Office de la Propriété Intellectuelle du Canada
Un organisme d'Industrie Canada

Canadian Intellectual Property Office
An agency of Industry Canada

CA 2633754 C 2013/06/18

(11)(21) 2 633 754

(12) BREVET CANADIEN
CANADIAN PATENT

(13) C

(86) Date de dépôt PCT/PCT Filing Date: 2007/01/03
(87) Date publication PCT/PCT Publication Date: 2007/07/19
(45) Date de délivrance/Issue Date: 2013/06/18
(85) Entrée phase nationale/National Entry: 2008/06/18
(86) N° demande PCT/PCT Application No.: US 2007/000159
(87) N° publication PCT/PCT Publication No.: 2007/081740
(30) Priorité/Priority: 2006/01/05 (US60/756,585)

(51) Cl.Int./Int.Cl. *C12Q 1/68* (2006.01), *C40B 30/04* (2006.01), *C40B 40/06* (2006.01), *G01N 33/48* (2006.01)

(72) Inventeurs/Inventors:
CROCE, CARLO M., US;
CALIN, GEORGE A., US;
VOLINIA, STEFANO, IT

(73) Propriétaire/Owner:
THE OHIO STATE UNIVERSITY RESEARCH FOUNDATION, US

(74) Agent: G. RONALD BELL & ASSOCIATES

(54) Titre : PROCEDES ET COMPOSITIONS FAISANT INTERVENIR DES MICRO-ARN POUR LE DIAGNOSTIC ET LE TRAITEMENT DES CANCERS SOLIDES

(54) Title: MICRORNA-BASED METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT OF SOLID CANCERS

(57) Abrégé/Abstract:
The present invention provides novel methods and compositions for the diagnosis and treatment of solid cancers. The invention also provides methods of identifying inhibitors of tumorigenesis.

ABSTRACT

The present invention provides novel methods and compositions for the diagnosis and treatment of solid cancers. The invention also provides methods of identifying inhibitors of tumorigenesis.

MicroRNA-BASED METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT OF SOLID CANCERS

BACKGROUND OF THE INVENTION

Cancer, the uncontrolled growth of malignant cells, is a major health problem of the modern medical era and is one of the leading causes of death in developed countries. In the United States, one in four deaths is caused by cancer (Jemal, A. *et al.*, *CA Cancer J. Clin. 52*:23-47 (2002)). Among cancers, those that arise from organs and solid tissues, known as solid cancers (e.g., colon cancer, lung cancer, breast cancer, stomach cancer, prostate cancer, pancreatic cancer) are among the most-commonly identified human cancers.

For example, prostate cancer is the most frequently diagnosed noncutaneous malignancy among men in industrialized countries, and, in the United States, 1 in 8 men will develop prostate cancer during his life (Simard, J. *et al.*, *Endocrinology 143(6)*:2029-40 (2002)). The incidence of prostate cancer has dramatically increased over the last decades and prostate cancer is now a leading cause of death in the United States and Western Europe (Peschel, R.E. and J.W. Colberg, *Lancet 4*:233-41 (2003); Nelson, W.G. *et al.*, *N. Engl. J. Med. 349(4)*:366-81 (2003)). An average 40% reduction in life expectancy affects males with prostate cancer. If detected early, prior to metastasis and local spread beyond the capsule, prostate cancer can often times be cured (e.g., using surgery). However, if diagnosed after spread and metastasis from the prostate, prostate cancer is typically a fatal disease with low cure rates. While prostate-specific antigen (PSA)-based screening has aided early diagnosis of prostate cancer, it is neither highly sensitive nor specific (Punglia *et al.*, *N. Engl. J. Med. 349(4)*:335-42

(2003)). This means that a high percentage of false negative and false positive diagnoses are associated with the test. The consequences are both many instances of missed cancers and unnecessary follow-up biopsies for those without cancer.

Breast cancer remains the second leading cause of cancer-related deaths in women, affecting more than 180,000 women in the United States each year. For women in North America, the life-time odds of getting breast cancer are now one in eight. Although the discovery of BRCA1 and BRCA2 were important steps in identifying key genetic factors involved in breast cancer, it has become clear that mutations in BRCA1 and BRCA2 account for only a fraction of inherited susceptibility to breast cancer (Nathanson, K.L., *et al.*, *Human Mol. Gen. 10(7)*:715-720 (2001); Anglican Breast Cancer Study Group. *Br. J. Cancer 83(10)*:1301-08 (2000); and Syrjakoski, K., *et al.*, *J. Natl. Cancer Inst. 92*:1529-31 (2000)). Despite considerable research into therapies for breast cancer, breast cancer remains difficult to diagnose and treat effectively, and the high mortality observed in breast cancer patients indicates that improvements are needed in the diagnosis, treatment and prevention of the disease.

Excluding skin cancer, colorectal cancer is the third most frequently diagnosed cancer in the United States and Canada (after lung and breast in women, and lung and prostate in men). The American Cancer Society estimates that there will be approximately 145,000 new cases of colorectal cancer diagnosed in the U.S. in 2005 (Cancer Facts and Figures 2005. Atlanta, GA: American Cancer Society, 2005.). Colorectal cancer is the second leading cause of cancer death among men and women in the United States and Canada (after lung cancer).

The annual incidence of pancreatic cancer is nearly equivalent to the annual mortality, estimated to be 31,860 and 31,270, respectively, in the U.S. in 2004 (Cancer Facts and Figures 2004. Atlanta, GA: American Cancer Society, 2004.). Patients with locally advanced and metastatic pancreatic cancer have poor prognoses, and diagnosis generally occurs too late for surgery or radiotherapy to be curative (Burr, H.A., *et al.*, *The Oncologist 10*(3): 183-190, (2005)). Chemotherapy can provide relief of symptoms for some patients with advanced pancreatic cancer, but its impact on survival has been modest to date.

In the United States, more than 20,000 individuals are diagnosed with stomach (gastric) cancer each year. The American Cancer Society estimates that there will be 22,710 new cases of colorectal cancer diagnosed in the U.S. in 2004 (Cancer Facts and Figures 2004. Atlanta, GA: American Cancer Society, 2004.). Because stomach cancer may occur without symptoms, it may be in advanced stages by the time the diagnosis is made. Treatment is then directed at making the patient more comfortable and improving quality of life.

Lung cancer causes more deaths worldwide than any other form of cancer (Goodman, G.E., *Thorax* 57:994-999 (2002)). In the United States, lung cancer is the primary cause of cancer death among both men and women. In 2002, the death rate from lung cancer was an estimated 134,900 deaths, exceeding the combined total for breast, prostate and colon cancer. *Id.* Lung cancer is also the leading cause of cancer death in all European countries, and numbers of lung cancer-related deaths are rapidly increasing in developing countries as well.

The five-year survival rate among all lung cancer patients, regardless of the stage of disease at diagnosis, is only about 13%. This contrasts with a five-year survival rate of 46% among cases detected while the disease is still localized. However, only 16% of lung cancers are discovered before the disease has spread. Early detection is difficult as clinical symptoms are often not observed until the disease has reached an advanced stage. Despite research into therapies for this and other cancers, lung cancer remains difficult to diagnose and treat effectively.

Clearly, the identification of markers and genes that are responsible for susceptibility to particular forms of solid cancer (e.g., prostate cancer, breast cancer, lung cancer, stomach cancer, colon cancer, pancreatic cancer) is one of the major challenges facing oncology today. There is a need to identify means for the early detection of individuals that have a genetic susceptibility to cancer so that more aggressive screening and intervention regimens may be instituted for the early detection and treatment of cancer. Cancer genes may also reveal key molecular pathways that may be manipulated (e.g., using small or large molecule weight drugs) and may lead to more effective treatments regardless of the cancer stage when a particular cancer is first diagnosed.

MicroRNAs are a class of small, non-coding RNAs that control gene expression by hybridizing to and triggering either translational repression or, less frequently, degradation of a messenger RNA (mRNA) target. The discovery and study of miRNAs has revealed miRNA-mediated gene regulatory mechanisms that play important roles in organismal development and various cellular processes, such as cell differentiation, cell growth and cell death (Cheng, A.M., *et al.*, *Nucleic Acids Res. 33*:1290-1297 (2005)). Recent studies suggest that aberrant expression of particular miRNAs may be involved in human diseases, such as neurological disorders (Ishizuka, A., *et al.*, *Genes Dev. 16*:2497-2508 (2002)) and cancer. In particular, misexpression of miR-16-1 and/or miR-15a has been found in human chronic lymphocytic leukemias (Calin, G.A., *et al.*, *Proc. Natl. Acad. Sci. U.S.A. 99*:15524-15529 (2002)).

Clearly, there is a great need in the art for improved methods for detecting and treating solid cancers (e.g., prostate cancer, breast cancer, lung cancer, stomach cancer, colon cancer, pancreatic cancer). The present invention provides novel methods and compositions for the diagnosis and treatment of solid cancers.

SUMMARY OF THE INVENTION

The present invention is based, in part, on the identification of specific miRNAs that have altered expression levels in particular solid cancers.

Accordingly, the invention encompasses methods of diagnosing whether a subject has, or is at risk for developing, a solid cancer. According to the methods of the invention, the level of at least one miR gene product in a test sample from the subject is compared to the level of a corresponding miR gene product in a control sample. An alteration (e.g., an increase, a decrease) in the level of the miR gene product in the test sample, relative to the level of a corresponding miR gene product in a control sample, is indicative of the subject either having, or being at risk for developing, a solid cancer. The solid cancer can be any cancer that arises from organs and solid tissues. In certain embodiments, the solid cancer is stomach cancer, breast cancer, pancreatic cancer, colon cancer, lung cancer or prostate cancer. In particular embodiments, the solid cancer is not breast cancer, lung cancer, prostate cancer, pancreatic cancer or gastrointestinal cancer.

In one embodiment, the at least one miR gene product measured in the test sample is selected from the group consisting of miR-21, miR-191, miR-17-5p and combinations thereof. In another embodiment, the at least one miR gene product measured in the test sample is selected from the group consisting of miR-21, miR-17-5p, miR-191, miR-29b-2, miR-223, miR-128b, miR-199a-1, miR-24-1, miR-24-2, miR-146, miR-155, miR-181b-1, miR-20a, miR-107, miR-32, miR-92-2, miR-214, miR-30c, miR-25, miR-221, miR-106a and combinations thereof.

In one embodiment, the solid cancer is breast cancer or lung cancer and the at least one miR gene product measured in the test sample is selected from the group consisting of miR-210, miR-213 and a combination thereof.

In another embodiment, the solid cancer is colon cancer, stomach cancer, prostate cancer or pancreas cancer and the at least one miR gene product measured in the test sample is miR-218-2.

