	CCDS5875.1
TRRAP	CCDS5659.1
TSAP6	CCDS2125.1
TSC2	CCDS10458.1
TSCOT	CCDS6786.1
TSGA10	CCDS2037.1
TTC12	CCDS8360.1
TTC18	CCDS7324.1
TTC6	NM_001007795
TTLL2	CCDS5301.1
TTLL5	NM_015072
TTN	NM_133378
TTN	NM_133432
TUBGCP3	CCDS9525.1
TUBGCP6	CCDS14087.1
TULP1	CCDS4807.1
TXNDC3	CCDS5452.1
TYR	CCDS8284.1
UBAP2L	CCDS1063.1
UBE2G2	CCDS13714.1
UCHL1	CCDS3462.1
UGCGL2	CCDS9480.1
UGDH	CCDS3455.1
UGT1A6	CCDS2510.1
ULK1	CCDS9274.1
UNQ2446	CCDS10850.1
UNQ3030	CCDS3319.1
UNQ689	CCDS3542.1
UPK3B	CCDS5588.1
URB1	ENST00000270201
USH2A	CCDS1516.1
USP11	CCDS14277.1
USP26	CCDS14635.1
USP8	CCDS10137.1
VANGL1	CCDS883.1
VCAM1	CCDS773.1
VCIP135	CCDS6192.1
VCL	CCDS7340.1
VDP	NM_003715
VDR	CCDS8757.1
VGCNL1	CCDS9498.1
VGLL2	CCDS5115.1
VIPR2	CCDS5950.1
VMD2	NM_004183
VN2R1P	ENST00000312652
VPS11	NM_021729
VPS13A	CCDS6655.1
VPS24	NM_001005753
VPS41	CCDS5457.1
VPS45A	CCDS944.1
VSIG2	CCDS8452.1
VWF	CCDS8539.1
WBSCR17	CCDS5540.1
WBSCR27	CCDS5561.1
WDFY3	CCDS3609.1
WDR21	CCDS9809.1
WDR22	NM_003861
WDR24	CCDS10420.1
WDR27	NM_182552
WDR32	CCDS6613.1
WDR34	CCDS6906.1
WDR42B	ENST00000329763
WDR52	CCDS2972.1
WDR6	CCDS2782.1
WDR70	NM_018034
WDTC1	CCDS296.1
WEE1	CCDS7800.1
WFS1	CCDS3386.1
WNK1	CCDS8506.1
WNK2	CCDS6704.1
WNT9A	NM_003395
XAB2	NM_020196
XDH	CCDS1775.1
XPO1	NM_003400
XPO7	NM_015024
XR_016172.1	ENST00000355015
XR_017335.1	ENST00000314295
YN004_HUMAN	ENST00000281581
YTHDC2	CCDS4113.1
YWHAH	CCDS13901.1
ZAN	NM_173059
ZBTB16	CCDS8367.1
ZBTB24	NM_014797
ZBTB4	CCDS11107.1
ZBTB9	NM_006772
ZC3H6	NM_198581
ZFPM1	NM_153813
ZFYVE9	CCDS563.1
ZIC1	CCDS3136.1
ZIK1	NM_001010879
ZMAT4	NM_024645
ZNF10	CCDS9283.1
ZNF160	CCDS12859.1
ZNF17	NM_006959
ZNF18	NM_144680
ZNF183L1	CCDS9486.1
ZNF189	CCDS6754.1
ZNF25	CCDS7195.1
ZNF286	CCDS11172.1
ZNF294	NM_015565
ZNF295	CCDS13678.1
ZNF30	NM_194325
ZNF31	NM_145238
ZNF313	NM_018683
ZNF318	CCDS4895.1
ZNF333	CCDS12316.1
ZNF339	CCDS13132.1
ZNF343	CCDS13028.1
ZNF358	NM_018083
ZNF366	CCDS4015.1
ZNF406	NM_001029939
ZNF440L	NM_001012753
ZNF473	NM_015428
ZNF487	ENST00000315429
ZNF496	CCDS1631.1
ZNF497	CCDS12977.1
ZNF507	NM_014910
ZNF545	CCDS12493.1
ZNF547	NM_173631
ZNF558	CCDS12208.1
ZNF585A	CCDS12499.1
ZNF628	NM_033113
ZNF67	ENST00000323012
ZNF79	CCDS6871.1
ZP2	CCDS10596.1
ZSCAN2	CCDS10329.1
ZSWIM4	NM_023072
ZW10	CCDS8363.1

Note: Gene symbols are standard symbols assigned by Entrz Gene (http://www.ncbi.nlm.nih.gov/sites/entrez?db=gene). Accession IDs

"NM_XXXX" are uniquely assigned to each gene by National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/sites/entrez?db=nuccore). Accession IDs "CCDSXXXX" are uniquely assigned to individual genes by National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/CCDS/). Accession IDs "ENSTXXXXXXXXXX" are uniquely assigned to individual genes by *Ensembl* (http://www.ensembl.org/index.html).

Table 13. Genes containing somatic mutations in pancreatic cancer adapted from the paper by Jones *et. al.* (Jones et al., 2008).

Gene Symbol	Accession ID
7h3	CCDS12324.1
AARS	NM_001605
ABCA1	CCDS6762.1
ABCA12	NM_015657
ABCA7	CCDS12055.1
ABCB5	CCDS5371.1
ABCD2	CCDS8734.1
ABLIM2	NM_032432
ACACB	NM_001093
ACD	CCDS10842.1
ACE	CCDS11637.1
ACOT9	NM_001033583
ACTL7B	CCDS6771.1
ADA	CCDS13335.1
ADAM11	CCDS11486.1
ADAM12	CCDS7653.1
ADAM19	CCDS4338.1
ADAM21	CCDS9804.1
ADAMTS10	CCDS12206.1
ADAMTS15	CCDS8488.1
ADAMTS16	NM_139056
ADAMTS18	CCDS10926.1
ADAMTS2	CCDS4444.1
ADAMTS20	NM_175851
ADAMTS20	NM_025003
ADAMTS5	CCDS13579.1
ADAMTSL3	CCDS10326.1
ADCY2	CCDS3872.1
ADCY4	CCDS9627.1
ADD2	CCDS1906.1
ADPRHL2	CCDS402.1
AFF3	NM_001025108
AHNAK	NM_024060
AHNAK	NM_001620
AHR	CCDS5366.1
AICDA	NM_020661
AIM2	CCDS1181.1
AK3	CCDS629.1
AKAP12	CCDS5229.1
ALDH18A1	CCDS7443.1
ALDH1A3	CCDS10389.1
ALDH3A1	CCDS11212.1
ALDH3B1	NM_000694
ALDH8A1	CCDS5171.1
ALG8	CCDS8258.1
ALMS1	NM_015120
ALOX5	CCDS7212.1
AMIGO3	NM_198722
ANAPC4	CCDS3434.1
ANK3	CCDS7258.1
ANKAR	ENST00000313581
ANKRD27	NM_032139
ANKRD6	NM_014942
ANKRD9	CCDS9973.1
ANXA13	NM_001003954
AOX1	NM_001159
AP3B2	NM_004644
APC2	CCDS12068.1
APG4A	CCDS14538.1
APOB	CCDS1703.1
APRIN	NM_015032
APXL2	CCDS4161.1
AQP8	CCDS10626.1
ARFGAP1	CCDS13515.1
ARHGAP10	NM_024605
ARHGAP21	CCDS7144.1
ARHGAP28	NM_001010000
ARHGEF11	CCDS1162.1
ARHGEF7	CCDS9521.1
ARHGEF9	NM_015185
ARID1A	CCDS285.1
ARMC7	CCDS11714.1
ARMCX1	CCDS14487.1
ARNT2	NM_014862
ARRDC2	CCDS12370.1
ARSA	CCDS14100.1
ARSI	NM_001012301
ARTS-1	CCDS4085.1
ASB2	CCDS9915.1
ASXL2	NM_018263
ATF2	CCDS2262.1
ATN1	NM_001940
ATP10A	NM_024490
ATP10B	ENST00000327245
ATP10D	CCDS3476.1
ATP11B	NM_014616
ATP1A3	CCDS12594.1
ATP1B2	NM_001678
ATP2A1	CCDS10643.1
ATP2B3	CCDS14722.1
ATP6V0A4	CCDS5849.1
AZU1	CCDS12044.1
B3GALT1	CCDS2227.1
B3GNTL1	NM_001009905
B4GALT7	CCDS4429.1
BACH2	CCDS5026.1
BAI1	NM_001702
BAI3	CCDS4968.1
BAIAP2L2	NM_025045
BAIAP3	CCDS10434.1
BC037295_3	NM_001005850
BCAN	CCDS1149.1
BCHE	CCDS3198.1
BCL2A1	CCDS10312.1
Beta4GalNAc-T4	CCDS7694.1
BMPR2	NM_001204
BOC	CCDS2971.1
BPIL3	CCDS13211.1
BRCA2	CCDS9344.1
BSN	CCDS2800.1
BTBD7	NM_001002860
C10orf113	NM_001010896
C10orf31	NM_001012713
C10orf93	CCDS7672.1
C10orf99	CCDS7371.1
C11orf16	CCDS7794.1
C13orf22	CCDS9336.1
C13orf25	CCDS9467.1
C14orf121	NM_138360
C14orf124	NM_020195
C15orf16	CCDS10026.1
C15orf41	NM_032499
C17orf27	NM_020914
C17orf38	NM_001010855
C19orf20	NM_033513
C19orf22	CCDS12048.1
C19orf28	NM_174983
C19orf35	CCDS12087.1
C19orf6	CCDS12052.1
C1orf113	NM_024676
C1orf129	NM_025063
C1orf14	NM_030933
C1orf25	CCDS1366.1
C1orf45	NM_001025231
C1QL2	NM_182528
C1RL	CCDS8573.1
C20orf134	NM_001024675
C20orf161	CCDS13377.1
C20orf26	NM_015585
C20orf42	CCDS13098.1
C20orf77	CCDS13301.1
C21orf29	CCDS13712.1
C21orf63	CCDS13614.1
C2orf10	CCDS2291.1
C2orf29	CCDS2050.1
C3	NM_000064
C3orf15	CCDS2994.1
C3orf18	CCDS2829.1
C4orf9	NM_003703
C6orf103	ENST00000326916
C6orf213	NM_001010852
C6orf54	CCDS5304.1
C6orf60	NM_024581
C7orf27	CCDS5334.1

Gene Symbol	Accession ID
C9orf138	CCDS6487.1
C9orf39	NM_017738
C9orf45	CCDS6850.1
C9orf91	CCDS6808.1
C9orf98	CCDS6954.1
CABLES2	NM_031215
CACNA1A	NM_000068
CACNA1E	NM_000721
CACNA2D1	CCDS5598.1
CACNG5	CCDS11666.1
CAD	CCDS1742.1
CALB1	CCDS6251.1
CALCR	CCDS5631.1
CAMSAP1	NM_015447
CAMTA1	NM_015215
CAND2	ENST00000295989
CAPN12	CCDS12519.1
CARD9	CCDS6997.1
CASKIN2	CCDS11723.1
CASP10	CCDS2338.1
CAT	CCDS7891.1
CBFA2T2	CCDS13221.1
CBLN4	CCDS13448.1
CCDC11	CCDS11940.1
CCDC18	NM_206886
CCKAR	CCDS3438.1
CCL2	CCDS11277.1
CCNB3	CCDS14331.1
CCNYL3	ENST00000332505
CCR1	CCDS2737.1
CCT6A	CCDS5523.1
CCT6B	NM_006584
CD163	CCDS8578.1
CD1A	CCDS1174.1
CD200R1	CCDS2969.1
CD44	CCDS7897.1
CD6	CCDS7999.1
CD79A	CCDS12589.1
CD86	CCDS3009.1
CDC42BPA	CCDS1558.1
CDH1	CCDS10869.1
CDH10	CCDS3892.1
CDH20	CCDS11977.1
CDH7	CCDS11993.1
CDKN2A	CCDS6510.1
CDSN	NM_001264
CEBPZ	CCDS1787.1
CEECAM1	CCDS6901.1
CEL	NM_001807
CELSR1	CCDS14076.1
CENTD1	CCDS3441.1
Cep192	NM_032142
CEP290	NM_025114

Gene Symbol	Accession ID
CFHR4	NM_006684
CGI-09	CCDS13093.1
CGN	CCDS999.1
CHD1	NM_001270
CHD5	CCDS57.1
CHD7	NM_017780
CHD3L1	CCDS1435.1
CHMP1B	NM_020412
CHPPR	CCDS6182.1
CHST1	CCDS7913.1
CHURC1	NM_145165
CIAS1	CCDS1632.1
CILP	CCDS10203.1
CKLFSF4	CCDS10817.1
CLEC4M	CCDS12187.1
CLIPR-59	CCDS12486.1
CLK1	CCDS2331.1
CLSTN2	CCDS3112.1
CLUAP1	NM_015041
CMAS	CCDS8696.1
CMYA1	CCDS2683.1
CMYA3	NM_152381
CMYA5	NM_153610
CNGB1	NM_001297
CNGB3	CCDS6244.1
CNTN4	CCDS2558.1
CNTN5	NM_014361
CNTN6	CCDS2557.1
CNTNAP2	CCDS5889.1
CNTNAP4	CCDS10924.1
COBLL1	CCDS2223.1
COCH	CCDS9640.1
COH1	CCDS6280.1
COL11A1	CCDS778.1
COL14A1	NM_021110
COL17A1	CCDS7554.1
COL22A1	CCDS6376.1
COL4A1	CCDS9511.1
COL4A4	NM_000092
COL5A1	CCDS6982.1
COL6A3	NM_004369
COLEC12	NM_130386
CORO2A	CCDS6735.1
CPAMD8	NM_015692
CPLX2	ENST00000274615
CPN1	CCDS7486.1
CPT1C	CCDS12779.1
CPZ	CCDS3404.1
CREBBP	CCDS10509.1
CSF2RB	CCDS13936.1
CSMD1	NM_033225
CSMD2	CCDS380.1
CSS3	NM_175856