In a certain embodiment, the solid cancer is breast cancer and the at least one miR gene product measured in the test sample is selected from the group consisting of miR-125b-1, miR-125b-2, miR-145, miR-21 and combinations thereof. In a related embodiment, the solid cancer is breast cancer and the at least one miR gene product in the test sample is selected from the group consisting of miR-21, miR-29b-2, miR-146, miR-125b-2, miR-125b-1, miR-10b, miR-145, miR-181a, miR-140, miR-213, miR-29a prec, miR-181b-1, miR-199b, miR-29b-1, miR-130a, miR-155, let-7a-2, miR-205, miR-29c, miR-224, miR-100, miR-31, miR-30c, miR-17-5p, miR-210, miR-122a, miR-16-2 and combinations thereof.

In another embodiment, the solid cancer is colon cancer and the at least one miR gene product in the test sample is selected from the group consisting of miR-24-1, miR-29b-2, miR-20a, miR-10a, miR-32, miR-203, miR-106a, miR-17-5p, miR-30c, miR-223, miR-126*, miR-128b, miR-21, miR-24-2, miR-99b prec, miR-155, miR-213, miR-150, miR-107, miR-191, miR-221, miR-9-3 and combinations thereof.

In yet another embodiment, the solid cancer is lung cancer and the miR gene product in the test sample is selected from the group consisting of miR-21, miR-205, miR-200b, miR-9-1, miR-210, miR-148, miR-141, miR-132, miR-215, miR-128b, let-7g, miR-16-2, miR-129-1/2 prec, miR-126*, miR-142-as, miR-30d, miR-30a-5p, miR-7-2, miR-199a-1, miR-127, miR-34a prec, miR-34a, miR-136, miR-202, miR-196-2,

miR-199a-2, let-7a-2, miR-124a-1, miR-149, miR-17-5p, miR-196-1 prec, miR-10a, miR-99b prec, miR-196-1, miR-199b, miR-191, miR-195, miR-155 and combinations thereof.

In an additional embodiment, the solid cancer is pancreatic cancer and the at least one miR gene product measured in the test sample is selected from the group consisting of miR-103-1, miR-103-2, miR-155, miR-204 and combinations thereof. In a related embodiment, the solid cancer is pancreatic cancer and the miR gene product in the test sample is selected from the group consisting of miR-103-2, miR-103-1, miR-24-2, miR-107, miR-100, miR-125b-2, miR-125b-1, miR-24-1, miR-191, miR-23a, miR-26a-1, miR-125a, miR-130a, miR-26b, miR-145, miR-221, miR-126*, miR-16-2, miR-146, miR-214, miR-99b, miR-128b, miR-155, miR-29b-2, miR-29a, miR-25, miR-16-1, miR-99a, miR-224, miR-30d, miR-92-2, miR-199a-1, miR-223, miR-29c, miR-30b, miR-129-1/2, miR-197, miR-17-5p, miR-30c, miR-7-1, miR-93-1, miR-140, miR-30a-5p, miR-132, miR-181b-1, miR-152 prec, miR-23b, miR-20a, miR-222, miR-27a, miR-92-1, miR-21, miR-129-1/2 prec, miR-150, miR-32, miR-106a, miR-29b-1 and combinations thereof.

In another embodiment, the solid cancer is prostate cancer and the miR gene product in the test sample is selected from the group consisting of let-7d, miR-128a prec, miR-195, miR-203, let-7a-2 prec, miR-34a, miR-20a, miR-218-2, miR-29a, miR-25, miR-95, miR-197, miR-135-2, miR-187, miR-196-1, miR-148, miR-191, miR-21, let-7i, miR-198, miR-199a-2, miR-30c, miR-17-5p, miR-92-2, miR-146, miR-181b-1 prec, miR-32, miR-206, miR-184 prec, miR-29a prec, miR-29b-2, miR-149, miR-181b-1, miR-196-1 prec, miR-93-1, miR-223, miR-16-1, miR-101-1, miR-124a-1, miR-26a-1, miR-214, miR-27a, miR-24-1, miR-106a, miR-199a-1 and combinations thereof.

In yet another embodiment, the solid cancer is stomach cancer and the miR gene product in the test sample is selected from the group consisting of miR-223, miR-21, miR-218-2, miR-103-2, miR-92-2, miR-25, miR-136, miR-191, miR-221, miR-125b-2, miR-103-1, miR-214, miR-222, miR-212 prec, miR-125b-1, miR-100, miR-107, miR-92-1, miR-96, miR-192, miR-23a, miR-215, miR-7-2, miR-138-2, miR-24-1, miR-99b, miR-33b, miR-24-2 and combinations thereof.

The level of the at least one miR gene product can be measured using a variety of techniques that are well known to those of skill in the art (e.g., quantitative or semi-

quantitative RT-PCR, Northern blot analysis, solution hybridization detection). In a particular embodiment, the level of at least one miR gene product is measured by reverse transcribing RNA from a test sample obtained from the subject to provide a set of target oligodeoxynucleotides, hybridizing the target oligodeoxynucleotides to one or more miRNA-specific probe oligonucleotides (e.g., hybridzing to a microarray that comprises several miRNA-specific probe oligonucleotides) to provide a hybridization profile for the test sample, and comparing the test sample hybridization profile to a hybridization profile from a control sample. An alteration in the signal of at least one miRNA in the test sample relative to the control sample is indicative of the subject either having, or being at risk for developing, a solid cancer. In a particular embodiment, target oligonucleotides are hybridized to a microarray comprising miRNA-specific probe oligonucleotides for one or more miRNAs selected from the group consisting of miR-21, miR-17-5p, miR-191, miR-29b-2, miR-223, miR-128b, miR-199a-1, miR-24-1, miR-24-2, miR-146, miR-155, miR-181b-1, miR-20a, miR-107, miR-32, miR-92-2, miR-214, miR-30c, miR-25, miR-221, miR-106a and combinations thereof.

The invention also encompasses methods of inhibiting tumorigenesis in a subject who has, or is suspected of having, a solid cancer (e.g., prostate cancer, stomach cancer, pancreatic cancer, lung cancer, breast cancer, colon cancer), wherein at least one miR gene product is deregulated (e.g., down-regulated, up-regulated) in the cancer cells of the subject. When the at least one isolated miR gene product is down-regulated in the cancer cells, the method comprises administering an effective amount of an isolated miR gene product, an isolated variant or a biologically-active fragment of the miR gene product or variant, such that proliferation of cancer cells in the subject is inhibited. In a further embodiment, the at least one isolated miR gene product is selected from the group consisting of miR-145, miR-155, miR-218-2 and combinations thereof. In a particular embodiment, the miR gene product is not miR-15a or miR-16-1. When the at least one isolated miR gene product is up-regulated in the cancer cells, the method comprises administering to the subject an effective amount of at least one compound for inhibiting expression of the at least one miR gene product (referred to herein as a "miR expression-inhibition compound"), such that proliferation of cancer cells in the subject is inhibited. In a particular embodiment, the at least one miR

expression-inhibition compound is specific for a miR gene product selected from the group consisting of miR-21, miR-17-5p, miR-191, miR-29b-2, miR-223, miR-128b, miR-199a-1, miR-24-1, miR-24-2, miR-146, miR-155, miR-181b-1, miR-20a, miR-107, miR-32, miR-92-2, miR-214, miR-30c, miR-25, miR-221, miR-106a and combinations thereof.

In a related embodiment, the methods of inhibiting tumorigenesis in a subject additionally comprise the step of determining the amount of at least one miR gene product in cancer cells from the subject, and comparing that level of the miR gene product in the cells to the level of a corresponding miR gene product in control cells. If expression of the miR gene product is deregulated (e.g., down-regulated, up-regulated) in cancer cells, the methods further comprise altering the amount of the at least one miR gene product expressed in the cancer cells. In one embodiment, the amount of the miR gene product expressed in the cancer cells is less than the amount of the miR gene product expressed in a control cell (e.g., control cells), and an effective amount of the down-regulated miR gene product, isolated variant or biologically-active fragment of the miR gene product or variant, is administered to the subject. Suitable miR gene products for this embodiment include miR-145, miR-155, miR-218-2 and combinations thereof, among others. In a particular embodiment, the miR gene product is not miR-15a or miR-16-1. In another embodiment, the amount of the miR gene product expressed in the cancer cells is greater than the amount of the miR gene product expressed in the control cell (e.g., control cells), and an effective amount of at least one compound for inhibiting expression of the at least one up-regulated miR gene product is administered to the subject. Suitable compounds for inhibiting expression of the at least one miR gene product include, but are not limited to, compounds that inhibit the expression of miR-21, miR-17-5p, miR-191, miR-29b-2, miR-223, miR-128b, miR-199a-1, miR-24-1, miR-24-2, miR-146, miR-155, miR-181b-1, miR-20a, miR-107, miR-32, miR-92-2, miR-214, miR-30c, miR-25, miR-221, miR-106a and combinations thereof.

The invention further provides pharmaceutical compositions for treating solid cancers (e.g., prostate cancer, stomach cancer, pancreatic cancer, lung cancer, breast cancer, colon cancer). In one embodiment, the pharmaceutical compositions comprise at least one isolated miR gene product and a pharmaceutically-acceptable carrier. In a

particular embodiment, the at least one miR gene product corresponds to a miR gene product that has a decreased level of expression in cancer cells relative to control cells. In certain embodiments the isolated miR gene product is selected from the group consisting of miR-145, miR-155, miR-218-2 and combinations thereof.

In another embodiment, the pharmaceutical compositions of the invention comprise at least one miR expression-inhibition compound and a pharmaceutically-acceptable carrier. In a particular embodiment, the at least one miR expression-inhibition compound is specific for a miR gene product whose expression is greater in cancer cells than in control cells. In certain embodiments, the miR expression-inhibition compound is specific for one or more miR gene products selected from the group consisting of miR-21, miR-17-5p, miR-191, miR-29b-2, miR-223, miR-128b, miR-199a-1, miR-24-1, miR-24-2, miR-146, miR-155, miR-181b-1, miR-20a, miR-107, miR-32, miR-92-2, miR-214, miR-30c, miR-25, miR-221, miR-106a and combinations thereof.

The invention also encompasses methods of identifying an inhibitor of tumorigenesis, comprising providing a test agent to a cell and measuring the level of at least one miR gene product in the cell. In one embodiment, the method comprises providing a test agent to a cell and measuring the level of at least one miR gene product associated with decreased expression levels in solid cancers (e.g., prostate cancer, stomach cancer, pancreatic cancer, lung cancer, breast cancer, colon cancer). An increase in the level of the miR gene product in the cell, relative to a suitable control cell, is indicative of the test agent being an inhibitor of tumorigenesis. In a particular embodiment, the at least one miR gene product associated with decreased expression levels in solid cancer cells is selected from the group consisting of miR-145, miR-155, miR-218-2 and combinations thereof.