Gene Symbol	Accession ID
CTAG2	CCDS14759.1
CTNNA2	NM_004389
CTNNA3	CCDS7269.1
CTNND2	CCDS3881.1
CUBN	CCDS7113.1
CUL4B	NM_003588
CUTL1	CCDS5720.1
CX40.1	CCDS7191.1
CXorf9	CCDS14614.1
CYFIP1	CCDS10009.1
CYFIP2	NM_014376
CYP1A1	CCDS10268.1
DACH2	CCDS14455.1
DAXX	CCDS4776.1
DBT	CCDS767.1
DCC1	CCDS6330.1
DCHS1	CCDS7771.1
DCHS2	CCDS3785.1
DCT	CCDS9470.1
DDX51	NM_175066
DDX58	CCDS6526.1
DEPDC2	CCDS6201.1
DEPDC5	NM_014662
DET1	NM_017996
DFNB31	CCDS6806.1
DGKA	CCDS8896.1
DGKD	CCDS2504.1
DGKK	NM_001013742
DGKZ	CCDS7918.1
DHCR24	CCDS600.1
DHX33	CCDS11072.1
DHX8	CCDS11464.1
DICER1	CCDS9931.1
DIP2B	NM_173602
DKFZp313G1735	CCDS4073.1
DKFZP434B0335	NM_015395
DKFZP434G1415	CCDS8743.1
DKFZP434L1717	CCDS3805.1
DKFZp434O0527	CCDS2430.1
DKFZP564J0863	NM_015459
DKFZp566O084	CCDS11215.1
DKFZP586P0123	NM_015531
DKFZp761A052	CCDS14313.1
DLC1	CCDS5989.1
DLEC1	ENST00000337335
DLG2	NM_001364
DLG3	CCDS14403.1
DLGAP1	CCDS11836.1
DMD	CCDS14228.1
DMP1	CCDS3623.1
DNA2L	ENST00000358410
DNAH11	NM_003777
DNAH5	CCDS3882.1

Gene Symbol	Accession ID
DNAH8	CCDS4838.1
DNAH9	CCDS11160.1
DNAPTP6	NM_015535
DNHD2	NM_178504
DNM1L	CCDS8728.1
DOCK2	CCDS4371.1
DOT1L	NM_032482
DP58	NM_001004441
DPP6	NM_130797
DRD2	CCDS8361.1
DRD3	CCDS2978.1
DUOX2	CCDS10117.1
DUSP15	CCDS13193.1
DUSP19	CCDS2289.1
DYSF	CCDS1918.1
EBF	CCDS4343.1
EBF3	NM_001005463
EDG8	CCDS12240.1
EFEMP1	CCDS1857.1
EHMT1	CCDS7050.1
EIF2AK2	CCDS1786.1
EIF5	CCDS9980.1
EIF5B	NM_015904
ELA2	CCDS12045.1
ELAVL4	CCDS553.1
ELN	CCDS5562.1
EME2	NM_001010865
EMILIN1	CCDS1733.1
EML1	NM_004434
ENC1	CCDS4021.1
ENST00000294635	ENST00000294635
ENST00000298876	ENST00000298876
ENST00000309390	ENST00000309390
ENST00000322493	ENST00000322493
ENST00000324303	ENST00000324303
ENST00000326382	ENST00000326382
ENST00000326952	ENST00000326952
ENST00000332477	ENST00000332477
ENST00000333971	ENST00000333971
ENST00000334548	ENST00000334548
ENST00000336168	ENST00000336168
ENST00000340260	ENST00000340260
ENST00000356555	ENST00000356555
ENTH	NM_014666
EP300	CCDS14010.1
EPB41L1	CCDS13271.1
EPC2	NM_015630
EPHA3	CCDS2922.1
EPHA7	CCDS5031.1
EPHB1	NM_004441
EPHB2	CCDS229.1
EPHB6	CCDS5873.1
EPM2A	CCDS5206.1
EPPK1	NM_031308
EPS8L2	NM_022772
ERCC2	NM_000400
ERCC4	NM_005236
ERCC6	CCDS7230.1
EST1B	CCDS1137.1
ETS2	CCDS13659.1
ETV6	CCDS8643.1
EVI1	CCDS3205.1
EVPL	CCDS11737.1
EXOC2	NM_018303
EXOSC8	NM_181503
F10	CCDS9530.1
F13A1	CCDS4496.1
F8	NM_000132
FAD158	CCDS725.1
FADD	CCDS8196.1
FADS1	CCDS8013.1
FADS2	CCDS8012.1
FAM132B	ENST00000344233
FAM47B	ENST00000329357
FAM50B	CCDS4487.1
FAM53B	CCDS7641.1
FAM54B	NM_019557
FAM55C	CCDS2945.1
FAT	NM_005245
FAT3	ENST00000298047
FAT4	CCDS3732.1
FBN2	NM_001999
FBN3	CCDS12196.1
FBXO15	CCDS12002.1
FBXO3	CCDS7887.1
FBXO41	ENST00000295133
FBXO9	NM_033481
FBXW7	CCDS3777.1
FBXW8	CCDS9182.1
FGD2	CCDS4829.1
FGD5	NM_152536
FKRP	CCDS12691.1
FKSG44	CCDS8102.1
FLJ10324	NM_018059
FLJ10407	CCDS583.1
FLJ10521	CCDS182.1
FLJ10647	CCDS406.1
FLJ12886	NM_019108
FLJ14011	CCDS12944.1
FLJ14299	CCDS6094.1
FLJ14490	CCDS446.1
FLJ14640	NM_032816
FLJ20032	CCDS3666.1
FLJ20035	NM_017631
FLJ20244	CCDS12293.1
FLJ20245	CCDS7041.1
FLJ20457	CCDS6774.1
FLJ20580	CCDS576.1
FLJ21628	CCDS4440.1
FLJ21816	NM_024675
FLJ21986	NM_024913
FLJ23420	CCDS12189.1
FLJ23577	ENST00000303168
FLJ23588	CCDS14049.1
FLJ25006	CCDS11237.1
FLJ25530	CCDS8456.1
FLJ26175	NM_001001668
FLJ31295	CCDS8763.1
FLJ32110	CCDS5613.1
FLJ32112	CCDS587.1
FLJ32416	CCDS12086.1
FLJ32685	CCDS2645.1
FLJ34969	NM_152678
FLJ35220	NM_173627
FLJ35843	CCDS9151.1
FLJ36180	CCDS3851.1
FLJ36748	NM_152406
FLJ37396	CCDS5072.1
FLJ38020	NM_001039775
FLJ38377	CCDS2164.1
FLJ39155	CCDS3924.1
FLJ39501	CCDS12331.1
FLJ39502	CCDS2281.1
FLJ40235	CCDS12827.1
FLJ41046	NM_207479
FLJ41993	NM_001001694
FLJ45231	NM_001039778
FLJ45909	CCDS12522.1
FLJ46072	CCDS6410.1
FLJ46365	CCDS6144.1
FLJ46481	CCDS3384.1
FLJ46536	NM_198483
FLJ90805	CCDS12603.1
FMN2	NM_020066
FMNL1	CCDS11497.1
FMNL3	NM_175736
FMR1	CCDS14682.1
FMR2	CCDS14684.1
FN1	CCDS2399.1
FOXJ1	NM_001454
FOXP2	CCDS5760.1
FREM1	NM_144966
FREM2	NM_207361
FRMPD4	NM_014728
FSTL5	CCDS3802.1
FTCD	CCDS13731.1
FTHL17	CCDS14227.1
GABRA1	CCDS4357.1
GABRR1	CCDS5019.1

Gene Symbol	Accession ID	Gene Symbol	Accession ID	Gene Symbol	Accession ID
GALNT13	CCDS2199.1	HELB	CCDS8976.1	KBTBD11	NM_014867
GALNT4	NM_003774	HELZ	NM_014877	KCNA3	CCDS828.1
GALNT8	CCDS8533.1	HIP1	NM_005338	KCNA4	NM_002233
GAS7	CCDS11152.1	HIST1H3A	CCDS4570.1	KCNB1	CCDS13418.1
GBP3	CCDS717.1	HIST1H4I	CCDS4620.1	KCNB2	CCDS6209.1
GDF6	NM_001001557	HKR2	CCDS12975.1	KCNC2	CCDS9005.1
GFAP	CCDS11491.1	HMGCLL1	NM_019036	KCNC3	CCDS12793.1
GFRA1	CCDS7593.1	HOXC10	CCDS8868.1	KCNJ3	CCDS2200.1
GH2	CCDS11648.1	HOXC9	CCDS8869.1	KCNK10	CCDS9880.1
GIMAP7	CCDS5903.1	HOXD4	CCDS2269.1	KCNMA1	CCDS7352.1
GJA3	CCDS9289.1	HPCAL1	CCDS1671.1	KCNT1	NM_020822
GLB1L3	ENST00000299136	HPS5	CCDS7836.1	KCTD15	CCDS12434.1
GLI1	CCDS8940.1	HRB2	CCDS9012.1	KEAP1	CCDS12239.1
GLI3	CCDS5465.1	HRPT2	CCDS1382.1	KIAA0082	CCDS4835.1
GLP1R	CCDS4839.1	HS3ST2	CCDS10606.1	KIAA0317	ENST00000338772
GLTSCR1	NM_015711	HS3ST5	NM_153612	KIAA0367	NM_015225
GNAT1	CCDS2812.1	HSGT1	CCDS7321.1	KIAA0372	CCDS4072.1
GOLGA3	CCDS9281.1	HTR1A	NM_000524	KIAA0590	CCDS10439.1
GPC2	CCDS5689.1	HYPC	CCDS8789.1	KIAA0774	NM_001033602
GPR	CCDS10051.1	IER5	CCDS1343.1	KIAA1024	NM_015206
GPR110	ENST00000326374	IL12RB1	NM_153701	KIAA1086	ENST00000262961
GPR133	CCDS9272.1	IL17RB	CCDS2874.1	KIAA1102	NM_014988
GPR151	NM_194251	IL17RC	CCDS2590.1	KIAA1109	ENST00000264501
GPR154	CCDS5443.1	IL18R1	CCDS2060.1	KIAA1219	CCDS13305.1
GPR158	NM_020752	IL2RG	CCDS14406.1	KIAA1543	ENST00000160298
GPR35	CCDS2541.1	ILK	CCDS7768.1	KIAA1704	CCDS9394.1
GPR54	CCDS12049.1	IMP5	NM_175882	KIAA1751	ENST00000270720
GPR73L1	CCDS13089.1	INHBB	CCDS2132.1	KIAA1755	NM_001029864
GPR82	CCDS14259.1	INO80	CCDS10071.1	KIAA1944	CCDS9266.1
GPRC5C	CCDS11699.1	INPP5D	NM_001017915	KIAA1957	ENST00000332235
GPS2	CCDS11100.1	INTS2	NM_020748	KIAA1961	NM_133372
GPX6	NM_182701	IQGAP1	CCDS10362.1	KIAA2013	ENST00000329923
GRCA	CCDS8563.1	IRGQ	NM_001007561	KIF21A	NM_017641
GRHL1	NM_198182	IRS4	CCDS14544.1	KIF25	CCDS5305.1
GRIA3	CCDS14604.1	IRX1	NM_024337	KIF3A	NM_007054
GRIK2	CCDS5048.1	ISYNA1	CCDS12379.1	KIN	CCDS7080.1
GRIN3A	CCDS6758.1	ITGA11	NM_001004439	KIRREL	CCDS1172.1
GRIP2	ENST00000273083	ITGA3	CCDS11557.1	KIT	CCDS3496.1
GRM6	CCDS4442.1	ITGA4	NM_000885	KLF5	CCDS9448.1
GRM8	CCDS5794.1	ITGA9	CCDS2669.1	KLHDC1	CCDS9692.1
GSDML	CCDS11354.1	ITGAE	NM_002208	KLHDC4	CCDS10963.1
GSR	NM_000637	ITGB4BP	CCDS13249.1	KLP1	CCDS12926.1
GTF3C1	NM_001520	ITIH2	NM_002216	KPNB1	CCDS11513.1
GTF3C3	CCDS2316.1	ITLN1	CCDS1211.1	KRAS	CCDS8702.1
GUCA2A	CCDS465.1	ITPR1	NM_002222	KRT13	CCDS11396.1
GUCY1A2	CCDS8335.1	IXL	NM_017592	KRT9	NM_000226
H1T2	CCDS8762.1	JAG1	CCDS13112.1	KRTAP11-1	CCDS13608.1
HAPLN4	CCDS12398.1	JM11	CCDS14316.1	L3MBTL4	CCDS11839.1
HAS1	CCDS12838.1	JMJD3	ENST00000254846	LAMA1	NM_005559
HBXIP	CCDS824.1	JPH3	CCDS10962.1	LAMA4	NM_002290
HCK	NM_002110	JPH4	CCDS9603.1	LAMA5	NM_005560
HECW1	CCDS5469.1	K6IRS2	CCDS8833.1	LAMC3	CCDS6938.1
HECW2	NM_020760	KAL1	CCDS14130.1	LARP	CCDS4328.1

Gene Symbol	Accession ID
LASS3	CCDS10384.1
LCT	CCDS2178.1
LENG8	CCDS12894.1
LGI4	CCDS12444.1
LGR6	CCDS1424.1
LIG3	CCDS11284.1
LIMR	CCDS8780.1
LIPH	CCDS3272.1
LMOD1	NM_012134
LMTK2	CCDS5654.1
LMX1A	CCDS1247.1
LOC113179	CCDS12076.1
LOC113386	NM_138781
LOC123872	CCDS10943.1
LOC126147	NM_145807
LOC128153	CCDS1519.1
LOC130951	NM_138804
LOC131873	ENST00000358511
LOC163131	NM_001005851
LOC167127	CCDS3914.1
LOC222967	ENST00000297186
LOC283219	NM_001029859
LOC283398	ENST00000342823
LOC284434	NM_001007525
LOC339768	CCDS2525.1
LOC340578	NM_001013628
LOC342979	ENST00000340790
LOC343521	NM_001013632
LOC387720	NM_001013633
LOC388135	NM_001039614
LOC392617	ENST00000333066
LOC399706	NM_001010910
LOC441136	NM_001013719
LOC441476	NM_001004353
LOC441722	ENST00000311061
LOC51334	CCDS4127.1
LOC63920	NM_022090
LOC89944	NM_138342
LPAL2	ENST00000342479
LPHN3	NM_015236
LPL	CCDS6012.1
LRFN5	CCDS9678.1
LRP1	CCDS8932.1
LRP1B	CCDS2182.1
LRP2	CCDS2232.1
LRP3	CCDS12430.1
LRP5	CCDS8181.1
LRRC16	NM_017640
LRRC18	NM_001006939
LRRC3B	CCDS2644.1
LRRC4	CCDS5799.1
LRRC48	NM_031294
LRRK2	NM_198578
LRRN3	CCDS5754.1
LRRTM4	NM_024993
MAGEE1	CCDS14433.1
MAMDC1	NM_182830
MAN2A1	NM_002372
MAP1A	NM_002373
MAP1B	CCDS4012.1
MAP2	CCDS2384.1
MAP2K6	CCDS11686.1
MAP4K2	CCDS8082.1
MAP4K3	CCDS1803.1
MAP4K4	ENST00000302217
MAPKBP1	NM_014994
MAPT	CCDS11499.1
MARLIN1	CCDS3385.1
MARS	CCDS8942.1
MASP2	CCDS123.1
MASS1	NM_032119
MAST2	NM_015112
MAT2B	CCDS4365.1
MBD3	CCDS12072.1
MCM7	CCDS5683.1
MCTP2	NM_018349
MEGF11	CCDS10213.1
MEP1A	CCDS4918.1
METTL3	NM_019852
MGC10731	CCDS171.1
MGC13125	CCDS8374.1
MGC15523	CCDS11780.1
MGC15875	CCDS4434.1
MGC20806	CCDS11797.1
MGC2494	CCDS10423.1
MGC26598	CCDS9036.1
MGC26988	CCDS4335.1
MGC29649	CCDS8033.1
MGC33407	CCDS12207.1
MGC34713	CCDS4070.1
MGC35138	CCDS7701.1
MGC35555	CCDS6307.1
MGC39581	CCDS12149.1
MGC4266	CCDS8522.1
MGC50721	CCDS10602.1
MGC5297	CCDS3873.1
MID1	CCDS14138.1
MIZF	CCDS8414.1
MKL2	NM_014048
MLC1	CCDS14083.1
MLL	NM_005933
MLL2	NM_003482
MLL3	CCDS5931.1
MLL5	NM_182931
MMP9	CCDS13390.1
MOBKL2C	CCDS539.1
MORC	CCDS2955.1
MORC2	NM_014941
MOXD1	CCDS5152.1
MPHOSPH1	CCDS7407.1
MPL	CCDS483.1
MPN2	CCDS1563.1
MPO	CCDS11604.1
MPZ	CCDS1229.1
MRGPRD	ENST00000309106
MRGX1	CCDS7846.1
MRPL38	CCDS11733.1
MRPS7	CCDS11718.1
MSLN	NM_013404
MTF1	NM_005955
MTMR12	NM_019061
MTMR2	CCDS8305.1
MTO1	CCDS4979.1
MTR	CCDS1614.1
MUC1	CCDS1098.1
MUC15	CCDS7859.1
MUC16	NM_024690
MUC2	NM_002457
MUF1	CCDS533.1
MUM1L1	NM_152423
MYBL1	ENST00000331406
MYBPHL	NM_001010985
MYCBPAP	NM_032133
MYH2	CCDS11156.1
MYH3	CCDS11157.1
MYH6	CCDS9600.1
MYH9	CCDS13927.1
MYLIP	CCDS4536.1
MYO10	NM_012334
MYO15A	NM_016239
MYO1G	NM_033054
MYO3A	CCDS7148.1
MYO6	NM_004999
MYO7B	ENST00000272666
MYO9A	CCDS10239.1
MYOM1	NM_003803
MYST3	CCDS6124.1
NAALAD2	CCDS8288.1
NAALADL2	NM_207015
NALP10	CCDS7784.1
NALP13	NM_176810
NALP14	CCDS7776.1
NALP4	CCDS12936.1
NAV2	CCDS7850.1
NAV3	NM_014903
NCDN	CCDS392.1
NCK1	CCDS3092.1
NCL	NM_005381
NCOA2	NM_006540

Gene Symbol	Accession ID
NEB	NM_004543
NEK8	NM_178170
NEO1	CCDS10247.1
NFATC3	CCDS10860.1
NFIA	CCDS615.1
NID	CCDS1608.1
NID2	CCDS9706.1
NIF3L1BP1	CCDS2900.1
NIPSNAP3B	CCDS6761.1
NKX2-2	CCDS13145.1
NLGN1	CCDS3222.1
NMUR1	CCDS2486.1
NOD3	NM_178844
NOL5A	CCDS13030.1
NOPE	CCDS10206.1
NOR1	CCDS409.1
NOS1	NM_000620
NOX5	NM_024505
NP_001035826.1	ENST00000331090
NP_001074311.1	ENST00000326096
NPD014	CCDS260.1
NPHP4	NM_015102
NPY1R	NM_000909
NRG2	CCDS4217.1
NRXN2	CCDS8077.1
NRXN3	CCDS9870.1
NSE1	CCDS1684.1
NTF3	CCDS8538.1
NTRK3	CCDS10340.1
NUDT5	CCDS7089.1
ENST00000318605	ENST00000318605
NUP210	NM_024923
NURIT	CCDS9399.1
NXN	CCDS10998.1
NXPH3	CCDS11550.1
OBSCN	CCDS1570.1
OBSL1	ENST00000265318
OCA2	CCDS10020.1
ODZ4	ENST00000278550
OGDHL	CCDS7234.1
OGFOD2	NM_024623
OGT	CCDS14414.1
OR10A3	ENST00000360759
OR10K2	NM_001004476
OR10P1	NM_206899
OR10R2	NM_001004472
OR10Z1	NM_001004478
OR11L1	NM_001001959
OR13C3	NM_001001961
OR13C5	NM_001004482
OR1J2	NM_054107
OR2AJ1	ENST00000318244
OR2T1	NM_030904
OR2W3	NM_001001957
OR4A16	NM_001005274
OR4B1	NM_001005470
OR4E2	NM_001001912
OR4L1	NM_001004717
OR4X1	NM_001004726
OR51B4	CCDS7757.1
OR51E1	NM_152430
OR51F2	NM_001004753
OR52I2	NM_001005170
OR52L1	ENST00000332249
OR5C1	NM_001001923
OR5D13	NM_001001967
OR5D3P	ENST00000333984
OR5F1	NM_003697
OR5J2	NM_001005492
OR5T1	NM_001004745
OR6A2	CCDS7772.1
OR6K2	NM_001005279
OR8D2	NM_001002918
OR8H1	NM_001005199
OR8K1	NM_001002907
OR8K5	NM_001004058
OR9I1	NM_001005211
OR9K2	NM_001005243
ORC5L	CCDS5734.1
OSBPL6	CCDS2277.1
OSCAR	CCDS12873.1
OSMR	CCDS3928.1
OSTN	CCDS3299.1
OTOF	CCDS1724.1
OTP	CCDS4039.1
OTX1	CCDS1873.1
OVCA2	NM_001383
OVCH1	NM_183378
P11	CCDS8754.1
PABPC5	CCDS14460.1
PACS2	NM_015197
PADI2	CCDS177.1
PALMD	CCDS758.1
PAPPA	CCDS6813.1
PARP10	NM_032789
PARP14	NM_017554
PARP2	NM_005484
PARP9	CCDS3014.1
PAX6	NM_000280
PB1	CCDS2859.1
PCDH15	CCDS7248.1
PCDH17	NM_014459
PCDH18	NM_019035
PCDH9	CCDS9443.1
PCDHA13	NM_031864
PCDHB16	CCDS4251.1
PCDHB2	CCDS4244.1
PCDHB3	CCDS4245.1
PCDHGA1	NM_031993
PCDHGA11	NM_032091
PCDHGA8	NM_014004
PCDHGC4	CCDS4260.1
PCNT	NM_006031
PCNXL2	ENST00000344698
PCSK2	CCDS13125.1
PCSK6	NM_138321
PDE6A	CCDS4299.1
PDZRN3	NM_015009
PDZRN4	CCDS8739.1
PEG3	CCDS12948.1
PER3	CCDS89.1
PFAS	CCDS11136.1
PGM5	CCDS6622.1
PGR	CCDS8310.1
PHACTR3	CCDS13480.1
PHB2	NM_007273
PIAS4	CCDS12118.1
PIGK	CCDS674.1
PIGT	CCDS13353.1
PIK3CG	CCDS5739.1
PIK3R2	CCDS12371.1
PIP5K3	CCDS2382.1
PITRM1	NM_014889
PKD1L2	NM_182740
PKHD1L1	NM_177531
PKIA	CCDS6222.1
PKP2	CCDS8731.1
PLCB2	NM_004573
PLCB3	CCDS8064.1
PLCB4	CCDS13104.1
PLEC1	NM_201380
PLEC1	NM_201378
PLEK2	CCDS9782.1
PLEKHA6	CCDS1444.1
PLEKHG2	NM_022835
PLK5_HUMAN	ENST00000334770
PLXNA1	NM_032242
PLXNB1	CCDS2765.1
PMP22CD	NM_001013743
PNPLA1	NM_001039725
PODN	CCDS573.1
PODXL	NM_001018111
POLR2A	NM_000937
POLRMT	CCDS12036.1
PON1	CCDS5638.1
PPA2	CCDS3667.1
PPFIA2	NM_003625
PPP1CA	CCDS8160.1
PPP1R15B	CCDS1445.1

Gene Symbol	Accession ID	Gene Symbol	Accession ID	Gene Symbol	Accession ID
PPP1R3A	CCDS5759.1	Q96M18_HUMAN	ENST00000335239	RODH	CCDS8925.1
PPP2R1A	CCDS12849.1	Q96MJ2_HUMAN	ENST00000327832	RP1	CCDS6160.1
PPP2R3A	CCDS3087.1	Q96QE0_HUMAN	ENST00000301647	RPGRIP1	NM_020366
PPP2R4	CCDS6920.1	Q96RX8_HUMAN	ENST00000301719	RREB1	NM_001003699
PPP5C	CCDS12684.1	Q96S27_HUMAN	ENST00000301682	RTL1	ENST00000331067
PRDM10	CCDS8484.1	Q9H557_HUMAN	ENST00000237253	RTTN	NM_173630
PRDM5	CCDS3716.1	Q9H5F0_HUMAN	ENST00000360484	RUNX1T1	CCDS6256.1
PRDM9	NM_020227	Q9H8A7_HUMAN	ENST00000053084	RYR1	NM_000540
PRELP	CCDS1438.1	Q9HA39_HUMAN	ENST00000329980	RYR2	NM_001035
PREX1	CCDS13410.1	Q9HCM3_HUMAN	ENST00000242365	SACS	CCDS9300.1
PRG-3	CCDS6751.1	Q9NSI0_HUMAN	ENST00000328881	SARS2	NM_017827
PRKACG	CCDS6625.1	Q9NT86_HUMAN	ENST00000314272	SART3	CCDS9117.1
PRKCG	CCDS12867.1	Q9P169_HUMAN	ENST00000342338	SBLF	CCDS1840.1
PRKD1	CCDS9637.1	Q9P193_HUMAN	ENST00000359406	SCAP2	CCDS5400.1
ProSAPiP1	CCDS13049.1	Q9P1M5_HUMAN	ENST00000303007	SCFD2	NM_152540
PRR12	ENST00000246798	Q9Y6V0-3	ENST00000333891	SCGN	CCDS4561.1
PRSS23	CCDS8278.1	QRICH2	NM_032134	SCN11A	NM_014139
PSMD3	CCDS11356.1	RAB6B	CCDS3082.1	SCN2A2	NM_021007
PSME4	NM_014614	RAD9B	CCDS9148.1	SCN4A	NM_000334
PTCHD2	ENST00000294484	RAG1	CCDS7902.1	SCN5A	NM_000335
PTCHD3	NM_001034842	RAG2	CCDS7903.1	SCN5A	NM_198056
PTF1A	CCDS7143.1	RaLP	CCDS10130.1	SCN7A	NM_002976
PTGER3	CCDS652.1	RANBP2	CCDS2079.1	SCNM1	CCDS987.1
PTN	CCDS5844.1	RARB	CCDS2642.1	SCNN1B	CCDS10609.1
PTPN12	CCDS5592.1	RARRES2	CCDS5902.1	SCNN1G	CCDS10608.1
PTPRK	CCDS5137.1	RASEF	ENST00000330861	SCRIB	CCDS6411.1
PTPRZ1	NM_002851	RASGRP3	NM_170672	SDPR	CCDS2313.1
PUM1	CCDS338.1	RASGRP4	NM_170603	SDS	CCDS9169.1
PWP2H	NM_005049	RASIP1	CCDS12731.1	SEC14L3	CCDS13877.1
PXDN	ENST00000252804	RASSF6	CCDS3558.1	SEMA4D	CCDS6685.1
PXDNL	NM_144651	RBAF600	CCDS189.1	SEMA5B	CCDS3019.1
PYHIN1	CCDS1178.1	RBBP6	CCDS10621.1	SENP1	NM_014554
Q08AG5_HUMAN	ENST00000334213	RBM27	ENST00000265271	SESN2	CCDS321.1
Q5JX50_HUMAN	ENST00000325076	RC74	NM_018250	SEZ6L	CCDS13833.1
Q5SYT8_HUMAN	ENST00000279434	RCHY1	CCDS3567.1	SF3A1	CCDS13875.1
Q6ZMX6_HUMAN	ENST00000269197	RDH8	CCDS12223.1	SF3B1	NM_012433
Q6ZT40_HUMAN	ENST00000296564	RELN	NM_005045	SFRS12	CCDS3991.1
Q7Z2Q7_HUMAN	ENST00000334994	RENBP	CCDS14738.1	SFRS16	CCDS12652.1
Q7Z7L8_HUMAN	ENST00000339446	REPIN1	NM_013400	SGEF	NM_015595
Q8N2V9_HUMAN	ENST00000324414	RFX1	CCDS12301.1	SH2D1B	NM_053282
Q8N5S4_HUMAN	ENST00000326474	RFX3	CCDS6449.1	SH3GL3	CCDS10325.1
Q8N6V7_HUMAN	ENST00000324928	RFXDC1	CCDS5113.1	SH3TC1	CCDS3399.1
Q8N800_HUMAN	ENST00000322516	RGS11	CCDS10403.1	SHANK2	CCDS8198.1
Q8N9F6_HUMAN	ENST00000317122	RGS17	CCDS5244.1	SHKBP1	CCDS12560.1
Q8N9G5_HUMAN	ENST00000313957	RHBDF1	NM_022450	SI	CCDS3196.1
Q8N9S5_HUMAN	ENST00000329388	RHOT2	CCDS10417.1	SIDT1	CCDS2974.1
Q8N9V7_HUMAN	ENST00000309765	RIC3	CCDS7788.1	SIGLEC11	CCDS12790.1
Q8N9Z1_HUMAN	ENST00000326413	RIMBP2	NM_015347	SIPA1L2	NM_020808
Q8NCK2_HUMAN	ENST00000325720	RIMS1	NM_014989	SIX2	CCDS1822.1
Q8NGP7_HUMAN	ENST00000341231	RIMS2	NM_014677	SKD3	CCDS8215.1
Q8NH06_HUMAN	ENST00000324144	RLF	CCDS448.1	SLC14A1	CCDS11925.1
Q8NH08_HUMAN	ENST00000327198	RNF175	NM_173662	SLC17A1	CCDS4565.1
Q96GK3_HUMAN	ENST00000315264	RNUT1	CCDS10281.1	SLC17A7	CCDS12764.1