In other embodiments, the method comprises providing a test agent to a cell and measuring the level of at least one miR gene product associated with increased expression levels in solid cancers. A decrease in the level of the miR gene product in the cell, relative to a suitable control cell, is indicative of the test agent being an inhibitor of tumorigenesis. In a particular embodiment, the at least one miR gene product associated with increased expression levels in solid cancer cells is selected from the group consisting of miR-21, miR-17-5p, miR-191, miR-29b-2, miR-223, miR-

128b, miR-199a-1, miR-24-1, miR-24-2, miR-146, miR-155, miR-181b-1, miR-20a, miR-107, miR-32, miR-92-2, miR-214, miR-30c, miR-25, miR-221, miR-106a and combinations thereof.

According to a further embodiment, there is provided a method of diagnosing whether a subject has, or is at risk for developing, a solid cancer, comprising: measuring in a test sample from the subject a level of miR-21 gene product and miR-191 gene product, comparing the level of miR-21 gene product and miR-191 gene product in the test sample to a control level of miR-21 gene product and miR-191 gene product; and diagnosing whether a subject has, or is at risk for developing, a solid cancer selected from the group consisting of colon, lung, pancreas, prostate and stomach cancer, wherein an increase in the level of miR-21 gene product and miR-191 gene product in the test sample, relative to the control level of a corresponding miR-21 gene product and miR-191 gene product is indicative of the subject either having, or being at risk for developing, a solid cancer selected from the group consisting of: colon, lung, pancreatic, prostate and stomach cancer.

According to a still further embodiment, there is provided a method of diagnosing whether a subject has, or is at risk for developing, a solid cancer, comprising: a. reverse transcribing miR-21 RNA and miR-191 RNA from a test sample obtained from the subject to provide miR-21 and miR-191 oligodeoxynucleotides; b. hybridizing the miR-21 and miR-191 oligodeoxynucleotides to a microarray comprising miRNA-specific probe oligonucleotides, which includes miR-21 and miR-191 specific probe oligonucleotides, to provide a hybridization profile for the test sample; and c. comparing the test sample hybridization profile to a control hybridization profile, wherein, if a signal of miR-21 RNA and miR-191 RNA in the test sample hybridization profile is greater than the signal of miR-21 RNA and miR-191 RNA in the control hybridization profile, than the subject either has, or is at risk for developing, a solid cancer selected from the group consisting of colon, lung, pancreas, prostate and stomach cancer.

BRIEF DESCRIPTION OF THE DRAWINGS

The patent or application file contains at least one drawing executed in color. Copies of this patent or patent application publication with color drawings will be provided by the Office upon request and payment of the necessary fee.

FIG. 1 depicts a clustering analysis of 540 samples, representing 6 solid cancers (top)

and the respective normal tissues. miRNAs included in the tree (n=137) represent those whose expression level (background-subtracted intensity) was higher than the threshold value (256) in at least 50% of the samples analyzed. Arrays were median-centered and normalized using Gene Cluster 2.0. Average linkage clustering was performed by using uncentered correlation metric. The shades indicate the difference in expression level from the median for the microRNAs in each sample.

FIG. 2 depicts unsupervised analysis of microRNA expression data. MicroRNA profiling of 540 samples (indicated at top of panel) covering breast, colon, lung, pancreas, prostate and stomach (normal tissues and tumors) were filtered, centered and normalized for each feature. The data were subject to hierarchical clustering on both the samples (horizontally-oriented) and the features (vertically-oriented) with average linkage and Pearson correlation as similarity measure. Sample names are indicated at the top of the figure and miRNA names on the left. The probe ID is indicated in parentheses, as the same microRNA can be measured by different oligonucleotides.

FIG. 3 depicts the expression of differentially-regulated miRNAS across solid cancers (top). Sixty-one microRNAs, which are present in at least 90% of the tissues solid cancers, are represented (right of panel). The tree displays the average absolute expression values for each of the listed microRNAs after $\log_2$ transformation. The mean was computed over all samples from the same tissue or tumor histotype. Genes were mean-centered and normalized using Gene Cluster 2.0. Average linkage clustering was performed using Euclidean distance.

FIG. 4 depicts fold changes in the expression of miRNAs present in at least 75% of the solid tumors with at least 1 tumor absolute value higher than 2 in different cancer samples (top), relative to normal samples. The tree displays the $\log_2$ transformation of average fold changes (cancer vs. normal). The mean was computed over all samples from the same tissue or tumor histotype. Arrays were mean-centered and normalized using Gene Cluster 2.0. Average linkage clustering was performed using uncentered correlation metric.

FIG. 5 depicts fold changes in the expression of miRNAs present in the signatures of at least 50% of the solid tumors in cancer vs. normal samples. The tree displays the $\log_2$ transformation of the average fold changes (cancer over normal). The mean was computed over all samples from the same tissue or tumor histotype. Arrays were mean centered and normalized using Gene Cluster 2.0. Average linkage clustering was performed using uncentered correlation metric.

FIG. 6A depicts bar graphs indicating that the 3'UTR of different genes encoding cancer protein enables cancer regulation by microRNA. The relative repression of firefly luciferase expression (Fold Change) standardized to a renilla luciferase control. PLAG1, pleiomorphic adenoma gene 1; TGFBR2, transforming growth factor beta receptor II; Rb, retinoblastoma gene. pGL-3 (Promega) was used as the empty vector. *miR-20a*, *miR-26a-1* and *miR-106* oligoRNAs (sense and scrambled) were used for transfections. A second experiment using mutated versions of each target mRNA, which lack the 5' miRNA-end complementarity site (MUT), as controls is shown in the bottom panel. All the experiments were performed twice in triplicate (n=6).

FIG. 6B depicts Western blots indicating that, in certain cancers (e.g., lung, breast, colon, gastric), the levels of RB1 (Rb) protein displays an inverse correlation with the level of *miR-106a* expression. β-Actin was used as a control for normalization. N1, normal sample; T1 and T2, tumor sample.

FIG. 7 depicts Northern blots showing down-regulation of miR-145 (top) and up-regulation of miR-21 (bottom) expression in breast cancer samples (P series and numbered series) relative to normal samples. Normalization was performed with a U6-specific probe.

FIG. 8 depicts Northern blots showing up-regulation of miR-103 and down-regulation miR-155 (top) expression in different endocrine pancreatic cancer samples (WDET, well differentiated pancreatic endocrine tumors, WDEC, well differentiated pancreatic endocrine carcinomas and ACC, pancreatic acinar cell carcinomas) relative to normal samples (K series), as well as up-regulation of miR-204 (bottom) expression in insulinomas (F series) relative to normal samples (K series) and non secreting/non functioning (NF-series) samples. Normalization was performed with a probe specific to 5S RNA.

DETAILED DESCRIPTION OF THE INVENTION

The present invention is based, in part, on the identification of particular microRNAs whose expression is altered in cancer cells associated with different solid cancers, such as colon, stomach, pancreatic, lung, breast and prostate cancer, relative to normal control cells.

As used herein interchangeably, a “miR gene product,” "microRNA," "miR," or “miRNA” refers to the unprocessed (e.g., precursor) or processed (e.g., mature) RNA transcript from a miR gene. As the miR gene products are not translated into protein, the term “miR gene products” does not include proteins. The unprocessed miR gene transcript is also called a “miR precursor” or "miR prec" and typically comprises an RNA transcript of about 70-100 nucleotides in length. The miR precursor can be processed by digestion with an RNAse (for example, Dicer, Argonaut, or RNAse III (e.g., *E. coli* RNAse III)) into an active 19-25 nucleotide RNA molecule. This active 19-25 nucleotide RNA molecule is also called the “processed" miR gene transcript or "mature" miRNA.

The active 19-25 nucleotide RNA molecule can be obtained from the miR precursor through natural processing routes (*e.g.*, using intact cells or cell lysates) or by synthetic processing routes (*e.g.*, using isolated processing enzymes, such as isolated Dicer, Argonaut, or RNAse III). It is understood that the active 19-25 nucleotide RNA molecule can also be produced directly by biological or chemical synthesis, without having been processed from the miR precursor. When a microRNA is referred to herein by name, the name corresponds to both the precursor and mature forms, unless otherwise indicated.

Tables 1a and 1b depict the nucleotide sequences of particular precursor and mature human microRNAs.

Table 1a- Human microRNA Precursor Sequences

Precursor Name	Sequence (5' To 3')*	SEQ ID NO.
let-7a-1	CACUGUGGGAUGAGGUAGUAGGUUGUAUAGUUU UAGGGUCACACCCACCACUGGGAGAUAACUAUA CAAUCUACUGUCUUUCCUAACGUG	1
let-7a-2	AGGUUGAGGUAGUAGGUUGUAUAGUUUAGAAUU ACAUCAAGGGAGAUAACUGUACAGCCUCCUAGC UUUCCU	2
let-7a-3	GGGUGAGGUAGUAGGUUGUAUAGUUUGGGGCUC UGCCCUGCUAUGGGAUAACUAUACAAUCUACUG UCUUUCCU	3
let-7a-4	GUGACUGCAUGCUCCCAGGUUGAGGUAGUAGGU UGUAUAGUUUAGAAUUACACAAGGGAGAUAACU GUACAGCCUCCUAGCUUUCCUUGGGUCUUGCAC UAAACAAC	4
let-7b	GGCGGGGUGAGGUAGUAGGUUGUGUGGUUUCAG GGCAGUGAUGUUGCCCCUCGGAAGAUAACUAUA CAACCUACUGCCUUCCCUG	5
let-7c	GCAUCCGGGUUGAGGUAGUAGGUUGUAUGGUUU AGAGUUACACCCUGGGAGUUAACUGUACAACCU UCUAGCUUUCCUUGGAGC	6
let-7d	CCUAGGAAGAGGUAGUAGGUUGCAUAGUUUUAG GGCAGGGAUUUUGCCCACAAGGAGGUAACUAUA CGACCUGCUGCCUUUCUUAGG	7
let-7d-v1	CUAGGAAGAGGUAGUAGUUUGCAUAGUUUUAGG GCAAAGAUUUUGCCCACAAGUAGUUAGCUAUAC GACCUGCAGCCUUUUGUAG	8
let-7d-v2	CUGGCUGAGGUAGUAGUUUGUGCUGUUGGUCGG GUUGUGACAUUGCCCGCUGUGGAGAUAACUGCG CAAGCUACUGCCUUGCUAG	9
let-7e	CCCGGGCUGAGGUAGGAGGUUGUAUAGUUGAGG AGGACACCCAAGGAGAUCACUAUACGGCCUCCU AGCUUUCCCCAGG	10
let-7f-1	UCAGAGUGAGGUAGUAGAUUGUAUAGUUGUGGG GUAGUGAUUUUACCCUGUUCAGGAGAUAACUAU ACAAUCUAUUGCCUUCCCUGA	11
let-7f-2-1	CUGUGGGAUGAGGUAGUAGAUUGUAUAGUUGUG GGGUAGUGAUUUUACCCUGUUCAGGAGAUAACU AUACAAUCUAUUGCCUUCCCUGA	12