Gene Symbol	Accession ID	Gene Symbol	Accession ID	Gene Symbol	Accession ID
SLC1A6	CCDS12321.1	STAC	CCDS2662.1	TM9SF4	CCDS13196.1
SLC22A15	NM_018420	STAC2	CCDS11335.1	TMCC2	NM_014858
SLC22A7	CCDS4893.1	STAMBP	CCDS1929.1	TMEFF2	CCDS2314.1
SLC25A26	CCDS2905.1	STARD13	CCDS9348.1	TMEM132B	NM_052907
SLC28A3	CCDS6670.1	STARD8	CCDS14390.1	TMEM16A	NM_018043
SLC2A1	CCDS477.1	STAT4	CCDS2310.1	TMEM16C	NM_031418
SLC2A3	CCDS8586.1	STIM1	CCDS7749.1	TMEM16G	NM_001001891
SLC2A5	CCDS99.1	STK10	NM_005990	TMEM63B	NM_018426
SLC33A1	CCDS3173.1	STK23	NM_014370	TMEM8	CCDS10407.1
SLC39A10	NM_020342	STK33	CCDS7789.1	TMEPAI	CCDS13462.1
SLC39A6	NM_012319	STMN4	CCDS6055.1	TMPO	CCDS9064.1
SLC45A1	ENST00000289877	STN2	CCDS9875.1	TMPRSS13	NM_032046
SLC4A10	NM_022058	SULF1	CCDS6204.1	TNF	CCDS4702.1
SLC4A8	CCDS8814.1	SULF2	CCDS13408.1	TNFRSF8	CCDS144.1
SLC4A9	NM_031467	SV2A	CCDS940.1	TNK1	NM_003985
SLC6A15	CCDS9026.1	SYNE1	CCDS5236.1	TNNI3	NM_000363
SLC6A17	NM_001010898	SYNE1	CCDS5237.1	TNR	CCDS1318.1
SLC6A2	CCDS10754.1	SYNE2	CCDS9761.1	TOR3A	CCDS1329.1
SLC6A3	CCDS3863.1	SYP	CCDS14321.1	TP53	CCDS11118.1
SLC9A5	NM_004594	SYT1	CCDS9017.1	TP53BP1	CCDS10096.1
SLCO1A2	CCDS8686.1	SYT6	CCDS871.1	TPO	CCDS1642.1
SLCO1B1	CCDS8685.1	SYT7	NM_004200	TREH	NM_007180
SLCO1C1	CCDS8683.1	T	CCDS5290.1	TRERF1	CCDS4867.1
SLCO4C1	NM_180991	TAF1B	NM_005680	TRIM37	NM_001005207
SLITRK2	CCDS14680.1	TAF1L	NM_153809	TRIM58	CCDS1636.1
SLITRK3	CCDS3197.1	TAF4	NM_003185	TRPM1	CCDS10024.1
SLITRK5	CCDS9465.1	TAS2R41	NM_176883	TRPM2	CCDS13710.1
SMAD3	CCDS10222.1	TATDN2	NM_014760	TRPM3	CCDS6634.1
SMAD4	CCDS11950.1	TBC1D14	CCDS3394.1	TSC2	CCDS10458.1
SMARCA4	CCDS12253.1	TBX15	NM_152380	TSP-NY	CCDS9237.1
SMOC1	CCDS9798.1	TBX18	ENST00000330469	TSTA3	CCDS6408.1
SMTN	CCDS13886.1	TBX5	CCDS9173.1	TTBK2	NM_173500
SN	CCDS13060.1	TBX6	CCDS10670.1	TTC12	CCDS8360.1
SNCAIP	CCDS4131.1	TCEB3B	CCDS11932.1	TTC21B	NM_024753
SNRPC	NM_003093	TCFL1	CCDS989.1	TTC24	ENST00000340086
SNX16	CCDS6234.1	TDRD7	CCDS6725.1	TTF1	CCDS6948.1
SNX26	CCDS12477.1	TENC1	CCDS8842.1	TTK	CCDS4993.1
SORL1	CCDS8436.1	TESSP2	NM_182702	TTN	NM_133378
SOX3	CCDS14669.1	TEX14	NM_198393	TTN	NM_133437
SP8	CCDS5372.1	TFCP2L1	CCDS2134.1	TUBB3	CCDS10988.1
SPAP1	CCDS1168.1	TFF2	CCDS13684.1	TXNDC6	CCDS3099.1
SPATA13	ENST00000360220	TFPI2	CCDS5632.1	UBE1L	CCDS2805.1
SPINLW1	CCDS13359.1	TFR2	NM_003227	UBE2M	CCDS12987.1
SPTAN1	CCDS6905.1	TFSM1_HUMAN	ENST00000314720	UBQLN4	CCDS1127.1
SPTBN2	CCDS8150.1	TG	NM_003235	UBR2	CCDS4870.1
SR140_HUMAN	ENST00000319822	TGFBR2	CCDS2648.1	UBXD7	ENST00000296328
SRCRB4D	CCDS5585.1	TGIF2	CCDS13278.1	UCP3	CCDS8229.1
SRRM2	NM_016333	THNSL1	CCDS7147.1	ULBP1	CCDS5223.1
SST	CCDS3288.1	THSD7B	ENST00000272643	UNC13C	ENST00000260323
ST6GAL2	CCDS2073.1	TIMELESS	CCDS8918.1	USP20	NM_001008563
ST6GALNAC5	CCDS673.1	TJP1	NM_175610	USP31	CCDS10607.1
ST8SIA5	CCDS11930.1	TLL2	CCDS7449.1	USP38	CCDS3758.1
STAB1	NM_015136	TM7SF4	CCDS6301.1	USP42	NM_032172

Gene Symbol	Accession ID
UTRN	NM_007124
VDAC2	CCDS7348.1
VGCNL1	CCDS9498.1
VIM	CCDS7120.1
VIT	NM_053276
VLDLR	CCDS6446.1
VMD2L1	NM_017682
VPS13A	CCDS6655.1
VPS13D	NM_018156
VPS16	CCDS13036.1
VPS39	CCDS10083.1
VSIG1	CCDS14535.1
VWF	CCDS8539.1
WASF3	CCDS9318.1
WBSCR14	CCDS5553.1
WBSCR17	CCDS5540.1
WDR1	NM_005112
WDR17	CCDS3825.1
WDR27	NM_182552
WDR42B	ENST00000329763
WDR44	CCDS14572.1
WHSC1	CCDS3357.1
WIRE	CCDS11364.1
WNT9A	NM_003395
WRNIP1	CCDS4475.1
XKR4	NM_052898
XPNPEP1	CCDS7560.1
XPO7	NM_015024
XR_017918.1	ENST00000258651
XYLT2	CCDS11563.1
YLPM1	ENST00000238571
YN002_HUMAN	ENST00000334389
ZAN	NM_173059
ZBTB24	NM_014797
ZBTB33	CCDS14596.1
ZBTB7	CCDS12119.1
ZC3H12B	NM_001010888
ZC3HDC7	CCDS10550.1
ZDHHC4	CCDS5352.1
ZFHX1B	CCDS2186.1
ZFP36	CCDS12534.1
ZHX3	CCDS13315.1
ZIM3	NM_052882
ZMAT4	NM_024645
ZNF133	CCDS13134.1
ZNF136	NM_003437
ZNF148	CCDS3031.1
ZNF238	CCDS1623.1
ZNF253	ENST00000327867
ZNF31	NM_145238
ZNF333	CCDS12316.1
ZNF334	NM_199441
ZNF365	CCDS7264.1
ZNF423	NM_015069
ZNF443	NM_005815
ZNF451	CCDS4960.1
ZNF507	NM_014910
ZNF537	CCDS12421.1
ZNF560	CCDS12214.1
ZNF614	CCDS12847.1
ZNF638	CCDS1917.1
ZNF645	CCDS14205.1
ZNF648	ENST00000339948
ZNF682	NM_033196
ZYG11B	NM_024646

Note: Gene symbols are standard symbols assigned by Entrz Gene (http://www.ncbi.nlm.nih.gov/sites/entrez?db=gene). Accession IDs "NM_XXXX" are uniquely assigned to each gene by National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/sites/entrez?db=nuccore). Accession IDs "CCDSXXXX" are uniquely assigned to individual genes by National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/CCDS/). Accession IDs "ENSTXXXXXXXXXXX" are uniquely assigned to individual genes by *Ensembl* (http://www.ensembl.org/index.html).

Table 14. Genes containing somatic mutations in breast cancer adapted from the paper by Wood *et. al.* (Wood et al., 2007).

Gene Symbol	Accession ID
ABCA12	NM_173076
ABCA3	NM_001089.1
ABCA4	NM_000350.1
ABCB10	NM_012089.1
ABCB6	NM_005689.1
ABCB8	NM_007188.2
ABL2	NM_007314
ABLIM1	NM_002313.4
ABP1	NM_001091
ACADM	NM_000016.2
ACO2	NM_001098.2
ACY1	NM_000666.1
ADAM12	NM_003474.2
ADAMTS16	NM_139056
ADAMTS19	NM_133638.1
ADAR	NM_001111.2
ADH1B	NM_000668
ADHFE1	NM_144650.1
ADRA1A	NM_033302.1
AEGP	NM_206920.1
AGBL4	NM_032785
AGC1	NM_001135
AGRN	NM_198576
AHRR	NM_020731
AHSA2	NM_152392.1
AIM1	NM_001624
AKAP6	NM_004274.3
AKAP8	NM_005858.2
AKAP9	NM_005751.3
ALCAM	NM_001627
ALMS1	NM_015120
ALS2	NM_020919
ALS2CL	NM_147129.2
ALS2CR12	NM_139163.1
ALS2CR19	NM_152526
AMFR	NM_001144.3
AMIGO1	NM_020703
AMOTL1	NM_130847
AMPD2	NM_139156.1
AMPD2	NM_004037.5
ANAPC5	NM_016237.3
ANK1	NM_020476.1
ANK2	NM_001148.2
ANKRD28	NM_015199
ANKRD29	NM_173505.1
ANKRD30A	NM_052997.1
ANKRD5	NM_198798.1
AP1M1	NM_032493.2
AP3B2	NM_004644
APBB1	NM_145689
APC2	NM_005883.1
APCS	NM_001639.2
APOC4	NM_001646.1
APOL1	NM_145343.1
APPL	NM_012096.1
APXL	NM_001649.2
AQP8	NM_001169.2
ARC	NM_015193
ARFGAP3	NM_014570.3
ARFGEF2	NM_006420.1
ARFRP1	NM_003224.2
ARHGAP11A	NM_014783.2
ARHGAP25	NM_001007231
ARHGEF4	NM_015320.2
ARID1B	NM_017519.1
ARRB1	NM_020251
ARRDC3	NM_020801
ARV1	NM_022786.1
ASB11	NM_080873.1
ASGR1	NM_001671.2
ASL	NM_000048.2
ASTN2	NM_014010.3
ATCAY	NM_033064
ATF2	NM_001880.2
ATN1	NM_001940
ATP10A	NM_024490
ATP12A	NM_001676
ATP2A3	NM_174955.1
ATP6AP1	NM_001183
ATP6V0B	NM_004047.2
ATP8B1	NM_005603.1
ATP8B4	NM_024837
ATRN	NM_139321.1
ATXN2	NM_002973
AVPI1	NM_021732.1
AVPR2	NM_000054.2
B3GALNT2	NM_152490.1
B3GALT4	NM_003782
BAI1	NM_001702
BAP1	NM_004656.2
BAT2	NM_080686.1
BAT3	NM_080703.1
BAZ1A	NM_013448.2
BAZ1B	NM_032408.1
BC002942	NM_033200.1
BCAR1	NM_014567.2
BCCIP	NM_016567.2
BCL11A	NM_018014.2
BCORL1	NM_021946.2
BGN	NM_001711.3
BLR1	NM_001716.2
BMP1	NM_006129.2
BOC	NM_033254.2
BRCA1	NM_007296.1
BRCA2	NM_000059.1
BSPRY	NM_017688
C10orf30	NM_152751.1
C10orf38	NM_001010924
C10orf39	NM_194303.1
C10orf45	NM_031453.2
C10orf54	NM_022153
C10orf56	NM_153367.1
C10orf64	NM_173524
C11orf37	NM_001007543
C11orf9	NM_013279
C13orf24	NM_006346
C14orf100	NM_016475
C14orf101	NM_017799.2
C14orf121	NM_138360
C14orf155	NM_032135.2
C14orf161	NM_024764
C14orf21	NM_174913.1
C14orf29	NM_181814.1
C14orf46	NM_001024674
C17orf47	NM_001038704
C17orf64	NM_181707
C18orf19	NM_152352.1
C19orf28	NM_174983
C19orf6	NM_033420.2
C1orf190	NM_001013615
C1orf2	NM_006589.2
C1QB	NM_000491.2
C20orf103	NM_012261.2
C20orf121	NM_024331.2
C20orf161	NM_033421.2
C20orf177	NM_022106.1
C20orf23	NM_024704.3
C20orf44	NM_018244.3
C22orf19	NM_003678.3
C4orf14	NM_032313.2
C5orf14	NM_024715.2
C6orf102	NM_145027.3
C6orf145	NM_183373.2
C6orf174	NM_001012279
C6orf204	NM_206921.1
C6orf21	NM_001003693
C6orf213	NM_001010852
C6orf31	NM_030651.2
C7orf11	NM_138701.1
C9orf126	NM_173690
C9orf37	NM_032937
C9orf67	NM_032728.2
CACNA1B	NM_000718

Gene Symbol	Accession ID	Gene Symbol	Accession ID	Gene Symbol	Accession ID
CACNA1F	NM_005183	COL11A1	NM_001854.2	DIP2B	NM_173602
CACNA1G	NM_198385	COL12A1	NM_004370	DKFZP564B1023	NM_031306.1
CACNA1H	NM_021098	COL19A1	NM_001858.3	DKFZP564J102	NM_001006655
CACNA1I	NM_001003406	COL4A4	NM_000092	DKFZp761I2123	NM_031449
CACNA2D3	NM_018398	COL7A1	NM_000094.2	DKFZp779B1540	NM_001010903
CAMTA1	NM_015215	COMMD7	NM_053041	DKK3	NM_015881.4
CAPN11	NM_007058	COPG	NM_016128	DLEC1	NM_007335.1
CBFB	NM_001755.2	COQ9	NM_020312	DMD	NM_004006.1
CCDC16	NM_052857	CPA3	NM_001870.1	DNAH17	NM_003727
CCDC18	NM_206886	CPAMD8	NM_015692	DNAH5	NM_001369.1
CCDC66	NM_001012506	CPEB1	NM_030594	DNAH9	NM_001372.2
CD2	NM_001767.2	CPS1	NM_001875.2	DNAJA3	NM_005147.3
CD74	NM_001025159	CPSF3	NM_016207.2	DNAJA5	NM_194283.1
CD97	NM_001784	CROCC	NM_014675	DNAJC10	NM_018981
CDC27	NM_001256.2	CRR9	NM_030782.2	DNAJC13	NM_015268
CDH10	NM_006727.2	CRSP2	NM_004229.2	DNASE1L3	NM_004944.1
CDH20	NM_031891.2	CRTC1	NM_025021	DNM2	NM_004945
CDH8	NM_001796.2	CRX	NM_000554.2	DNM3	NM_015569
CDKL2	NM_003948.2	CRYAA	NM_000394.2	DOCK1	NM_001380
CDON	NM_016952.2	CSEN	NM_013434.3	DPAGT1	NM_001382.2
CDS1	NM_001263.2	CSMD1	NM_033225	DPAGT1	NM_203316.1
CENPE	NM_001813	CSMD3	NM_198123.1	DPP10	NM_020868
CENTB1	NM_014716.2	CSNK1D	NM_001893.3	DPP6	NM_130797
CENTD3	NM_022481.4	CSPP1	NM_024790	DPYD	NM_000110
CENTG1	NM_014770.2	CST4	NM_001899.2	DRIM	NM_014503.1
CEP290	NM_025114	CTF8	NM_001039690	DSCR6	NM_018962.1
CFHL5	NM_030787.1	CTNNA1	NM_001903	DSG2	NM_001943
CFL2	NM_138638.1	CTNNA2	NM_004389	DTNA	NM_032978.4
CGI-14	NM_015944.2	CTNND1	NM_001331	DTX3L	NM_138287.2
CGI-37	NM_016101.2	CUBN	NM_001081.2	DUOX1	NM_017434
CHD1	NM_001270	CUTC	NM_015960.1	DVL3	NM_004423.3
CHD5	NM_015557.1	CUTL1	NM_001913.2	DYSF	NM_003494.2
CHD7	NM_017780	CUTL2	NM_015267	ECT2	NM_018098.4
CHD8	NM_020920	CYP1A1	NM_000499.2	EDEM1	NM_014674
CHD9	NM_025134	CYP1A2	NM_000761	EDNRA	NM_001957.1
CHRND	NM_000751.1	CYP26A1	NM_000783.2	EEF1G	NM_001404
CIC	NM_015125.2	CYP2D6	NM_000106	EGFL6	NM_015507.2
CLCA2	NM_006536.3	CYP4A22	NM_001010969	EHBP1	NM_015252.2
CLCN1	NM_000083.1	DACH1	NM_080759	EHMT1	NM_024757.3
CLCN3	NM_001829	DAZAP1	NM_018959.2	EIF4A2	NM_001967.2
CLEC6A	NM_001007033	DBN1	NM_004395.2	EIF4B	NM_001417
CLSPN	NM_022111.2	DC2	NM_021227.2	EIF5	NM_183004.3
CLUAP1	NM_015041	DDO	NM_003649.2	ELA1	NM_001971.3
CMYA1	NM_194293.2	DDX10	NM_004398.2	ELAVL3	NM_001420
CMYA4	NM_173167.1	DDX18	NM_006773.3	ENPEP	NM_001977.2
CNGA2	NM_005140.1	DDX3X	NM_024005.1	EOMES	NM_005442.2
CNGB1	NM_001297	DEFB128	NM_001037732	EP400	NM_015409
CNNM4	NM_020184.2	DENND2A	NM_015689	EPC2	NM_015630
CNTN3	NM_020872	DGKB	NM_004080	ERCC3	NM_000122.1
CNTN5	NM_014361	DGKE	NM_003647.1	ERCC6	NM_000124.1
CNTN6	NM_014461.2	DGKG	NM_001346.1	EREG	NM_001432.1
COG3	NM_031431.2	DHX32	NM_018180.2	ETV5	NM_004454
COH1	NM_017890.3	DIP	NM_015124	EVI2A	NM_001003927

Gene Symbol	Accession ID
EVI5	NM_005665
EXOC2	NM_018303
EXOC5	NM_006544
EXOSC3	NM_016042
FAAH	NM_001441.1
FABP4	NM_001442.1
FAM44A	NM_148894.1
FAM47B	NM_152631.1
FAM80B	NM_020734
FANCA	NM_000135
FANCM	NM_020937
FARP1	NM_005766.1
FBXO40	NM_016298
FBXO8	NM_012180.1
FBXW11	NM_012300
FCHO1	NM_015122
FCMD	NM_006731.1
FCRH3	NM_052939.2
FEM1C	NM_020177.2
FER1L3	NM_133337
FGD3	NM_033086
FGD6	NM_018351
FGFR2	NM_022970.1
FHOD1	NM_013241.1
FHOD3	NM_025135
FLG2	NM_001014342
FLJ10241	NM_018035
FLJ10292	NM_018048.2
FLJ10324	NM_018059
FLJ10458	NM_018096.2
FLJ10726	NM_018195.2
FLJ10874	NM_018252.1
FLJ13089	NM_024953.2
FLJ13231	NM_023073
FLJ13479	NM_024706.3
FLJ13868	NM_022744.1
FLJ14503	NM_152780.2
FLJ14624	NM_032813.1
FLJ16331	NM_001004326
FLJ20152	NM_019000
FLJ20184	NM_017700.1
FLJ20422	NM_017814.1
FLJ20584	NM_017891.2
FLJ20604	NM_017897.1
FLJ21839	NM_021831.3
FLJ21945	NM_025203.1
FLJ23584	NM_024588
FLJ25955	NM_178821.1
FLJ31413	NM_152557.3
FLJ32115	NM_152321.1
FLJ32363	NM_198566.1
FLJ32440	NM_173685.1
FLJ32830	NM_152781.1
FLJ34521	NM_001039787
FLJ36180	NM_178556.3
FLJ36748	NM_152406
FLJ40342	NM_152347.3
FLJ40869	NM_182625.2
FLJ41821	NM_001001697
FLJ45455	NM_207386
FLJ46321	NM_001001670
FLJ46354	NM_198547.1
FLJ46481	NM_207405.1
FLJ90579	NM_173591.1
FLNA	NM_001456
FLNB	NM_001457.1
FLNC	NM_001458
FMNL3	NM_175736
FMOD	NM_002023
FN1	NM_002026.2
FNDC3B	NM_022763.2
FOLR2	NM_000803.2
FOXP2	NM_014491.1
FOXP4	NM_138457.1
FREM1	NM_144966
FRMPD1	NM_014907.1
FUCA1	NM_000147.2
FUS	NM_004960.1
FXR1	NM_005087.1
G3BP2	NM_203505.1
G6PC	NM_000151.1
GA17	NM_006360.2
GAB1	NM_002039.2
GABRA4	NM_000809.2
GABRP	NM_014211.1
GALK2	NM_001001556
GALNT17	NM_001034845
GALNT5	NM_014568.1
GALNTL2	NM_054110
GARNL1	NM_194301
GDF6	NM_001001557
GGA1	NM_013365.2
GGA3	NM_014001.2
GIMAP1	NM_130759.2
GIMAP8	NM_175571
GIOT-1	NM_153257
GIPC3	NM_133261
GJA8	NM_005267
GJB1	NM_000166.2
GKN1	NM_019617.2
GLG1	NM_012201
GLI1	NM_005269.1
GLT25D2	NM_015101.1
GMCL1L	NM_022471.2
GNB1L	NM_053004.1
GNPAT	NM_014236.1
GOLGA7	NM_016099
GOLGB1	NM_004487.1
GOLPH4	NM_014498.2
GORASP2	NM_015530
GP5	NM_004488.1
GPC1	NM_002081.1
GPC2	NM_152742.1
GPHB5	NM_145171
GPNMB	NM_002510.1
GPR115	NM_153838.1
GPR45	NM_007227.3
GPR7	NM_005285.1
GPR81	NM_032554.2
GRIK2	NM_021956.2
GRIK3	NM_000831.2
GRIN2C	NM_000835
GRIN2D	NM_000836.1
GRIPAP1	NM_207672
GRM6	NM_000843.2
GSDML	NM_018530.1
GSN	NM_000177.3
GTF2A1	NM_015859.2
GTF3C1	NM_001520
GUCY2F	NM_001522.1
HADHB	NM_000183.1
HCN3	NM_020897.1
HDAC4	NM_006037.2
HDAC7A	NM_015401.1
HDLBP	NM_203346.1
HEBP1	NM_015987
HEL308	NM_133636.1
HIST1H4L	NM_003546.2
HIST2H2AB	NM_175065.2
HK3	NM_002115.1
HLCS	NM_000411.4
HM13	NM_030789.2
HMG2L1	NM_001003681
HOMER2	NM_199331
HOOK1	NM_015888.3
HOOK2	NM_013312
HOOK3	NM_032410.2
HOXA3	NM_153631.1
HOXA4	NM_002141.2
HS3ST4	NM_006040
HSD11B1	NM_181755.1
HSD17B8	NM_014234.3
HSHIN1	NM_199324.1
HSPA14	NM_016299.1
HSPA1B	NM_005346
HSPC049	NM_014149
HTF9C	NM_182984.2
HUMCYT2A	NM_015848.1
HUWE1	NM_031407

Gene Symbol	Accession ID	Gene Symbol	Accession ID	Gene Symbol	Accession ID
ICAM5	NM_003259.2	KIAA1324	NM_020775.2	LOC652968	NM_001037666
IFNA2	NM_000605.2	KIAA1377	NM_020802	LOC88523	NM_033111
IFNB1	NM_002176.1	KIAA1414	NM_019024	LOC90529	NM_178122.2
IKBKAP	NM_003640.2	KIAA1632	NM_020964.1	LOC91461	NM_138370
IKBKB	NM_001556.1	KIAA1797	NM_017794	LOXL2	NM_002318
IL1RAPL2	NM_017416.1	KIAA1826	NM_032424	LPO	NM_006151
IL7R	NM_002185.2	KIAA1914	NM_001001936	LRBA	NM_006726.1
INA	NM_032727.2	KIAA1946	NM_177454	LRRC16	NM_017640
INHBE	NM_031479.3	KIBRA	NM_015238.1	LRRC4	NM_022143.3
IPLA2(GAMMA)	NM_015723	KIF14	NM_014875	LRRC43	NM_152759
IPO7	NM_006391	KIR2DS4	NM_012314.2	LRRC7	NM_020794.1
IQSEC2	NM_015075	KLHL10	NM_152467	LRRFIP1	NM_004735.1
IRF8	NM_002163.1	KLHL15	NM_030624	LUZP5	NM_017760
IRS4	NM_003604.1	KLK15	NM_017509.2	LYST	NM_000081
IRTA2	NM_031281.1	KPNA5	NM_002269.2	LYST	NM_001005736
ITGA9	NM_002207.1	KRTAP10-8	NM_198695.1	LZTS2	NM_032429.1
ITGAE	NM_002208	KRTAP20-1	NM_181615.1	MACF1	NM_012090.3
ITGAL	NM_002209	KTN1	NM_182926.1	MAGEA1	NM_004988.3
ITGB2	NM_000211.1	LAMA1	NM_005559	MAGEA4	NM_002362.3
ITPR1	NM_002222	LAMA2	NM_000426.2	MAGEB10	NM_182506
ITR	NM_180989.3	LAMA4	NM_002290	MAGEC2	NM_016249.2
JARID1B	NM_006618	LAMB4	NM_007356	MAGED2	NM_201222.1
JMJD1A	NM_018433.3	LAP1B	NM_015602.2	MAGEE1	NM_020932.1
JMJD1C	NM_004241	LDHB	NM_002300.3	MAGI1	NM_173515.1
JUP	NM_021991.1	LEPREL1	NM_018192.2	MANEA	NM_024641.2
KCNA5	NM_002234.2	LGALS2	NM_006498.1	MAOA	NM_000240.2
KCNC2	NM_139136.2	LHCGR	NM_000233.1	MAP1A	NM_002373
KCNJ1	NM_000220.2	LIP8	NM_053051.1	MAP3K6	NM_004672.3
KCNJ15	NM_170737.1	LIPE	NM_005357.2	MAPK13	NM_002754.3
KCNQ3	NM_004519	LLGL1	NM_004140	MAPKBP1	NM_014994
KEAP1	NM_203500.1	LMO6	NM_006150.3	MASP1	NM_001879
KIAA0100	NM_014680	LOC112703	NM_138411	MAZ	NM_002383
KIAA0143	NM_015137	LOC113179	NM_138422.1	MCAM	NM_006500
KIAA0256	NM_014701	LOC113828	NM_138435.1	MCART1	NM_033412.1
KIAA0284	NM_015005	LOC123876	NM_001010845	MCF2L2	NM_015078.2
KIAA0367	NM_015225	LOC126248	NM_173479.2	MCOLN1	NM_020533.1
KIAA0427	NM_014772.1	LOC200420	NM_145300	MDC1	NM_014641
KIAA0467	NM_015284	LOC220929	NM_182755.1	MED12	NM_005120
KIAA0513	NM_014732	LOC253012	NM_198151.1	MEF2C	NM_002397
KIAA0528	NM_014802	LOC255374	NM_203397	MFAP5	NM_003480.2
KIAA0664	NM_015229	LOC283849	NM_178516.2	MGC11332	NM_032718.2
KIAA0672	NM_014859	LOC339123	NM_001005920	MGC17299	NM_144626.1
KIAA0676	NM_015043.3	LOC339745	NM_001001664	MGC21688	NM_144635.3
KIAA0703	NM_014861	LOC340156	NM_001012418	MGC24047	NM_178840.2
KIAA0774	NM_001033602	LOC374955	NM_198546.1	MGC27019	NM_144705.2
KIAA0789	NM_014653	LOC388595	NM_001013641	MGC33212	NM_152773
KIAA0863	NM_014913	LOC388915	NM_001010902	MGC33370	NM_173807.2
KIAA0913	NM_015037	LOC389151	NM_001013650	MGC33657	NM_001029996
KIAA0934	NM_014974.1	LOC389549	NM_001024613	MGC34837	NM_152377.1
KIAA0999	NM_025164.3	LOC440925	NM_001013712	MGC42174	NM_152383
KIAA1012	NM_014939.2	LOC440944	NM_001013713	MGC5297	NM_024091.2
KIAA1117	NM_015018.2	LOC441070	NM_001013715	MIA2	NM_054024.3
KIAA1161	NM_020702	LOC646870	NM_001039790	MICAL1	NM_022765.2