Precursor Name	Sequence (5' To 3')*	SEQ ID NO.
let-7f-2-2	CUGUGGGAUGAGGUAGUAGAUUGUAUAGUUUUA GGGUCAUACCCCAUCUUGGAGAUAACUAUACAG UCUACUGUCUUUCCCACGG	13
let-7g	UUGCCUGAUUCCAGGCUGAGGUAGUAGUUUGUA CAGUUUGAGGGUCUAUGAUACCACCCGGUACAG GAGAUAACUGUACAGGCCACUGCCUUGCCAGGA ACAGCGCGC	14
let-7i	CUGGCUGAGGUAGUAGUUUGUGCUGUUGGUCGG GUUGUGACAUUGCCCGCUGUGGAGAUAACUGCG CAAGCUACUGCCUUGCUAG	15
miR-1b-1-1	ACCUACUCAGAGUACAUACUUCUUUAUGUACCC AUAUGAACAUACAAUGCUAUGGAAUGUAAAGAA GUAUGUAUUUUUGGUAGGC	16
miR-1b-1-2	CAGCUAACAACUUAGUAAUACCUACUCAGAGUA CAUACUUCUUUAUGUACCCAUAUGAACAUACAA UGCUAUGGAAUGUAAAGAAGUAUGUAUUUUUGG UAGGCAAUA	17
miR-1b-2	GCCUGCUUGGGAAACAUACUUCUUUAUAUGCCC AUAUGGACCUGCUAAGCUAUGGAAUGUAAAGAA GUAUGUAUCUCAGGCCGGG	18
miR-1b	UGGGAAACAUACUUCUUUAUAUGCCCAUAUGGA CCUGCUAAGCUAUGGAAUGUAAAGAAGUAUGUA UCUCA	19
miR-1d	ACCUACUCAGAGUACAUACUUCUUUAUGUACCC AUAUGAACAUACAAUGCUAUGGAAUGUAAAGAA GUAUGUAUUUUUGGUAGGC	20
miR-7-1a	UGGAUGUUGGCCUAGUUCUGUGUGGAAGACUAG UGAUUUUGUUGUUUUUAGAUAACUAAAUCGACA ACAAAUCACAGUCUGCCAUAUGGCACAGGCCAU GCCUCUACA	21
miR-7-1b	UUGGAUGUUGGCCUAGUUCUGUGUGGAAGACUA GUGAUUUUGUUGUUUUUAGAUAACUAAAUCGAC AACAAAUCACAGUCUGCCAUAUGGCACAGGCCA UGCCUCUACAG	22
miR-7-2	CUGGAUACAGAGUGGACCGGCUGGCCCCAUCUG GAAGACUAGUGAUUUUGUUGUUGUCUUACUGCG CUCAACAACAAAUCCCAGUCUACCUAAUGGUGC CAGCCAUCGCA	23
miR-7-3	AGAUUAGAGUGGCUGUGGUCUAGUGCUGUGUGG AAGACUAGUGAUUUUGUUGUUCUGAUGUACUAC GACAACAAGUCACAGCCGGCCUCAUAGCGCAGA CUCCCUUCGAC	24
miR-9-1	CGGGGUUGGUUGUUAUCUUUGGUUAUCUAGCUG UAUGAGUGGUGUGGAGUCUUCAUAAAGCUAGAU AACCGAAAGUAAAAAUAACCCCA	25

Precursor Name	Sequence (5' To 3')*	SEQ ID NO.
miR-9-2	GGAAGCGAGUUGUUAUCUUUGGUUAUCUAGCUG UAUGAGUGUAUUGGUCUUCAUAAAGCUAGAUAA CCGAAAGUAAAAACUCCUUCA	26
miR-9-3	GGAGGCCCGUUUCUCUCUUUGGUUAUCUAGCUG UAUGAGUGCCACAGAGCCGUCAUAAAGCUAGAU AACCGAAAGUAGAAAUGAUUCUCA	27
miR-10a	GAUCUGUCUGUCUUCUGUAUAUACCCUGUAGAU CCGAAUUUGUGUAAGGAAUUUUGUGGUCACAAA UUCGUAUCUAGGGGAAUAUGUAGUUGACAUAAA CACUCCGCUCU	28
miR-10b	CCAGAGGUUGUAACGUUGUCUAUAUAUACCCUG UAGAACCGAAUUUGUGUGGUAUCCGUAUAGUCA CAGAUUCGAUUCUAGGGGAAUAUAUGGUCGAUG CAAAAACUUCA	29
miR-15a-2	GCGCGAAUGUGUGUUUAAAAAAAAUAAAACCUU GGAGUAAAGUAGCAGCACAUAAUGGUUUGUGGA UUUUGAAAAGGUGCAGGCCAUAUUGUGCUGCCU CAAAAAUAC	30
miR-15a	CCUUGGAGUAAAGUAGCAGCACAUAAUGGUUUG UGGAUUUUGAAAAGGUGCAGGCCAUAUUGUGCU GCCUCAAAAAUACAAGG	31
miR-15b-1	CUGUAGCAGCACAUCAUGGUUUACAUGCUACAG UCAAGAUGCGAAUCAUUAUUUGCUGCUCUAG	32
miR-15b-2	UUGAGGCCUUAAAGUACUGUAGCAGCACAUCAU GGUUUACAUGCUACAGUCAAGAUGCGAAUCAUU AUUUGCUGCUCUAGAAAUUUAAGGAAAUUCAU	33
miR-16-1	GUCAGCAGUGCCUUAGCAGCACGUAAAUAUUGG CGUUAAGAUUCUAAAAUUAUCUCCAGUAUUAAC UGUGCUGCUGAAGUAAGGUUGAC	34
miR-16-2	GUUCCACUCUAGCAGCACGUAAAUAUUGGCGUA GUGAAAUAUAUAUUAAACACCAAUAUUACUGUG CUGCUUUAGUGUGAC	35
miR-16-13	GCAGUGCCUUAGCAGCACGUAAAUAUUGGCGUU AAGAUUCUAAAAUUAUCUCCAGUAUUAACUGUG CUGCUGAAGUAAGGU	36
miR-17	GUCAGAAUAAUGUCAAAGUGCUUACAGUGCAGG UAGUGAUAUGUGCAUCUACUGCAGUGAAGGCAC UUGUAGCAUUAUGGUGAC	37
miR-18	UGUUCUAAGGUGCAUCUAGUGCAGAUAGUGAAG UAGAUUAGCAUCUACUGCCCUAAGUGCUCCUUC UGGCA	38
miR-18-13	UUUUUGUUCUAAGGUGCAUCUAGUGCAGAUAGU GAAGUAGAUUAGCAUCUACUGCCCUAAGUGCUC CUUCUGGCAUAAGAA	39

Precursor Name	Sequence (5' To 3')*	SEQ ID NO.
miR-19a	GCAGUCCUCUGUUAGUUUUGCAUAGUUGCACUA CAAGAAGAAUGUAGUUGUGCAAAUCUAUGCAAA ACUGAUGGUGGCCUGC	40
miR-19a-13	CAGUCCUCUGUUAGUUUUGCAUAGUUGCACUAC AAGAAGAAUGUAGUUGUGCAAAUCUAUGCAAAA CUGAUGGUGGCCUG	41
miR-19b-1	CACUGUUCUAUGGUUAGUUUUGCAGGUUUGCAU CCAGCUGUGUGAUAUUCUGCUGUGCAAAUCCAU GCAAAACUGACUGUGGUAGUG	42
miR-19b-2	ACAUUGCUACUUACAAUUAGUUUUGCAGGUUUG CAUUUCAGCGUAUAUAUGUAUAUGUGGCUGUGC AAAUCCAUGCAAAACUGAUUGUGAUAAUGU	43
miR-19b-13	UUCUAUGGUUAGUUUUGCAGGUUUGCAUCCAGC UGUGUGAUAUUCUGCUGUGCAAAUCCAUGCAAA ACUGACUGUGGUAG	44
miR-19b-X	UUACAAUUAGUUUUGCAGGUUUGCAUUUCAGCG UAUAUAUGUAUAUGUGGCUGUGCAAAUCCAUGC AAAACUGAUUGUGAU	45
miR-20 (miR-20a)	GUAGCACUAAAGUGCUUAUAGUGCAGGUAGUGU UUAGUUAUCUACUGCAUUAUGAGCACUUAAAGU ACUGC	46
miR-21	UGUCGGGUAGCUUAUCAGACUGAUGUUGACUGU UGAAUCUCAUGGCAACACCAGUCGAUGGGCUGU CUGACA	47
miR-21-17	ACCUUGUCGGGUAGCUUAUCAGACUGAUGUUGA CUGUUGAAUCUCAUGGCAACACCAGUCGAUGGG CUGUCUGACAUUUUG	48
miR-22	GGCUGAGCCGCAGUAGUUCUUCAGUGGCAAGCU UUAUGUCCUGACCCAGCUAAAGCUGCCAGUUGA AGAACUGUUGCCCUCUGCC	49
miR-23a	GGCCGGCUGGGGUUCCUGGGGAUGGGAUUUGCU UCCUGUCACAAAUCACAUUGCCAGGGAUUUCCA ACCGACC	50
miR-23b	CUCAGGUGCUCUGGCUGCUUGGGUUCCUGGCAU GCUGAUUUGUGACUUAAGAUUAAAAUCACAUUG CCAGGGAUUACCACGCAACCACGACCUUGGC	51
miR-23-19	CCACGGCCGGCUGGGGUUCCUGGGGAUGGGAUU UGCUUCCUGUCACAAAUCACAUUGCCAGGGAUU UCCAACCGACCCUGA	52
miR-24-1	CUCCGGUGCCUACUGAGCUGAUAUCAGUUCUCA UUUUACACACUGGCUCAGUUCAGCAGGAACAGG AG	53
miR-24-2	CUCUGCCUCCCGUGCCUACUGAGCUGAAACACAG UUGGUUUGUGUACACUGGCUCAGUUCAGCAGGA ACAGGG	54