Gene Symbol	Accession ID
MICAL-L1	NM_033386.1
MKLN1	NM_013255
MLL4	NM_014727
MLLT2	NM_005935.1
MMP10	NM_002425.1
MMP15	NM_002428.2
MOGAT1	NM_058165
MOSPD1	NM_019556.1
MPFL	NM_001025190
MRE11A	NM_005590.2
MSI1	NM_002442.2
MTA1	NM_004689
MTAC2D1	NM_152332.2
MTL5	NM_004923.2
MTMR3	NM_021090.2
MTMR8	NM_017677.2
MUC16	NM_024690
MUC2	NM_002457
MUF1	NM_006369.3
MULK	NM_018238.2
MYBPC2	NM_004533
MYCBP2	NM_015057
MYH1	NM_005963.2
MYH7B	NM_020884
MYH9	NM_002473.2
MYLC2PL	NM_138403
MYO15A	NM_016239
MYO18B	NM_032608
MYO1G	NM_033054
MYO7A	NM_000260
MYO9B	NM_004145
MYOD1	NM_002478.3
MYR8	NM_015011
MYST4	NM_012330.1
N4BP2	NM_018177.2
NAG6	NM_022742
NALP1	NM_014922
NALP14	NM_176822.2
NALP8	NM_176811.2
NALP9	NM_176820.2
NAV3	NM_014903
NCAM1	NM_000615
NCB5OR	NM_016230.2
NCOA6	NM_014071.2
NDRG2	NM_201541.1
NDST1	NM_001543
NDUFA2	NM_002488.2
NDUFA3	NM_004542.1
NDUFA8	NM_014222.2
NEB	NM_004543
NEDD4	NM_198400.1
NEF3	NM_005382.1
NET1	NM_005863.2
NF1	NM_000267.1
NF2	NM_000268.2
NFASC	NM_015090
NFIX	NM_002501
NFKB1	NM_003998.2
NFKBIA	NM_020529.1
NFKBIE	NM_004556
NFYC	NM_014223.2
NGLY1	NM_018297
NHS	NM_198270.2
NID2	NM_007361.2
NIPBL	NM_133433.2
NOD27	NM_032206.2
NOS2A	NM_000625.3
NOTCH1	NM_017617
NOTCH4	NM_004557
NOX5	NM_024505
NRCAM	NM_005010.2
NRK	NM_198465
NRXN3	NM_004796.3
NUFIP2	NM_020772
NUP133	NM_018230.2
NUP188	NM_015354
NUP205	NM_015135
NUP214	NM_005085.2
NUP98	NM_016320.2
NXN	NM_022463.3
NYD-SP21	NM_032597
OATL1	NM_002536
OBSCN	NM_052843.1
OCA2	NM_000275.1
ODZ1	NM_014253.1
OR10A2	NM_001004460
OR10H4	NM_001004465
OR12D3	NM_030959.2
OR1J2	NM_054107
OR1N1	NM_012363.1
OR1S1	NM_001004458
OR2AK2	NM_001004491
OR2M4	NM_017504
OR2W3	NM_001001957
OR2W5	NM_001004698
OR4D2	NM_001004707
OR52A1	NM_012375
OR52H1	NM_001005289
OR56A1	NM_001001917
OR5H1	NM_001005338
OR5J2	NM_001005492
OR5M11	NM_001005245
OR8B12	NM_001005195
OR8D2	NM_001002918
OR8I2	NM_001003750
OR9Q2	NM_001005283
OSBP2	NM_030758
OSBPL11	NM_022776.3
OTC	NM_000531.3
OTOF	NM_194323.1
P15RS	NM_018170.2
PADI3	NM_016233.1
PADI6	NM_207421
PANX2	NM_052839.2
PAPPA2	NM_020318
PARP1	NM_001618.2
PCDH19	NM_020766
PCDH20	NM_022843.2
PCDH8	NM_002590.2
PCDHA10	NM_031859
PCDHA11	NM_031861
PCDHA5	NM_031501
PCDHB15	NM_018935.2
PCDHGA1	NM_031993
PCDHGA3	NM_032011
PCDHGA6	NM_032086
PCDHGB1	NM_032095
PCDHGB5	NM_032099
PCM1	NM_006197
PCNT	NM_006031
PDCD11	NM_014976
PDCD4	NM_014456.3
PDCD6	NM_013232.2
PDE2A	NM_002599.1
PDLIM7	NM_005451.3
PDPR	NM_017990
PDZD7	NM_024895
PDZK2	NM_024791.2
PDZK4	NM_032512.2
PEBP4	NM_144962
PER1	NM_002616.1
PER2	NM_022817.1
PEX14	NM_004565
PFC	NM_002621.1
PFKFB4	NM_004567.2
PGBD3	NM_170753.1
PHACS	NM_032592.1
PHC1	NM_004426.1
PHF19	NM_015651
PHF7	NM_016483.4
PHKB	NM_000293.1
PIGN	NM_176787
PIGS	NM_033198.2
PIK3C2G	NM_004570
PIK3CA	NM_006218
PIK3R1	NM_181523.1
PIK3R4	NM_014602.1
PKD1L1	NM_138295
PKD1L2	NM_052892

Gene Symbol	Accession ID
PKDREJ	NM_006071.1
PKHD1L1	NM_177531
PKN1	NM_213560
PLA2G4A	NM_024420.1
PLB1	NM_153021
PLCB1	NM_015192.2
PLCB2	NM_004573
PLCD3	NM_133373
PLCG1	NM_002660.2
PLD2	NM_002663.2
PLEKHA8	NM_032639.2
PLEKHG2	NM_022835
PLOD1	NM_000302.2
PLS3	NM_005032.3
PLXNB1	NM_002673.3
PNCK	NM_198452.1
PNLIPRP1	NM_006229.1
PNPLA1	NM_001039725
PODXL	NM_001018111
POLH	NM_006502.1
POLR2F	NM_021974.2
POP1	NM_015029.1
POU2F1	NM_002697.2
POU4F2	NM_004575
PP	NM_021129.2
PPAPDC1A	NM_001030059
PPFIBP2	NM_003621
PPHLN1	NM_201439.1
PPM1E	NM_014906.3
PPM1F	NM_014634.2
PPP1R12A	NM_002480
PPP1R3A	NM_002711.2
PRDM13	NM_021620
PRDM4	NM_012406.3
PRDX5	NM_012094.3
PRKAA1	NM_006251.4
PRKAA2	NM_006252.2
PRODH	NM_016335.2
PRPF39	NM_017922.2
PRPF4B	NM_176800.1
PRPS1	NM_002764.2
PRPS1L1	NM_175886
PRRG1	NM_000950.1
PRSS7	NM_002772.1
PSD	NM_002779
PSME4	NM_014614
PSPC1	NM_018282
PSRC2	NM_144982
PTD004	NM_013341.2
PTHLH	NM_198964.1
PTPN14	NM_005401.3
PTPN6	NM_080548
PTPRC	NM_002838.2
PTRF	NM_012232.2
PURG	NM_013357.2
PUS1	NM_025215.3
PUS7	NM_019042
RAB41	NM_001032726
RABEP2	NM_024816
RAC2	NM_002872.3
RAI17	NM_020338.1
RANBP1	NM_002882.2
RANBP3	NM_007321
RANBP3	NM_007322
RAP1GA1	NM_002885.1
RAPH1	NM_213589.1
RARG	NM_000966.3
RASAL2	NM_170692.1
RASGRF2	NM_006909.1
RASL10B	NM_033315.2
RBAF600	NM_020765.1
RBM25	NM_021239
RCE1	NM_005133.1
RFC4	NM_181573.1
RFX2	NM_000635.2
RG9MTD2	NM_152292.2
RGL1	NM_015149.2
RGS22	NM_015668
RHAG	NM_000324.1
RHD	NM_016124.2
RIF1	NM_018151.1
RIMS1	NM_014989
RIMS2	NM_014677
RLTPR	NM_001013838
RNF123	NM_022064
RNF127	NM_024778.3
RNF149	NM_173647.2
RNU3IP2	NM_004704.2
ROBO3	NM_022370
ROR1	NM_005012.1
RP1L1	NM_178857
RPGRIP1	NM_020366
RPL3	NM_000967.2
RPRC1	NM_018067
RPS26	NM_001029
RPS6KA3	NM_004586.1
RPS9	NM_001013.2
RPUSD4	NM_032795.1
RREB1	NM_001003699
RSN	NM_002956.2
RTP1	NM_153708.1
RTTN	NM_173630
RUFY1	NM_025158.2
RYR1	NM_000540
RYR2	NM_001035
SAMD9	NM_017654
SAPS1	NM_014931
SATL1	NM_001012980
SBNO1	NM_018183.2
SCARF2	NM_153334.3
SCGB3A2	NM_054023.2
SCML1	NM_006746.2
SCN2A2	NM_021007
SCN3A	NM_006922
SCNN1B	NM_000336.1
SCP2	NM_002979.2
SEC31L1	NM_014933.2
SEMA3A	NM_006080.1
SEMA4B	NM_198925
SEMA4G	NM_017893.2
SEMA5B	NM_018987.1
SEMA6D	NM_153616
SEMA7A	NM_003612.1
SEPHS2	NM_012248
SERPINB1	NM_030666.2
SERPINB11	NM_080475
SERPINE2	NM_006216.2
SF3B1	NM_012433
SF3B2	NM_006842
SFRS1	NM_006924.3
SFRS16	NM_007056.1
SGKL	NM_013257.3
SH2D3A	NM_005490.1
SH3RF1	NM_020870
SHCBP1	NM_024745.2
SIGLEC5	NM_003830
SIPA1L1	NM_015556.1
SIX4	NM_017420.1
SKIP	NM_016532.2
SKIV2L	NM_006929.3
SLAMF1	NM_003037.1
SLC12A3	NM_000339.1
SLC16A2	NM_006517.1
SLC17A6	NM_020346.1
SLC22A2	NM_003058.2
SLC22A9	NM_080866.2
SLC25A30	NM_001010875
SLC35A2	NM_005660.1
SLC35F1	NM_001029858
SLC38A3	NM_006841
SLC39A12	NM_152725.1
SLC4A3	NM_005070.1
SLC6A3	NM_001044.2
SLC6A5	NM_004211.1
SLC7A7	NM_003982.2
SLC8A3	NM_033262.3
SLC8A3	NM_182932.1
SLC9A10	NM_183061
SLC9A2	NM_003048.3

Gene Symbol	Accession ID
SLCO2B1	NM_007256.2
SLFN13	NM_144682
SLICK	NM_198503.2
SMARCAL1	NM_014140.2
SMC4L1	NM_005496.2
SMC6L1	NM_024624.2
SMOX	NM_175839.1
SN	NM_023068.2
SNTG2	NM_018968
SNX25	NM_031953
SOHLH1	NM_001012415
SORBS1	NM_015385.1
SORCS1	NM_052918.2
SORL1	NM_003105.3
SOX13	NM_005686
SOX15	NM_006942
SP110	NM_004509.2
SPAG6	NM_012443.2
SPATS2	NM_023071
SPCS2	NM_014752
SPEN	NM_015001.2
SPG4	NM_014946.3
SPINK5	NM_006846
SPO11	NM_012444.2
SPOCD1	NM_144569.3
SPTA1	NM_003126
SPTAN1	NM_003127.1
SPTBN1	NM_178313
SPTLC1	NM_006415.2
SPTY2D1	NM_194285
SREBF2	NM_004599.2
SRGAP3	NM_014850.1
SSFA2	NM_006751.3
SSNA1	NM_003731.1
ST8SIA3	NM_015879
STAB1	NM_015136
STARD8	NM_014725.2
STAT1	NM_007315.2
STAT4	NM_003151.2
STATIP1	NM_018255.1
STRBP	NM_018387.2
STX12	NM_177424.1
STX5A	NM_003164.2
SULF2	NM_018837.2
SULT6B1	NM_001032377
SUPT3H	NM_181356
SURF1	NM_003172.2
SUSD3	NM_145006.2
SUV39H2	NM_024670.3
SYNE2	NM_182914.1
SYT3	NM_032298.1
SYTL2	NM_032943
TAC4	NM_170685
TACC2	NM_206862.1
TAF1	NM_004606.2
TAF1B	NM_005680
TA-KRP	NM_032505.1
TAS2R13	NM_023920.1
TAX1BP1	NM_006024.4
TBC1D19	NM_018317.1
TBC1D2B	NM_015079
TBX1	NM_005992.1
TBXAS1	NM_001061.2
TCEAL5	NM_001012979
TCF1	NM_000545.3
TCF7L1	NM_031283.1
TCFL1	NM_005997.1
TCP1	NM_030752.1
TCP10	NM_004610
TDRD6	NM_001010870
TECTA	NM_005422.1
TEK	NM_000459.1
TESK1	NM_006285.1
TESK2	NM_007170
TEX11	NM_031276
TFAP2D	NM_172238.1
TG	NM_003235
TGM3	NM_003245
THBS3	NM_007112.3
THG-1	NM_030935.3
TIAM2	NM_001010927
TIFA	NM_052864
TIMELESS	NM_003920.1
TLL1	NM_012464.3
TLN1	NM_006289
TLN2	NM_015059
TM4SF7	NM_003271.3
TMED1	NM_006858.2
TMEM123	NM_052932
TMEM132B	NM_052907
TMEM28	NM_015686
TMEM37	NM_183240
TMEM39A	NM_018266.1
TMEM62	NM_024956
TMEM63A	NM_014698
TMPRSS3	NM_024022.1
TMPRSS6	NM_153609.1
TNFRSF25	NM_003790.2
TNS	NM_022648.2
TOP1	NM_003286.2
TOP2B	NM_001068
TP53	NM_000546.2
TPM4	NM_003290.1
TPTE	NM_199261.1
TRAD	NM_007064.1
TREM1	NM_018643.2
TREML1	NM_178174.2
TREML4	NM_198153
TRIAD3	NM_207116
TRIF	NM_182919.1
TRIM25	NM_005082.3
TRIM29	NM_012101.2
TRIM36	NM_018700.2
TRIOBP	NM_001039141
TRIP12	NM_004238
TRPC4	NM_016179.1
TRPM5	NM_014555
TSN	NM_004622
TTC15	NM_016030.5
TTC21B	NM_024753
TTC3	NM_003316.2
TTC7A	NM_020458
TTN	NM_133378
TXNDC3	NM_016616.2
UBE2I	NM_194261.1
UBE2O	NM_022066
UGT1A9	NM_021027.2
UNQ9356	NM_207410.1
UQCR	NM_006830.2
USP29	NM_020903
USP34	NM_014709
USP54	NM_152586.2
UTP14C	NM_021645
UTS2R	NM_018949.1
VAV3	NM_006113.3
VEPH1	NM_024621.1
VGCNL1	NM_052867.1
VWF	NM_000552.2
WARS	NM_173701.1
WBP4	NM_007187.3
WBSCR28	NM_182504
WDR48	NM_020839
WDR53	NM_182627.1
WDR60	NM_018051
WDSOF1	NM_015420
WFDC1	NM_021197.2
WNK1	NM_018979.1
WNT2	NM_003391.1
XAB2	NM_020196
XBP1	NM_005080.2
XDH	NM_000379.2
XKR7	NM_001011718
XPO5	NM_020750
XPO7	NM_015024
YY2	NM_206923.1
ZBTB3	NM_024784.2
ZBTB39	NM_014830
ZCCHC14	NM_015144.1
ZCSL3	NM_181706.3

Gene Symbol	Accession ID	Gene Symbol	Accession ID	Gene Symbol	Accession ID
ZDHHC4	NM_018106.2	ZNF142	NM_005081	ZNF281	NM_012482.3
ZFHX4	NM_024721	ZNF161	NM_007146	ZNF318	NM_014345.1
ZFP64	NM_199427.1	ZNF183	NM_006978.1	ZNF37A	NM_001007094
ZFYVE26	NM_015346.2	ZNF22	NM_006963.2	ZNF425	NM_001001661
ZIC3	NM_003413.2	ZNF25	NM_145011.2	ZNF432	NM_014650.2
ZNF10	NM_015394.4	ZNF267	NM_003414	ZNF436	NM_030634.1
ZNF124	NM_003431	ZNF277	NM_021994.1	ZNF529	NM_020951
ZNF532	NM_018181.3				
ZNF541	NM_032255.1	ZNF674	NM_001039891		
ZNF546	NM_178544.2	ZNF694	NM_001012981		
ZNF548	NM_152909	ZNF707	NM_173831		
ZNF569	NM_152484.2	ZNF75A	NM_153028.1		
ZNF644	NM_201269.1	ZNHIT2	NM_014205.2		
ZNF646	NM_014699.2				

Note: Gene symbols are standard symbols assigned by Entrz Gene (http://www.ncbi.nlm.nih.gov/sites/entrez?db=gene). Accession IDs "NM_XXXX" are uniquely assigned to each gene by National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/sites/entrez?db=nuccore)

Table 15. Genes containing somatic mutations in colorectal cancer adapted from the paper by Wood *et. al.*(Wood et al., 2007).