Precursor Name	Sequence (5' To 3')*	SEQ ID NO.
miR-24-19	CCCUGGGCUCUGCCUCCCGUGCCUACUGAGCUGAAACACAGUUGGUUUGUGUACACUGGCUCAGUUCAGCAGGAACAGGGG	55
miR-24-9	CCCUCCGGUGCCUACUGAGCUGAUAUCAGUUCUCAUUUUACACACUGGCUCAGUUCAGCAGGAACAGCAUC	56
miR-25	GGCCAGUGUUGAGAGGCGGAGACUUGGGCAAUUGCUGGACGCUGCCCUGGGCAUUGCACUUGUCUCGGUCUGACAGUGCCGGCC	57
miR-26a	AGGCCGUGGCCUCGUUCAAGUAAUCCAGGAUAGGCUGUGCAGGUCCCAAUGGCCUAUCUUGGUUACUUGCACGGGGACGCGGGCCU	58
miR-26a-1	GUGGCCUCGUUCAAGUAAUCCAGGAUAGGCUGUGCAGGUCCCAAUGGGCCUAUUCUUGGUUACUUGCACGGGGACGC	59
miR-26a-2	GGCUGUGGCUGGAUUCAAGUAAUCCAGGAUAGGCUGUUUCCAUCUGUGAGGCCUAUUCUUGAUUACUUGUUUCUGGAGGCAGCU	60
miR-26b	CCGGGACCCAGUUCAAGUAAUUCAGGAUAGGUUGUGUGCUGUCCAGCCUGUUCUCCAUUACUUGGCUCGGGGACCGG	61
miR-27a	CUGAGGAGCAGGGCUUAGCUGCUUGUGAGCAGGGUCCACACCAAGUCGUGUUCACAGUGGCUAAGUUCCGCCCCCCAG	62
miR-27b-1	AGGUGCAGAGCUUAGCUGAUUGGUGAACAGUGAUUGGUUUCCGCUUUGUUCACAGUGGCUAAGUUCUGCACCU	63
miR-27b-2	ACCUCUCUAACAAGGUGCAGAGCUUAGCUGAUUGGUGAACAGUGAUUGGUUUCCGCUUUGUUCACAGUGGCUAAGUUCUGCACCUGAAGAGAAGGUG	64
miR-27-19	CCUGAGGAGCAGGGCUUAGCUGCUUGUGAGCAGGGUCCACACCAAGUCGUGUUCACAGUGGCUAAGUUCCGCCCCCCAGG	65
miR-28	GGUCCUUGCCCUCAAGGAGCUCACAGUCUAUUGAGUUACCUUUCUGACUUUCCCACUAGAUUGUGAGCUCCUGGAGGGCAGGCACU	66
miR-29a-2	CCUUCUGUGACCCCUUAGAGGAUGACUGAUUUCUUUUGGUGUUCAGAGUCAAUAUAAUUUUCUAGCACCAUCUGAAAUCGGUUAUAAUGAUUGGGGAAGAGCACCAUG	67
miR-29a	AUGACUGAUUUCUUUUGGUGUUCAGAGUCAAUAUAAUUUUCUAGCACCAUCUGAAAUCGGUUAU	68
miR-29b-1	CUUCAGGAAGCUGGUUUCAUAUGGUGGUUUAGAUUUAAAUAGUGAUUGUCUAGCACCAUUUGAAAUCAGUGUUCUUGGGGG	69

Precursor Name	Sequence (5' To 3')*	SEQ ID NO.
miR-29b-2	CUUCUGGAAGCUGGUUUCACAUGGUGGCUUAGA UUUUUCCAUCUUUGUAUCUAGCACCAUUUGAAA UCAGUGUUUUAGGAG	70
miR-29c	ACCACUGGCCCAUCUCUUACACAGGCUGACCGAU UUCUCCUGGUGUUCAGAGUCUGUUUUUGUCUAG CACCAUUUGAAAUCGGUUAUGAUGUAGGGGGAA AAGCAGCAGC	71
miR-30a	GCGACUGUAAACAUCCUCGACUGGAAGCUGUGA AGCCACAGAUGGGCUUUCAGUCGGAUGUUUGCA GCUGC	72
miR-30b-1	AUGUAAACAUCCUACACUCAGCUGUAAUACAUG GAUUGGCUGGGAGGUGGAUGUUUACGU	73
miR-30b-2	ACCAAGUUUCAGUUCAUGUAAACAUCCUACACU CAGCUGUAAUACAUGGAUUGGCUGGGAGGUGGA UGUUUACUUCAGCUGACUUGGA	74
miR-30c	AGAUACUGUAAACAUCCUACACUCUCAGCUGUG GAAAGUAAGAAAGCUGGGAGAAGGCUGUUUACU CUUUCU	75
miR-30d	GUUGUUGUAAACAUCCCCGACUGGAAGCUGUAA GACACAGCUAAGCUUUCAGUCAGAUGUUUGCUG CUAC	76
miR-30e	CUGUAAACAUCCUUGACUGGAAGCUGUAAGGUG UUCAGAGGAGCUUUCAGUCGGAUGUUUACAG	77
miR-31	GGAGAGGAGGCAAGAUGCUGGCAUAGCUGUUGA ACUGGGAACCUGCUAUGCCAACAUAUUGCCAUC UUUCC	78
miR-32	GGAGAUAUUGCACAUUACUAAGUUGCAUGUUGU CACGGCCUCAAUGCAAUUUAGUGUGUGUGAUAU UUUC	79
miR-33b	GGGGGCCGAGAGAGGCGGGCGGCCCCGCGGUGC AUUGCUGUUGCAUUGCACGUGUGUGAGGCGGGU GCAGUGCCUCGGCAGUGCAGCCCGGAGCCGGCCC CUGGCACCAC	80
miR-33b-2	ACCAAGUUUCAGUUCAUGUAAACAUCCUACACU CAGCUGUAAUACAUGGAUUGGCUGGGAGGUGGA UGUUUACUUCAGCUGACUUGGA	81
miR-33	CUGUGGUGCAUUGUAGUUGCAUUGCAUGUUCUG GUGGUACCCAUGCAAUGUUUCCACAGUGCAUCA CAG	82
miR-34-a	GGCCAGCUGUGAGUGUUUCUUUGGCAGUGUCUU AGCUGGUUGUUGUGAGCAAUAGUAAGGAAGCAA UCAGCAAGUAUACUGCCCUAGAAGUGCUGCACG UUGUGGGGCCC	83

Precursor Name	Sequence (5' To 3')*	SEQ ID NO.
miR-34-b	GUGCUCGGUUUGUAGGCAGUGUCAUUAGCUGAUUGUACUGUGGUGGUUACAAUCACUAACUCCACUGCCAUCAAAACAAGGCAC	84
miR-34-c	AGUCUAGUUACUAGGCAGUGUAGUUAGCUGAUUGCUAAUAGUACCAAUCACUAACCACACGGCCAGGUAAAAAGAUU	85
miR-91-13	UCAGAAUAAUGUCAAAGUGCUUACAGUGCAGGUAGUGAUAUGUGCAUCUACUGCAGUGAAGGCACUUGUAGCAUUAUGGUGA	86
miR-92-1	CUUUCUACACAGGUUGGGAUCGGUUGCAAUGCUGUGUUUCUGUAUGGUAUUGCACUUGUCCCGGCCUGUUGAGUUUGG	87
miR-92 -2	UCAUCCCUGGGUGGGGAUUUGUUGCAUUACUUGUGUUCUAUAUAAAGUAUUGCACUUGUCCCGGCCUGUGGAAGA	88
miR-93-1 (miR-93-2)	CUGGGGGCUCCAAAGUGCUGUUCGUGCAGGUAGUGUGAUUACCCAACCUACUGCUGAGCUAGCACUUCCCGAGCCCCCGG	89
miR-95-4	AACACAGUGGGCACUCAAUAAAUGUCUGUUGAAUUGAAAUGCGUUACAUUCAACGGGUAUUUAUUGAGCACCCACUCUGUG	90
miR-96-7	UGGCCGAUUUUGGCACUAGCACAUUUUUGCUUGUGUCUCUCCGCUCUGAGCAAUCAUGUGCAGUGCCAAUAUGGGAAA	91
miR-97-6 (miR-30)*	GUGAGCGACUGUAAACAUCCUCGACUGGAAGCUGUGAAGCCACAGAUGGGCUUUCAGUCGGAUGUUUGCAGCUGCCUACU	92
miR-98	GUGAGGUAGUAAGUUGUAUUGUUGUGGGGUAGGGAUAUUAGGCCCCAAUUAGAAGAUAACUAUACAACUUACUACUUUCC	93
miR-99b	GGCACCCACCCGUAGAACCGACCUUGCGGGGCCUUCGCCGCACACAAGCUCGUGUCUGUGGGUCCGUGUC	94
miR-99a	CCCAUUGGCAUAAACCCGUAGAUCCGAUCUUGUGGUGAAGUGGACCGCACAAGCUCGCUUCUAUGGGUCUGUGUCAGUGUG	95
miR-100-1/2	AAGAGAGAAGAUAUUGAGGCCUGUUGCCACAAACCCGUAGAUCCGAACUUGUGGUAUUAGUCCGCACAAGCUUGUAUCUAUAGGUAUGUGUCUGUUAGGCAAUCUCAC	96
miR-100-11	CCUGUUGCCACAAACCCGUAGAUCCGAACUUGUGGUAUUAGUCCGCACAAGCUUGUAUCUAUAGGUAUGUGUCUGUUAGG	97

Precursor Name	Sequence (5' To 3')*	SEQ ID NO.
miR-101-1 /2	AGGCUGCCCUGGCUCAGUUAUCACAGUGCUGAUGCUGUCUAUUCUAAAGGUACAGUACUGUGAUAACUGAAGGAUGGCAGCCAUCUUACCUUCCAUCAGAGGAGCCUCAC	98
miR-101	UCAGUUAUCACAGUGCUGAUGCUGUCCAUUCUAAAGGUACAGUACUGUGAUAACUGA	99
miR-101-1	UGCCCUGGCUCAGUUAUCACAGUGCUGAUGCUGUCUAUUCUAAAGGUACAGUACUGUGAUAACUGAAGGAUGGCA	100
miR-101-2	ACUGUCCUUUUUCGGUUAUCAUGGUACCGAUGCUGUAUAUCUGAAAGGUACAGUACUGUGAUAACUGAAGAAUGGUGGU	101
miR-101-9	UGUCCUUUUUCGGUUAUCAUGGUACCGAUGCUGUAUAUCUGAAAGGUACAGUACUGUGAUAACUGAAGAAUGGUG	102
miR-102-1	CUUCUGGAAGCUGGUUUCACAUGGUGGCUUAGAUUUUUCCAUCUUUGUAUCUAGCACCAUUUGAAAUCAGUGUUUUAGGAG	103
miR-102-7.1 (miR-102-7.2)	CUUCAGGAAGCUGGUUUCAUAUGGUGGUUUAGAUUUAAAUAGUGAUUGUCUAGCACCAUUUGAAAUCAGUGUUCUUGGGGG	104
miR-103-2	UUGUGCUUUCAGCUUCUUUACAGUGCUGCCUUGUAGCAUUCAGGUCAAGCAACAUUGUACAGGGCUAUGAAAGAACCA	105
miR-103-1	UACUGCCCUCGGCUUCUUUACAGUGCUGCCUUGUUGCAUAUGGAUCAAGCAGCAUUGUACAGGGCUAUGAAGGCAUUG	106
miR-104-17	AAAUGUCAGACAGCCCAUCGACUGGUGUUGCCAUGAGAUUCAACAGUCAACAUCAGUCUGAUAAGCUACCCGACAAGG	107
miR-105-1	UGUGCAUCGUGGUCAAAUGCUCAGACUCCUGUGGUGGCUGCUCAUGCACCACGGAUGUUUGAGCAUGUGCUACGGUGUCUA	108
miR-105-2	UGUGCAUCGUGGUCAAAUGCUCAGACUCCUGUGGUGGCUGCUUAUGCACCACGGAUGUUUGAGCAUGUGCUAUGGUGUCUA	109
miR-106-a	CCUUGGCCAUGUAAAAGUGCUUACAGUGCAGGUAGCUUUUUGAGAUCUACUGCAAUGUAAGCACUUCUUACAUUACCAUGG	110
miR-106-b	CCUGCCGGGGCUAAAGUGCUGACAGUGCAGAUAGUGGUCCUCUCCGUGCUACCGCACUGUGGGUACUUGCUGCUCCAGCAGG	111
miR-107	CUCUCUGCUUUCAGCUUCUUUACAGUGUUGCCUUGUGGCAUGGAGUUCAAGCAGCAUUGUACAGGGCUAUCAAAGCACAGA	112