Gene Symbol	Accession ID
ABCA1	NM_005502.2
ABCA6	NM_080284.2
ABCB1	NM_000927.3
ABCB11	NM_003742
ABCB5	NM_178559.3
ABCC5	NM_005688
ABCD4	NM_005050.1
ABI3BP	NM_015429
ACACA	NM_198839.1
ACIN1	NM_014977.1
ACSL4	NM_022977.1
ACSL5	NM_016234.3
AD026	NM_020683.5
ADAM19	NM_033274.1
ADAM29	NM_014269.2
ADAM33	NM_025220.2
ADAM8	NM_001109
ADAMTS1	NM_006988
ADAMTS15	NM_139055.1
ADAMTS18	NM_199355.1
ADAMTS20	NM_025003
ADAMTS20	NM_175851
ADAMTSL3	NM_207517.1
ADARB2	NM_018702.1
ADCY8	NM_001115.1
ADCY9	NM_001116
ADD3	NM_016824.2
ADORA1	NM_000674.1
AFMID	NM_001010982
AGTPBP1	NM_015239.1
AIM1	NM_001624
AKAP12	NM_005100.2
AKAP3	NM_006422.2
AKAP6	NM_004274.3
AKAP9	NM_005751.3
ALDH1L1	NM_012190.2
ALG9	NM_024740
ALK	NM_004304
ALS2CR11	NM_152525.3
ALS2CR8	NM_024744
AMACO	NM_198496.1
AMOTL2	NM_016201
AMPD1	NM_000036.1
AMPD3	NM_000480.1
ANAPC4	NM_013367.2
ANK2	NM_001148.2
ANKFN1	NM_153228
ANKRD11	NM_013275
ANKRD26	NM_014915
APBB2	NM_173075
APC	NM_000038.2
APG5L	NM_004849.1
API5	NM_006595
APIN	NM_017855.2
APOB	NM_000384.1
APOB48R	NM_182804
AQR	NM_014691
ARAF	NM_001654
ARFGEF1	NM_006421.2
ARHGEF1	NM_199002.1
ARHGEF10	NM_014629
ARHGEF9	NM_015185
ARR3	NM_004312.1
ASCC3L1	NM_014014.2
ASE-1	NM_012099.1
ATAD1	NM_032810.2
ATP11A	NM_032189
ATP11C	NM_173694.2
ATP12A	NM_001676
ATP13A1	NM_020410
ATP13A5	NM_198505
ATP13A5	NM_198505
ATP6V1E1	NM_001696.2
ATP8A2	NM_016529
ATP8B4	NM_024837
AVPR1B	NM_000707
AZI1	NM_001009811
BCAP29	NM_001008405
BCAS2	NM_005872.1
BCL11B	NM_022898.1
BCL9	NM_004326
BICD1	NM_001714.1
BMP6	NM_001718.2
BMPR2	NM_001204
BPIL1	NM_025227.1
BRAF	NM_004333.2
BRF1	NM_001519.2
BRUNOL6	NM_052840.2
BTBD4	NM_025224.1
BTF3L4	NM_152265
C10orf137	NM_015608.2
C10orf28	NM_014472
C10orf64	NM_173524
C10orf72	NM_144984.1
C12orf11	NM_018164.1
C13orf7	NM_024546
C14orf115	NM_018228.1
C15orf2	NM_018958.1
C17orf27	NM_020914
C17orf46	NM_152343
C17orf49	NM_174893
C18orf4	NM_032160.1
C1QR1	NM_012072.2
C20orf23	NM_024704.3
C21orf18	NM_017438.1
C21orf29	NM_144991.2
C21orf88	NM_153754
C2orf10	NM_194250.1
C2orf16	NM_032266
C2orf33	NM_020194.4
C4BPA	NM_000715.2
C4orf15	NM_024511
C6orf191	NM_001010876
C6orf29	NM_025257.1
C8B	NM_000066
C9orf21	NM_153698
Cab45	NM_016547.1
CACNA1A	NM_000068
CACNA1B	NM_000718
CACNA2D3	NM_018398
CACNB1	NM_199247.1
CACNB2	NM_201596.1
CAD	NM_004341.3
CAPN10	NM_023086.1
CAPN13	NM_144575
CAPN6	NM_014289.2
CARD12	NM_021209
CBFA2T3	NM_005187.4
CCAR1	NM_018237.2
CCNB3	NM_033031.1
CD109	NM_133493.1
CD248	NM_020404.2
CD99L2	NM_134445.1
CDC14A	NM_003672.2
CDH13	NM_001257
CDH18	NM_004934.2
CDH23	NM_022124
CDH6	NM_004932.2
CDKL5	NM_003159.1
CDO1	NM_001801.1
CDS1	NM_001263.2
CEACAM20	NM_198444
CENPF	NM_016343
CENPH	NM_022909.3
CENTB1	NM_014716.2
CENTB2	NM_012287
CENTD3	NM_022481.4
CGI-14	NM_015944.2
CHD7	NM_017780
CHD8	NM_020920
CHL1	NM_006614.2
CHR415SYT	NM_001014372
CHST8	NM_022467.3
CINP	NM_032630.2

Gene Symbol	Accession ID
CIR	NM_004882.3
CLIC2	NM_001289.3
CLSTN2	NM_022131.1
CLSTN3	NM_014718.2
CMKOR1	NM_020311.1
CNKSR2	NM_014927.2
CNOT6L	NM_144571
CNTN1	NM_001843.2
CNTN4	NM_175613.1
COL12A1	NM_004370
COL3A1	NM_000090.2
COL4A6	NM_001847.1
CORO1B	NM_020441.1
CORO2B	NM_006091.1
CPAMD8	NM_015692
CPE	NM_001873.1
CPO	NM_173077.1
CRB1	NM_201253.1
CRNKL1	NM_016652
CSDA	NM_003651.3
CSE1L	NM_001316.2
CSMD1	NM_033225
CSMD3	NM_198123.1
CSNK1A1L	NM_145203.2
CTCFL	NM_080618.2
CTEN	NM_032865.3
CTNNA1	NM_001903
CTNND2	NM_001332.2
CTSH	NM_004390.2
CUBN	NM_001081.2
CUTL1	NM_001913.2
CX40.1	NM_153368.1
CXorf53	NM_024332
CYP4F8	NM_007253
DACT1	NM_016651.4
DBC1	NM_014618.1
DCC	NM_005215.1
DCHS1	NM_003737.1
DDEFL1	NM_017707.2
DDHD2	NM_015214
DDI1	NM_001001711
DDIT3	NM_004083.3
DDN	NM_015086
DDX53	NM_182699
DEFA4	NM_001925.1
DEFB111	NM_001037497
DENND1C	NM_024898
DEPDC2	NM_024870.2
DGCR2	NM_005137
DHRS2	NM_005794.2
DJ167A19.1	NM_018982.3
DKFZp761I2123	NM_031449
DLG3	NM_021120.1
DMD	NM_004021.1
DMD	NM_004006.1
DMRTA1	NM_022160.1
DNAH1	NM_015512
DNAH11	NM_003777
DNAH3	NM_017539.1
DNAH8	NM_001371.1
DNAJC10	NM_018981
DNAJC6	NM_014787
DNALI1	NM_003462.3
DNAPTP6	NM_015535
DNASE1L3	NM_004944.1
DPEP1	NM_004413.1
DPP10	NM_020868
DPYSL2	NM_001386.3
DSCAML1	NM_020693.2
DSTN	NM_006870.2
DTNB	NM_183361
DUSP21	NM_022076.2
DUX4C	NM_001023569
EDA	NM_001399.3
EDD1	NM_015902
EFS	NM_005864.2
EIF2S2	NM_003908.2
EIF4G1	NM_198241.1
EML1	NM_004434
EML2	NM_012155.1
EN1	NM_001426.2
ENPP2	NM_006209.2
EPHA3	NM_005233.3
EPHA4	NM_004438.3
EPHA7	NM_004440.2
EPHB1	NM_004441
EPHB6	NM_004445.1
ERCC6	NM_000124.1
ESSPL	NM_183375
ETAA16	NM_019002.2
ETFDH	NM_004453.1
EVC2	NM_147127.2
EVL	NM_016337.1
EYA4	NM_004100.2
EZH2	NM_004456.3
F5	NM_000130.2
F8	NM_000132
FAM102B	NM_001010883
FAM19A5	NM_015381
FAM26A	NM_182494
FAM3A	NM_021806
FAM40A	NM_033088
FANCG	NM_004629.1
FAT	NM_005245
FBN1	NM_000138
FBN2	NM_001999
FBXL2	NM_012157.2
FBXO30	NM_032145.3
FBXW7	NM_033632.1
FCN1	NM_002003.2
FCN2	NM_004108.1
FERD3L	NM_152898.2
FGF13	NM_033642.1
FGF14	NM_175929.1
FHOD3	NM_025135
FIGN	NM_018086.1
FLJ10241	NM_018035
FLJ10404	NM_019057
FLJ10490	NM_018111
FLJ10521	NM_018125.2
FLJ10560	NM_018138.1
FLJ10665	NM_018173.1
FLJ10996	NM_019044.2
FLJ11000	NM_018295.1
FLJ12770	NM_032174.3
FLJ13305	NM_032180
FLJ14803	NM_032842
FLJ16171	NM_001004348
FLJ16542	NM_001004301
FLJ20294	NM_017749
FLJ20729	NM_017953.2
FLJ21019	NM_024927.3
FLJ21986	NM_024913
FLJ22679	NM_032227.1
FLJ25477	NM_199138.1
FLJ32252	NM_182510
FLJ32312	NM_144709.1
FLJ33534	NM_182586.1
FLJ34633	NM_152365.1
FLJ34922	NM_152270.2
FLJ35834	NM_178827.3
FLJ36119	NM_153254.1
FLJ38964	NM_173527
FLJ40142	NM_207435.1
FLJ42418	NM_001001695
FLJ43339	NM_207380.1
FLJ43980	NM_001004299
FLJ44653	NM_001001678
FLJ45273	NM_198461.1
FLJ46082	NM_207417.1
FLJ46154	NM_198462.1
FLNC	NM_001458
FMN2	NM_020066
FN1	NM_002026.2
FNDC1	NM_032532
FOLH1	NM_004476.1
FRAS1	NM_025074
FRAS1	NM_032863
FRMPD2	NM_152428.2

Gene Symbol	Accession ID
FRMPD4	NM_014728
FRY	NM_023037
FSTL5	NM_020116.2
FZD4	NM_012193.2
GAB4	NM_001037814
GABPB2	NM_016654.2
GABRA6	NM_000811.1
GALGT2	NM_153446.1
GALNS	NM_000512.2
GDAP1L1	NM_024034.3
GFI1	NM_005263
GFI1B	NM_004188.2
GHRHR	NM_000823.1
GJA8	NM_005267
GLB1	NM_000404
GLI3	NM_000168.2
GLIPR1	NM_006851.1
GMCL1L	NM_022471.2
GNAS	NM_000516.3
GNRH1	NM_000825
GPBP1	NM_022913
GPR112	NM_153834
GPR124	NM_032777.6
GPR158	NM_020752
GPR50	NM_004224
GPR8	NM_005286.2
GPR87	NM_023915.2
GPX1	NM_000581
GRID1	NM_017551
GRID2	NM_001510.1
GRIK1	NM_175611
GRIK3	NM_000831.2
GRM1	NM_000838.2
GTF2B	NM_001514.2
GUCY1A2	NM_000855.1
HAPIP	NM_003947.1
HAPLN1	NM_001884.2
HAT1	NM_003642.1
HBXIP	NM_006402.2
HCAP-G	NM_022346.2
HDC	NM_002112.1
HECTD1	NM_015382
HIC1	NM_006497
HIST1H1B	NM_005322.2
HIST1H1E	NM_005321.2
HIST1H2BM	NM_003521.2
HIVEP1	NM_002114
HIVEP3	NM_024503.1
HK3	NM_002115.1
HOXC9	NM_006897.1
HPS3	NM_032383.3
HR	NM_005144.2
HRH1	NM_000861.2
HS3ST4	NM_006040
HSPG2	NM_005529
HTR3C	NM_130770.2
HTR5A	NM_024012.1
HUWE1	NM_031407
IDH1	NM_005896.2
IGFBP3	NM_000598.2
IGSF22	NM_173588
IGSF9	NM_020789.2
IK	NM_006083
IL6ST	NM_002184.2
IQSEC3	NM_015232
IREM2	NM_181449.1
IRS2	NM_003749.2
IRS4	NM_003604.1
ISLR	NM_201526.1
ITGAE	NM_002208
ITGB3	NM_000212.2
ITPR1	NM_002222
K6IRS3	NM_175068.2
KCNA10	NM_005549.2
KCNB2	NM_004770.2
KCNC4	NM_004978.2
KCND3	NM_004980.3
KCNH4	NM_012285.1
KCNQ5	NM_019842.2
KCNT1	NM_020822
KCTD16	NM_020768
KDR	NM_002253.1
KIAA0182	NM_014615.1
KIAA0367	NM_015225
KIAA0415	NM_014855
KIAA0528	NM_014802
KIAA0555	NM_014790.3
KIAA0556	NM_015202
KIAA0789	NM_014653
KIAA0934	NM_014974.1
KIAA1078	NM_203459.1
KIAA1185	NM_020710.1
KIAA1285	NM_015694
KIAA1409	NM_020818.1
KIAA1468	NM_020854.2
KIAA1529	NM_020893
KIAA1727	NM_033393
KIAA1875	NM_032529
KIAA2022	NM_001008537
KIF13A	NM_022113
KL	NM_004795.2
KLF5	NM_001730.2
KLRF1	NM_016523
KRAS	NM_004985.3
KRT20	NM_019010.1
KRTAP10-2	NM_198693
KRTAP10-8	NM_198695.1
KSR2	NM_173598
LAMA1	NM_005559
LAMA4	NM_002290
LAMB3	NM_000228.1
LAMB4	NM_007356
LAMC1	NM_002293.2
LAS1L	NM_031206.2
LCN10	NM_001001712
LCN9	NM_001001676
LDB1	NM_003893.3
LDLRAD1	NM_001010978
LEF1	NM_016269.2
LGR6	NM_021636.1
LIFR	NM_002310.2
LIG1	NM_000234.1
LIG3	NM_013975.1
LILRB1	NM_006669
LMNB2	NM_032737.2
LMO7	NM_005358.3
LOC122258	NM_145248.2
LOC126147	NM_145807
LOC129531	NM_138798.1
LOC157697	NM_207332.1
LOC167127	NM_174914.2
LOC223075	NM_194300.1
LOC388199	NM_001013638
LOC91807	NM_182493.1
LPIN1	NM_145693.1
LPPR2	NM_022737.1
LRCH4	NM_002319
LRP1	NM_002332.1
LRP2	NM_004525.1
LRRC4	NM_022143.3
LRRN6D	NM_001004432
LRTM2	NM_001039029
LSP1	NM_001013253
LZTS2	NM_032429.1
MAMDC1	NM_182830
MAN2A2	NM_006122
MAP1B	NM_005909.2
MAP2	NM_002374.2
MAP2K7	NM_145185
MAPK8IP2	NM_012324
MARLIN1	NM_144720.2
MAST1	NM_014975
MCF2L2	NM_015078.2
MCM3AP	NM_003906.3
MCP	NM_172350.1
MCRS1	NM_006337.3
MED12L	NM_053002
MEF2C	NM_002397
MEGF6	NM_001409