Precursor Name	Sequence (5' To 3')*	SEQ ID NO.
miR-108-1-small	ACACUGCAAGAACAAUAAGGAUUUUUAGGGGCAUUAUGACUGAGUCAGAAAACACAGCUGCCCCUGAAAGUCCCUCAUUUUUCUUGCUGU	113
miR-108-2-small	ACUGCAAGAGCAAUAAGGAUUUUUAGGGGCAUUAUGAUAGUGGAAUGGAAACACAUCUGCCCCCAAAAGUCCCUCAUUUU	114
miR-122a-1	CCUUAGCAGAGCUGUGGAGUGUGACAAUGGUGUUUGUGUCUAAACUAUCAAACGCCAUUAUCACACUAAAUAGCUACUGCUAGGC	115
miR-122a-2	AGCUGUGGAGUGUGACAAUGGUGUUUGUGUCCAAACUAUCAAACGCCAUUAUCACACUAAAUAGCU	116
miR-123	ACAUUAUUACUUUUGGUACGCGCUGUGACACUUCAAACUCGUACCGUGAGUAAUAAUGCGC	117
miR-124a-1	AGGCCUCUCUCUCCGUGUUCACAGCGGACCUUGAUUUAAAUGUCCAUACAAUUAAGGCACGCGGUGAAUGCCAAGAAUGGGGCUG	118
miR-124a-2	AUCAAGAUUAGAGGCUCUGCUCUCCGUGUUCACAGCGGACCUUGAUUUAAUGUCAUACAAUUAAGGCACGCGGUGAAUGCCAAGAGCGGAGCCUACGGCUGCACUUGAAG	119
miR-124a-3	UGAGGGCCCCUCUGCGUGUUCACAGCGGACCUUGAUUUAAUGUCUAUACAAUUAAGGCACGCGGUGAAUGCCAAGAGAGGCGCCUCC	120
miR-124a	CUCUGCGUGUUCACAGCGGACCUUGAUUUAAUGUCUAUACAAUUAAGGCACGCGGUGAAUGCCAAGAG	121
miR-124b	CUCUCCGUGUUCACAGCGGACCUUGAUUUAAUGUCAUACAAUUAAGGCACGCGGUGAAUGCCAAGAG	122
miR-125a-1	UGCCAGUCUCUAGGUCCCUGAGACCCUUUAACCUGUGAGGACAUCCAGGGUCACAGGUGAGGUUCUUGGGAGCCUGGCGUCUGGCC	123
miR-125a-2	GGUCCCUGAGACCCUUUAACCUGUGAGGACAUCCAGGGUCACAGGUGAGGUUCUUGGGAGCCUGG	124
miR-125b-1	UGCGCUCCUCUCAGUCCCUGAGACCCUAACUUGUGAUGUUUACCGUUUAAAUCCACGGGUUAGGCUCUUGGGAGCUGCGAGUCGUGCU	125
miR-125b-2	ACCAGACUUUUCCUAGUCCCUGAGACCCUAACUUGUGAGGUAUUUUAGUAACAUCACAAGUCAGGCUCUUGGGACCUAGGCGGAGGGGA	126
miR-126-1	CGCUGGCGACGGGACAUUAUUACUUUUGGUACGCGCUGUGACACUUCAAACUCGUACCGUGAGUAAUAAUGCGCCGUCCACGGCA	127
miR-126-2	ACAUUAUUACUUUUGGUACGCGCUGUGACACUUCAAACUCGUACCGUGAGUAAUAAUGCGC	128

Precursor Name	Sequence (5' To 3')*	SEQ ID NO.
miR-127-1	UGUGAUCACUGUCUCCAGCCUGCUGAAGCUCAG AGGGCUCUGAUUCAGAAAGAUCAUCGGAUCCGU CUGAGCUUGGCUGGUCGGAAGUCUCAUCAUC	129
miR-127-2	CCAGCCUGCUGAAGCUCAGAGGGCUCUGAUUCA GAAAGAUCAUCGGAUCCGUCUGAGCUUGGCUGG UCGG	130
miR-128a	UGAGCUGUUGGAUUCGGGGCCGUAGCACUGUCU GAGAGGUUUACAUUUCUCACAGUGAACCGGUCU CUUUUUCAGCUGCUUC	131
miR-128b	GCCCGGCAGCCACUGUGCAGUGGGAAGGGGGGC CGAUACACUGUACGAGAGUGAGUAGCAGGUCUC ACAGUGAACCGGUCUCUUUCCCUACUGUGUCAC ACUCCUAAUGG	132
miR-128	GUUGGAUUCGGGGCCGUAGCACUGUCUGAGAGG UUUACAUUUCUCACAGUGAACCGGUCUCUUUUU CAGC	133
miR-129-1	UGGAUCUUUUUGCGGUCUGGGCUUGCUGUUCCU CUCAACAGUAGUCAGGAAGCCCUUACCCCAAAA AGUAUCUA	134
miR-129-2	UGCCCUUCGCGAAUCUUUUUGCGGUCUGGGCUU GCUGUACAUAACUCAAUAGCCGGAAGCCCUUAC CCCAAAAAGCAUUUGCGGAGGGCG	135
miR-130a	UGCUGCUGGCCAGAGCUCUUUUCACAUUGUGCU ACUGUCUGCACCUGUCACUAGCAGUGCAAUGUU AAAAGGGCAUUGGCCGUGUAGUG	136
miR-131-1	GCCAGGAGGCGGGGUUGGUUGUUAUCUUUGGUU AUCUAGCUGUAUGAGUGGUGUGGAGUCUUCAUA AAGCUAGAUAACCGAAAGUAAAAAUAACCCCAU ACACUGCGCAG	137
miR-131-3	CACGGCGCGGCAGCGGCACUGGCUAAGGGAGGC CCGUUUCUCUCUUUGGUUAUCUAGCUGUAUGAG UGCCACAGAGCCGUCAUAAAGCUAGAUAACCGA AAGUAGAAAUG	138
miR-131	GUUGUUAUCUUUGGUUAUCUAGCUGUAUGAGUG UAUUGGUCUUCAUAAAGCUAGAUAACCGAAAGU AAAAAC	139
miR-132-1	CCGCCCCCGCGUCUCCAGGGCAACCGUGGCUUUC GAUUGUUACUGUGGGAACUGGAGGUAACAGUCU ACAGCCAUGGUCGCCCCGCAGCACGCCCACGCGC	140
miR-132-2	GGGCAACCGUGGCUUUCGAUUGUUACUGUGGGA ACUGGAGGUAACAGUCUACAGCCAUGGUCGCCC	141
miR-133a-1	ACAAUGCUUUGCUAGAGCUGGUAAAAUGGAACC AAAUCGCCUCUUCAAUGGAUUUGGUCCCCUUCA ACCAGCUGUAGCUAUGCAUUGA	142

Precursor Name	Sequence (5' To 3')*	SEQ ID NO.
miR-133a-2	GGGAGCCAAAUGCUUUGCUAGAGCUGGUAAAAU GGAACCAAAUCGACUGUCCAAUGGAUUUGGUCC CCUUCAACCAGCUGUAGCUGUGCAUUGAUGGCG CCG	143
miR-133	GCUAGAGCUGGUAAAAUGGAACCAAAUCGCCUC UUCAAUGGAUUUGGUCCCCUUCAACCAGCUGUA GC	144
miR-133b	CCUCAGAAGAAAGAUGCCCCCUGCUCUGGCUGG UCAAACGGAACCAAGUCCGUCUUCCUGAGAGGU UUGGUCCCCUUCAACCAGCUACAGCAGGGCUGG CAAUGCCCAGUCCUUGGAGA	145
miR-133b-small	GCCCCCUGCUCUGGCUGGUCAAACGGAACCAAG UCCGUCUUCCUGAGAGGUUUGGUCCCCUUCAAC CAGCUACAGCAGGG	146
miR-134-1	CAGGGUGUGUGACUGGUUGACCAGAGGGGCAUG CACUGUGUUCACCCUGUGGGCCACCUAGUCACCA ACCCUC	147
miR-134-2	AGGGUGUGUGACUGGUUGACCAGAGGGGCAUGC ACUGUGUUCACCCUGUGGGCCACCUAGUCACCA ACCCU	148
miR-135a-1	AGGCCUCGCUGUUCUCUAUGGCUUUUUAUUCCU AUGUGAUUCUACUGCUCACUCAUAUAGGGAUUG GAGCCGUGGCGCACGGCGGGGACA	149
miR-135a-2 (miR-135-2)	AGAUAAAUUCACUCUAGUGCUUUAUGGCUUUUU AUUCCUAUGUGAUAGUAAUAAAGUCUCAUGUAG GGAUGGAAGCCAUGAAAUACAUUGUGAAAAAUC A	150
miR-135	CUAUGGCUUUUUAUUCCUAUGUGAUUCUACUGC UCACUCAUAUAGGGAUUGGAGCCGUGG	151
miR-135b	CACUCUGCUGUGGCCUAUGGCUUUUCAUUCCUA UGUGAUUGCUGUCCCAAACUCAUGUAGGGCUAA AAGCCAUGGGCUACAGUGAGGGGCGAGCUCC	152
miR-136-1	UGAGCCCUCGGAGGACUCCAUUUGUUUUGAUGA UGGAUUCUUAUGCUCCAUCAUCGUCUCAAAUGA GUCUUCAGAGGGUUCU	153
miR-136-2	GAGGACUCCAUUUGUUUUGAUGAUGGAUUCUUA UGCUCCAUCAUCGUCUCAAAUGAGUCUUC	154
miR-137	CUUCGGUGACGGGUAUUCUUGGGUGGAUAAUAC GGAUUACGUUGUUAUUGCUUAAGAAUACGCGUA GUCGAGG	155
miR-138-1	CCCUGGCAUGGUGUGGUGGGGCAGCUGGUGUUG UGAAUCAGGCCGUUGCCAAUCAGAGAACGGCUA CUUCACAACACCAGGGCCACACCACACUACAGG	156