Gene Symbol	Accession ID
MET	NM_000245
MFN1	NM_033540.2
MGC13125	NM_032725.2
MGC15730	NM_032880.2
MGC16943	NM_080663.1
MGC20470	NM_145053
MGC26733	NM_144992
MGC29671	NM_182538.3
MGC32124	NM_144611.2
MGC33407	NM_178525.2
MGC33846	NM_175885
MGC39325	NM_147189.1
MGC39545	NM_203452.1
MGC48628	NM_207491
MGC52022	NM_198563.1
MGC52282	NM_178453.2
MGC5242	NM_024033.1
MGC8685	NM_178012.3
MKRN3	NM_005664.1
MLF2	NM_005439.1
MLL3	NM_170606.1
MMP11	NM_005940.2
MMP2	NM_004530.1
MMRN2	NM_024756.1
MN1	NM_002430
MPO	NM_000250.1
MPP3	NM_001932
MRGPRE	NM_001039165
MRPL23	NM_021134
MS4A5	NM_023945.2
MTHFD1L	NM_015440.3
MUC1	NM_002456.3
MUC16	NM_024690
MYADML	NM_207329.1
MYO18B	NM_032608
MYO1B	NM_012223.2
MYO1D	NM_015194
MYO5C	NM_018728
MYOHD1	NM_001033579
MYR8	NM_015011
NALP7	NM_139176.2
NALP8	NM_176811.2
NAV3	NM_014903
NBEA	NM_015678
NCDN	NM_014284.1
NCR1	NM_004829.3
NDST3	NM_004784.1
NDUFA1	NM_004541.2
NEB	NM_004543
NELL1	NM_006157.2
NEUGRIN	NM_016645.1
NF1	NM_000267.1
NFATC1	NM_006162.3
NID	NM_002508.1
NLGN4X	NM_181332.1
NODAL	NM_018055.3
NOS3	NM_000603.2
NR3C2	NM_000901.1
NTNG1	NM_014917
NUP210	NM_024923
NUP210L	NM_207308
OBSCN	NM_052843.1
ODZ1	NM_014253.1
OLFM2	NM_058164.1
OMA1	NM_145243.2
OR10G3	NM_001005465
OR13F1	NM_001004485
OR1E2	NM_003554.1
OR2T33	NM_001004695
OR2T34	NM_001001821
OR4A16	NM_001005274
OR4K14	NM_001004712
OR51E1	NM_152430
OR51T1	NM_001004759
OR5H6	NM_001005479
OR5J2	NM_001005492
OR5K1	NM_001004736
OR6C1	NM_001005182
OR6C6	NM_001005493
OR6C75	NM_001005497
OR8K3	NM_001005202
OSBP	NM_002556.2
OSBPL5	NM_020896
OSBPL5	NM_145638
OTOP2	NM_178160.1
OVCH1	NM_183378
OVGP1	NM_002557.2
OXCT1	NM_000436.2
P2RX7	NM_002562.4
P2RY14	NM_014879.2
PAK6	NM_020168.3
PANK4	NM_018216.1
PAOX	NM_207128.1
PARP8	NM_024615.2
PBEF1	NM_005746.1
PBX4	NM_025245.1
PBXIP1	NM_020524.2
PCDH11X	NM_032968.2
PCDHA9	NM_014005
PCDHGA7	NM_032087
PCDHGB4	NM_032098
PCP4	NM_006198
PCSK2	NM_002594.2
PDE11A	NM_016953
PDGFD	NM_033135.2
PDILT	NM_174924.1
PDZD2	NM_178140
PDZRN3	NM_015009
PDZRN4	NM_013377.2
PEBP4	NM_144962
PEG3	NM_006210.1
PER1	NM_002616.1
PERQ1	NM_022574
PEX5L	NM_016559.1
PF6	NM_206996.1
PHIP	NM_017934.4
PHKB	NM_000293.1
PIGO	NM_032634.2
PIK3CA	NM_006218
PIK3R5	NM_014308.1
PKHD1	NM_138694.2
PKHD1L1	NM_177531
PKNOX1	NM_004571.3
PLA2G4B	NM_005090
PLA2G4D	NM_178034
PLB1	NM_153021
PLCG2	NM_002661
PLEC1	NM_201378
PLXND1	NM_015103
PNLIPRP2	NM_005396
PNMA3	NM_013364
PNPLA1	NM_001039725
PPM1F	NM_014634.2
PPP1R12A	NM_002480
PQBP1	NM_005710.1
PQLC1	NM_025078.3
PRDM9	NM_020227
PRF1	NM_005041.3
PRG2	NM_002728.4
PRIMA1	NM_178013.1
PRKCE	NM_005400.2
PRKCZ	NM_002744.2
PRKD1	NM_002742.1
PRKDC	NM_006904
PRNPIP	NM_024066
PRO0149	NM_014117.2
PROL1	NM_021225
PROS1	NM_000313.1
PRPS1	NM_002764.2
PRSS1	NM_002769.2
PRTG	NM_173814
PSMA2	NM_002787.3
PSMC5	NM_002805.4
PTEN	NM_000314
PTPRD	NM_130391.1
PTPRH	NM_002842
PTPRN2	NM_002847.2
PTPRS	NM_130853.1
PTPRU	NM_005704.2

Gene Symbol	Accession ID
PTPRZ1	NM_002851
PZP	NM_002864.1
QKI	NM_006775.1
RAB38	NM_022337.1
RAB5C	NM_201434.1
RABEP1	NM_004703
RALGDS	NM_006266.2
RAPGEF4	NM_007023
RARB	NM_000965.2
RASAL2	NM_170692.1
RASGRF2	NM_006909.1
RASGRP1	NM_005739
RASSF2	NM_170774.1
RASSF4	NM_032023.3
RAVER2	NM_018211
RB1CC1	NM_014781
RBM10	NM_005676.3
RBP3	NM_002900.1
RCN1	NM_002901.1
RDH13	NM_138412
RELN	NM_005045
RET	NM_020975.2
REV3L	NM_002912.1
RFC4	NM_181573.1
RHEB	NM_005614.2
RHPN1	NM_052924
RIC3	NM_024557.2
RIMBP2	NM_015347
RIMS2	NM_014677
RNF182	NM_152737.1
RNF31	NM_017999
RNPEPL1	NM_018226.2
ROBO1	NM_002941
ROBO2	NM_002942
RORA	NM_002943.2
RPA3	NM_002947.2
RPAP1	NM_015540.2
RPL6	NM_000970.2
RPS6KB1	NM_003161.1
RREB1	NM_001003699
RTN4	NM_207521.1
RUNX1T1	NM_175634.1
RYR2	NM_001035
SACS	NM_014363.3
SALL2	NM_005407
SALL3	NM_171999.1
SCN10A	NM_006514
SCN1A	NM_006920
SCN3B	NM_018400.2
SCN7A	NM_002976
SCNN1B	NM_000336.1
SCNN1G	NM_001039.2
SDBCAG84	NM_015966.2
SDCBP2	NM_080489.2
SDK1	NM_152744
SEC24B	NM_006323
SEC8L1	NM_021807.2
SEMA3D	NM_152754
SERPINA3	NM_001085
SETBP1	NM_015559.1
SEZ6	NM_178860
SF3A1	NM_005877.3
SFMBT2	NM_001029880
SFRS6	NM_006275.4
SGEF	NM_015595
SH3TC1	NM_018986.2
SHANK1	NM_016148.1
SHQ1	NM_018130
SIGLEC7	NM_014385.1
SKIP	NM_030623
SKIV2L	NM_006929.3
SLB	NM_015662.1
SLC11A2	NM_000617.1
SLC12A5	NM_020708.3
SLC12A7	NM_006598
SLC1A7	NM_006671.3
SLC22A15	NM_018420
SLC22A9	NM_080866.2
SLC26A10	NM_133489.1
SLC29A1	NM_004955.1
SLC33A1	NM_004733.2
SLC37A4	NM_001467
SLC39A7	NM_006979
SLC4A9	NM_031467
SLCO1A2	NM_134431.1
SLCO1B3	NM_019844.1
SLITRK4	NM_173078.2
SLITRK6	NM_032229
SMAD2	NM_005901.2
SMAD3	NM_005902.2
SMAD4	NM_005359.3
SMTN	NM_006932.3
SNRPB2	NM_198220.1
SNTG2	NM_018968
SNX5	NM_152227.1
SNX8	NM_013321.1
SOCS6	NM_004232.2
SORL1	NM_003105.3
SPOCK3	NM_016950
SPTBN2	NM_006946.1
ST8SIA4	NM_005668.3
STAB1	NM_015136
STAM	NM_003473.2
STK32C	NM_173575.2
STMN4	NM_030795.2
STX17	NM_017919.1
SUHW4	NM_001002843
SYNE1	NM_182961.1
SYNPO	NM_007286.3
SYT9	NM_175733.2
SYTL2	NM_206927
T3JAM	NM_025228.1
TAF1L	NM_153809
TAF2	NM_003184
TAIP-2	NM_024969.2
TA-KRP	NM_032505.1
TBC1D2B	NM_015079
TBX1	NM_005992.1
TBX15	NM_152380
TBX22	NM_016954.2
TCEB3B	NM_016427.2
TCERG1L	NM_174937.1
TCF3	NM_003200.1
TCF7L2	NM_030756.1
TCFL5	NM_006602.2
TCOF1	NM_000356.1
TFEC	NM_012252.1
TFG	NM_006070.3
TGFBR2	NM_003242.3
TGM2	NM_004613.2
TGM3	NM_003245
THAP9	NM_024672.2
THRAP1	NM_005121
TIAM1	NM_003253.1
TLR8	NM_138636.2
TLR9	NM_017442.2
TM7SF4	NM_030788.2
TMEM132B	NM_052907
TMEM16B	NM_020373
TMPRSS4	NM_019894
TNFRSF9	NM_001561.4
TNN	NM_022093
TNNI3K	NM_015978.1
TOP2A	NM_001067
TP53	NM_000546.2
TP53BP1	NM_005657.1
TPX2	NM_012112.4
TREX2	NM_080701
TRIM3	NM_033278.2
TRIM71	NM_001039111
TRMT5	NM_020810
TSKS	NM_021733.1
TSN	NM_004622
TSP-NY	NM_032573.3
TSPYL5	NM_033512
TTID	NM_006790.1
TTLL3	NM_015644.1
TTN	NM_133378
TTYH2	NM_032646

Gene Symbol	Accession ID
TXLNB	NM_153235
TYSND1	NM_173555
UBE3C	NM_014671
UGDH	NM_003359.1
UHRF2	NM_152896.1
UNC13B	NM_006377.2
UNC84B	NM_015374.1
UNQ689	NM_212557.1
UQCRC2	NM_003366.1
USP28	NM_020886
USP32	NM_032582
USP52	NM_014871.2
UTP14C	NM_021645
UTX	NM_021140.1
VEST1	NM_052958.1
VIM	NM_003380.1
VPS13A	NM_033305.1
WAC	NM_016628.2
WDR19	NM_025132
WDR49	NM_178824.3
WNK1	NM_018979.1
WNT16	NM_016087.2
WNT8B	NM_003393.2
WRN	NM_000553.2
XKR3	NM_175878
XPO4	NM_022459
XRCC1	NM_006297.1
YEATS2	NM_018023
ZAN	NM_173059
ZBTB8	NM_144621.2
ZD52F10	NM_033317.2
ZDHHC7	NM_017740.1
ZFHX1B	NM_014795.2
ZFHX4	NM_024721
ZFPM2	NM_012082
ZNF155	NM_198089.1
ZNF217	NM_006526.2
ZNF232	NM_014519.2
ZNF235	NM_004234
ZNF262	NM_005095.2
ZNF291	NM_020843
ZNF43	NM_003423.1
ZNF435	NM_025231.1
ZNF442	NM_030824.1
ZNF471	NM_020813.1
ZNF480	NM_144684.1
ZNF521	NM_015461
ZNF536	NM_014717
ZNF540	NM_152606.2
ZNF560	NM_152476.1
ZNF568	NM_198539
ZNF572	NM_152412.1
ZNF582	NM_144690
ZNF624	NM_020787.1
ZNF659	NM_024697.1
ZNF714	NM_182515
ZNHIT1	NM_006349.2
ZNRF4	NM_181710
ZSCAN5	NM_024303.1
ZZZ3	NM_015534.3

Note: Gene symbols are standard symbols assigned by Entrz Gene (http://www.ncbi.nlm.nih.gov/sites/entrez?db=gene). Accession IDs "NM_XXXX" are uniquely assigned to each gene by National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/sites/entrez?db=nuccore).

Claims

1. A method for aiding in the diagnosis, prognosis, or monitoring of a prostate cancer in a human male subject, comprising the steps of:

(a) isolating a microvesicle fraction from a urine sample from a human male subject; and

(b) analyzing mRNA in the microvesicle fraction for the presence of a gene specific to prostate cancer, wherein the gene is PCA-3 and/or TMPRSS2-ERG, to thereby indicate the presence or absence of prostate cancer in the subject.

2. The method of claim 1, wherein step (a) is performed by size exclusion chromatography, density gradient centrifugation, differential centrifugation, nanomembrane ultrafiltration, immunoabsorbent capture, affinity purification, microfluidic separation, or combinations thereof.

3. The method of claim 1 or claim 2, wherein step (b) is performed by microarray analysis, PCR, hybridization with allele-specific probes, enzymatic mutation detection, ligation chain reaction (LCR), oligonucleotide ligation assay (OLA), flow-cytometric heteroduplex analysis, chemical cleavage of mismatches, mass spectrometry, nucleic acid sequencing, single strand conformation polymorphism (SSCP), denaturing gradient gel electrophoresis (DGGE), temperature gradient gel electrophoresis (TGGE), restriction fragment polymorphisms, serial analysis of gene expression (SAGE), or combinations thereof.

4. The method of any previous claim, wherein the isolating step includes filtering.

5. The method of claim 4, wherein filtering is through a 0.8 micron filter.

6. The method of claim 4 or claim 5, wherein filtering is followed by ultrafiltration concentration.

7. The method of any of claims 4 to 6, wherein filtering is followed by ultracentrifugation.