Precursor Name	Sequence (5' To 3')*	SEQ ID NO.
miR-138-2	CGUUGCUGCAGCUGGUGUUGUGAAUCAGGCCGACGAGCAGCGCAUCCUCUUACCCGGCUAUUUCACGACACCAGGGUUGCAUCA	157
miR-138	CAGCUGGUGUUGUGAAUCAGGCCGACGAGCAGCGCAUCCUCUUACCCGGCUAUUUCACGACACCAGGGUUG	158
miR-139	GUGUAUUCUACAGUGCACGUGUCUCCAGUGUGGCUCGGAGGCUGGAGACGCGGCCCUGUUGGAGUAAC	159
miR-140	UGUGUCUCUCUCUGUGUCCUGCCAGUGGUUUUACCCUAUGGUAGGUUACGUCAUGCUGUUCUACCACAGGGUAGAACCACGGACAGGAUACCGGGGCACC	160
miR-140as	UCCUGCCAGUGGUUUUACCCUAUGGUAGGUUACGUCAUGCUGUUCUACCACAGGGUAGAACCACGGACAGGA	161
miR-140s	CCUGCCAGUGGUUUUACCCUAUGGUAGGUUACGUCAUGCUGUUCUACCACAGGGUAGAACCACGGACAGG	162
miR-141-1	CGGCCGGCCCUGGGUCCAUCUUCCAGUACAGUGUUGGAUGGUCUAAUUGUGAAGCUCCUAACACUGUCUGGUAAAGAUGGCUCCCGGGUGGGUUC	163
miR-141-2	GGGUCCAUCUUCCAGUACAGUGUUGGAUGGUCUAAUUGUGAAGCUCCUAACACUGUCUGGUAAAGAUGGCCC	164
miR-142	ACCCAUAAAGUAGAAAGCACUACUAACAGCACUGGAGGGUGUAGUGUUUCCUACUUUAUGGAUG	165
miR-143-1	GCGCAGCGCCCUGUCUCCCAGCCUGAGGUGCAGUGCUGCAUCUCUGGUCAGUUGGGAGUCUGAGAUGAAGCACUGUAGCUCAGGAAGAGAGAAGUUGUUCUGCAGC	166
miR-143-2	CCUGAGGUGCAGUGCUGCAUCUCUGGUCAGUUGGGAGUCUGAGAUGAAGCACUGUAGCUCAGG	167
miR-144-1	UGGGGCCCUGGCUGGGAUAUCAUCAUAUACUGUAAGUUUGCGAUGAGACACUACAGUAUAGAUGAUGUACUAGUCCGGGCACCCCC	168
miR-144-2	GGCUGGGAUAUCAUCAUAUACUGUAAGUUUGCGAUGAGACACUACAGUAUAGAUGAUGUACUAGUC	169
miR-145-1	CACCUUGUCCUCACGGUCCAGUUUUCCCAGGAAUCCCUUAGAUGCUAAGAUGGGGAUUCCUGGAAAUACUGUUCUUGAGGUCAUGGUU	170
miR-145-2	CUCACGGUCCAGUUUUCCCAGGAAUCCCUUAGAUGCUAAGAUGGGGAUUCCUGGAAAUACUGUUCUUGAG	171

Precursor Name	Sequence (5' To 3')*	SEQ ID NO.
miR-146-1	CCGAUGUGUAUCCUCAGCUUUGAGAACUGAAUU CCAUGGGUUGUGUCAGUGUCAGACCUCUGAAAU UCAGUUCUUCAGCUGGGAUAUCUCUGUCAUCGU	172
miR-146-2	AGCUUUGAGAACUGAAUUCCAUGGGUUGUGUCA GUGUCAGACCUGUGAAAUUCAGUUCUUCAGCU	173
miR-147	AAUCUAAAGACAACAUUUCUGCACACACACCAG ACUAUGGAAGCCAGUGUGUGGAAAUGCUUCUGC UAGAUU	174
miR-148a (miR-148)	GAGGCAAAGUUCUGAGACACUCCGACUCUGAGU AUGAUAGAAGUCAGUGCACUACAGAACUUUGUC UC	175
miR-148b	CAAGCACGAUUAGCAUUUGAGGUGAAGUUCUGU UAUACACUCAGGCUGUGGCUCUCUGAAAGUCAG UGCAUCACAGAACUUUGUCUCGAAAGCUUUCUA	176
miR-148b-small	AAGCACGAUUAGCAUUUGAGGUGAAGUUCUGUU AUACACUCAGGCUGUGGCUCUCUGAAAGUCAGU GCAU	177
miR-149-1	GCCGGCGCCCGAGCUCUGGCUCCGUGUCUUCACU CCCGUGCUUGUCCGAGGAGGGAGGGAGGGACGG GGGCUGUGCUGGGGCAGCUGGA	178
miR-149-2	GCUCUGGCUCCGUGUCUUCACUCCCGUGCUUGUC CGAGGAGGGAGGGAGGGAC	179
miR-150-1	CUCCCCAUGGCCCUGUCUCCCAACCCUUGUACCA GUGCUGGGCUCAGACCCUGGUACAGGCCUGGGG GACAGGGACCUGGGGAC	180
miR-150-2	CCCUGUCUCCCAACCCUUGUACCAGUGCUGGGCU CAGACCCUGGUACAGGCCUGGGGGACAGGG	181
miR-151	UUUCCUGCCCUCGAGGAGCUCACAGUCUAGUAU GUCUCAUCCCCUACUAGACUGAAGCUCCUUGAG GACAGG	182
miR-151-2	CCUGUCCUCAAGGAGCUUCAGUCUAGUAGGGGA UGAGACAUACUAGACUGUGAGCUCCUCGAGGGC AGG	183
miR-152-1	UGUCCCCCCCGGCCCAGGUUCUGUGAUACACUCC GACUCGGGCUCUGGAGCAGUCAGUGCAUGACAG AACUUGGGCCCGGAAGGACC	184
miR-152-2	GGCCCAGGUUCUGUGAUACACUCCGACUCGGGC UCUGGAGCAGUCAGUGCAUGACAGAACUUGGGC CCCGG	185
miR-153-1-1	CUCACAGCUGCCAGUGUCAUUUUUGUGAUCUGC AGCUAGUAUUCUCACUCCAGUUGCAUAGUCACA AAAGUGAUCAUUGGCAGGUGUGGC	186
miR-153-1-2	UCUCUCUCUCCCUCACAGCUGCCAGUGUCAUUGU CACAAAAGUGAUCAUUGGCAGGUGUGGCUGCUG CAUG	187

Precursor Name	Sequence (5' To 3')*	SEQ ID NO.
miR-153-2-1	AGCGGUGGCCAGUGUCAUUUUUGUGAUGUUGCA GCUAGUAAUAUGAGCCCAGUUGCAUAGUCACAA AAGUGAUCAUUGGAAACUGUG	188
miR-153-2-2	CAGUGUCAUUUUUGUGAUGUUGCAGCUAGUAAU AUGAGCCCAGUUGCAUAGUCACAAAAGUGAUCA UUG	189
miR-154-1	GUGGUACUUGAAGAUAGGUUAUCCGUGUUGCCU UCGCUUUAUUUGUGACGAAUCAUACACGGUUGA CCUAUUUUUCAGUACCAA	190
miR-154-2	GAAGAUAGGUUAUCCGUGUUGCCUUCGCUUUAU UUGUGACGAAUCAUACACGGUUGACCUAUUUUU	191
miR-155	CUGUUAAUGCUAAUCGUGAUAGGGGUUUUUGCC UCCAACUGACUCCUACAUAUUAGCAUUAACAG	192
miR-156 = miR-157=overlap miR-141	CCUAACACUGUCUGGUAAAGAUGGCUCCCGGGU GGGUUCUCUCGGCAGUAACCUUCAGGGAGCCCU GAAGACCAUGGAGGAC	193
miR-158-small = miR-192	GCCGAGACCGAGUGCACAGGGCUCUGACCUAUG AAUUGACAGCCAGUGCUCUCGUCUCCCCUCUGGC UGCCAAUUCCAUAGGUCACAGGUAUGUUCGCCU CAAUGCCAGC	194
miR-159-1-small	UCCCGCCCCCUGUAACAGCAACUCCAUGUGGAAG UGCCCACUGGUUCCAGUGGGGCUGCUGUUAUCU GGGGCGAGGGCCA	195
miR-161-small	AAAGCUGGGUUGAGAGGGCGAAAAAGGAUGAGG UGACUGGUCUGGGCUACGCUAUGCUGCGGCGCU CGGG	196
miR-163-1b-small	CAUUGGCCUCCUAAGCCAGGGAUUGUGGGUUCG AGUCCCACCCGGGGUAAAGAAAGGCCGAAUU	197
miR-163-3-small	CCUAAGCCAGGGAUUGUGGGUUCGAGUCCCACC UGGGGUAGAGGUGAAAGUUCCUUUUACGGAAUU UUUU	198
miR-162	CAAUGUCAGCAGUGCCUUAGCAGCACGUAAAUA UUGGCGUUAAGAUUCUAAAAUUAUCUCCAGUAU UAACUGUGCUGCUGAAGUAAGGUUGACCAUACU CUACAGUUG	199
miR-175-small=miR-224	GGGCUUUCAAGUCACUAGUGGUUCCGUUUAGUA GAUGAUUGUGCAUUGUUUCAAAAUGGUGCCCUA GUGACUACAAAGCCC	200
miR-177-small	ACGCAAGUGUCCUAAGGUGAGCUCAGGGAGCAC AGAAACCUCCAGUGGAACAGAAGGGCAAAAGCU CAUU	201
miR-180-small	CAUGUGUCACUUUCAGGUGGAGUUUCAAGAGUC CCUUCCUGGUUCACCGUCUCCUUUGCUCUUCCAC AAC	202

Precursor Name	Sequence (5' To 3')*	SEQ ID NO.
miR-181a	AGAAGGGCUAUCAGGCCAGCCUUCAGAGGACUC CAAGGAACAUUCAACGCUGUCGGUGAGUUUGGG AUUUGAAAAAACCACUGACCGUUGACUGUACCU UGGGGUCCUUA	203
miR-181b-1	CCUGUGCAGAGAUUAUUUUUUAAAAGGUCACAA UCAACAUUCAUUGCUGUCGGUGGGUUGAACUGU GUGGACAAGCUCACUGAACAAUGAAUGCAACUG UGGCCCCGCUU	204
miR-181b-2	CUGAUGGCUGCACUCAACAUUCAUUGCUGUCGG UGGGUUUGAGUCUGAAUCAACUCACUGAUCAAU GAAUGCAAACUGCGGACCAAACA	205
miR-181c	CGGAAAAUUUGCCAAGGGUUUGGGGGAACAUUC AACCUGUCGGUGAGUUUGGGCAGCUCAGGCAAA CCAUCGACCGUUGAGUGGACCCUGAGGCCUGGA AUUGCCAUCCU	206
miR-182-as	GAGCUGCUUGCCUCCCCCCGUUUUUGGCAAUGG UAGAACUCACACUGGUGAGGUAACAGGAUCCGG UGGUUCUAGACUUGCCAACUAUGGGGCGAGGAC UCAGCCGGCAC	207
miR-182	UUUUUGGCAAUGGUAGAACUCACACUGGUGAGG UAACAGGAUCCGGUGGUUCUAGACUUGCCAACU AUGG	208
miR-183	CCGCAGAGUGUGACUCCUGUUCUGUGUAUGGCA CUGGUAGAAUUCACUGUGAACAGUCUCAGUCAG UGAAUUACCGAAGGGCCAUAAACAGAGCAGAGA CAGAUCCACGA	209
miR-184-1	CCAGUCACGUCCCCUUAUCACUUUUCCAGCCCAG CUUUGUGACUGUAAGUGUUGGACGGAGAACUGA UAAGGGUAGGUGAUUGA	210
miR-184-2	CCUUAUCACUUUUCCAGCCCAGCUUUGUGACUG UAAGUGUUGGACGGAGAACUGAUAAGGGUAGG	211
miR-185-1	AGGGGGCGAGGGAUUGGAGAGAAAGGCAGUUCC UGAUGGUCCCCUCCCCCAGGGGCUGGCUUUCCUCU GGUCCUUCCCUCCCA	212
miR-185-2	AGGGAUUGGAGAGAAAGGCAGUUCCUGAUGGUC CCCUCCCCCAGGGGCUGGCUUUCCUCUGGUCCUU	213
miR-186-1	UGCUUGUAACUUUCCAAAGAAUUCUCCUUUUGG GCUUUCUGGUUUUAUUUUAAGCCCAAAGGUGAA UUUUUUGGGAAGUUUGAGCU	214
miR-186-2	ACUUUCCAAAGAAUUCUCCUUUUGGGCUUUCUG GUUUUAUUUUAAGCCCAAAGGUGAAUUUUUUGG GAAGU	215

Precursor Name	Sequence (5' To 3')*	SEQ ID NO.
miR-187	GGUCGGGCUCACCAUGACACAGUGUGAGACUCG GGCUACAACACAGGACCCGGGGCGCUGCUCUGA CCCCUCGUGUCUUGUGUUGCAGCCGGAGGGACG CAGGUCCGCA	216
miR-188-1	UGCUCCCUCUCUCACAUCCCUUGCAUGGUGGAG GGUGAGCUUUCUGAAAACCCCUCCCACAUGCAG GGUUUGCAGGAUGGCGAGCC	217
miR-188-2	UCUCACAUCCCUUGCAUGGUGGAGGGUGAGCUU UCUGAAAACCCCUCCCACAUGCAGGGUUUGCAG GA	218
miR-189-1	CUGUCGAUUGGACCCGCCCUCCGGUGCCUACUGA GCUGAUAUCAGUUCUCAUUUUACACACUGGCUC AGUUCAGCAGGAACAGGAGUCGAGCCCUUGAGC AA	219
miR-189-2	CUCCGGUGCCUACUGAGCUGAUAUCAGUUCUCA UUUUACACACUGGCUCAGUUCAGCAGGAACAGG AG	220
miR-190-1	UGCAGGCCUCUGUGUGAUAUGUUUGAUAUAUUA GGUUGUUAUUUAAUCCAACUAUAUAUCAAACAU AUUCCUACAGUGUCUUGCC	221
miR-190-2	CUGUGUGAUAUGUUUGAUAUAUUAGGUUGUUAU UUAAUCCAACUAUAUAUCAAACAUAUUCCUACA G	222
miR-191-1	CGGCUGGACAGCGGGCAACGGAAUCCCAAAAGC AGCUGUUGUCUCCAGAGCAUUCCAGCUGCGCUU GGAUUUCGUCCCCUGCUCUCCUGCCU	223
miR-191-2	AGCGGGCAACGGAAUCCCAAAAGCAGCUGUUGU CUCCAGAGCAUUCCAGCUGCGCUUGGAUUUCGU CCCCUGCU	224
miR-192-2/3	CCGAGACCGAGUGCACAGGGCUCUGACCUAUGA AUUGACAGCCAGUGCUCUCGUCUCCCCUCUGGCU GCCAAUUCCAUAGGUCACAGGUAUGUUCGCCUC AAUGCCAG	225
miR-192	GCCGAGACCGAGUGCACAGGGCUCUGACCUAUG AAUUGACAGCCAGUGCUCUCGUCUCCCCUCUGGC UGCCAAUUCCAUAGGUCACAGGUAUGUUCGCCU CAAUGCCAGC	226
miR-193-1	CGAGGAUGGGAGCUGAGGGCUGGGUCUUUGCGG GCGAGAUGAGGGUGUCGGAUCAACUGGCCUACA AAGUCCCAGUUCUCGGCCCCCG	227
miR-193-2	GCUGGGUCUUUGCGGGCGAGAUGAGGGUGUCGG AUCAACUGGCCUACAAAGUCCCAGU	228
miR-194-1	AUGGUGUUAUCAAGUGUAACAGCAACUCCAUGU GGACUGUGUACCAAUUUCCAGUGGAGAUGCUGU UACUUUUGAUGGUUACCAA	229

Precursor Name	Sequence (5' To 3')*	SEQ ID NO.
miR-194-2	GUGUAACAGCAACUCCAUGUGGACUGUGUACCA AUUUCCAGUGGAGAUGCUGUUACUUUUGAU	230
miR-195-1	AGCUUCCCUGGCUCUAGCAGCACAGAAAUAUUG GCACAGGGAAGCGAGUCUGCCAAUAUUGGCUGU GCUGCUCCAGGCAGGGUGGUG	231
miR-195-2	UAGCAGCACAGAAAUAUUGGCACAGGGAAGCGA GUCUGCCAAUAUUGGCUGUGCUGCU	232
miR-196-1	CUAGAGCUUGAAUUGGAACUGCUGAGUGAAUUA GGUAGUUUCAUGUUGUUGGGCCUGGGUUUCUGA ACACAACAACAUUAAACCACCCGAUUCACGGCA GUUACUGCUCC	233
miR-196a-1	GUGAAUUAGGUAGUUUCAUGUUGUUGGGCCUGG GUUUCUGAACACAACAACAUUAAACCACCCGAU UCAC	234
miR-196a-2 (miR-196-2)	UGCUCGCUCAGCUGAUCUGUGGCUUAGGUAGUU UCAUGUUGUUGGGAUUGAGUUUUGAACUCGGCA ACAAGAAACUGCCUGAGUUACAUCAGUCGGUUU UCGUCGAGGGC	235
miR-196	GUGAAUUAGGUAGUUUCAUGUUGUUGGGCCUGG GUUUCUGAACACAACAACAUUAAACCACCCGAU UCAC	236
miR-196b	ACUGGUCGGUGAUUUAGGUAGUUUCCUGUUGUU GGGAUCCACCUUUCUCUCGACAGCACGACACUGC CUUCAUUACUUCAGUUG	237
miR-197	GGCUGUGCCGGGUAGAGAGGGCAGUGGGAGGUA AGAGCUCUUCACCCUUCACCACCUUCUCCACCCA GCAUGGCC	238
miR-197-2	GUGCAUGUGUAUGUAUGUGUGCAUGUGCAUGUG UAUGUGUAUGAGUGCAUGCGUGUGUGC	239
miR-198	UCAUUGGUCCAGAGGGGAGAUAGGUUCCUGUGA UUUUUCCUUCUUCUCUAUAGAAUAAAUGA	240
miR-199a-1	GCCAACCCAGUGUUCAGACUACCUGUUCAGGAG GCUCUCAAUGUGUACAGUAGUCUGCACAUUGGU UAGGC	241
miR-199a-2	AGGAAGCUUCUGGAGAUCCUGCUCCGUCGCCCC AGUGUUCAGACUACCUGUUCAGGACAAUGCCGU UGUACAGUAGUCUGCACAUUGGUUAGACUGGGC AAGGGAGAGCA	242
miR-199b	CCAGAGGACACCUCCACUCCGUCUACCCAGUGUU UAGACUAUCUGUUCAGGACUCCCAAAUUGUACA GUAGUCUGCACAUUGGUUAGGCUGGGCUGGGUU AGACCCUCGG	243
miR-199s	GCCAACCCAGUGUUCAGACUACCUGUUCAGGAG GCUCUCAAUGUGUACAGUAGUCUGCACAUUGGU UAGGC	244

Precursor Name	Sequence (5' To 3')*	SEQ ID NO.
miR-200a	GCCGUGGCCAUCUUACUGGGCAGCAUUGGAUGG AGUCAGGUCUCUAAUACUGCCUGGUAAUGAUGA CGGC	245
miR-200b	CCAGCUCGGGCAGCCGUGGCCAUCUUACUGGGC AGCAUUGGAUGGAGUCAGGUCUCUAAUACUGCC UGGUAAUGAUGACGGCGGAGCCCUGCACG	246
miR-200c	CCCUCGUCUUACCCAGCAGUGUUUGGGUGCGGU UGGGAGUCUCUAAUACUGCCGGGUAAUGAUGGA GG	247
miR-202	GUUCCUUUUUCCUAUGCAUAUACUUCUUUGAGG AUCUGGCCUAAAGAGGUAUAGGGCAUGGGAAGA UGGAGC	248
miR-203	GUGUUGGGGACUCGCGCGCUGGGUCCAGUGGUU CUUAACAGUUCAACAGUUCUGUAGCGCAAUUGU GAAAUGUUUAGGACCACUAGACCCGGCGGGCGC GGCGACAGCGA	249
miR-204	GGCUACAGUCUUUCUUCAUGUGACUCGUGGACU UCCCUUUGUCAUCCUAUGCCUGAGAAUAUAUGA AGGAGGCUGGGAAGGCAAAGGGACGUUCAAUUG UCAUCACUGGC	250
miR-205	AAAGAUCCUCAGACAAUCCAUGUGCUUCUCUUG UCCUUCAUUCCACCGGAGUCUGUCUCAUACCCAA CCAGAUUUCAGUGGAGUGAAGUUCAGGAGGCAU GGAGCUGACA	251
miR-206-1	UGCUUCCCGAGGCCACAUGCUUCUUUAUAUCCCC AUAUGGAUUACUUUGCUAUGGAAUGUAAGGAAG UGUGUGGUUUCGGCAAGUG	252
miR-206-2	AGGCCACAUGCUUCUUUAUAUCCCCAUAUGGAU UACUUUGCUAUGGAAUGUAAGGAAGUGUGUGGU UUU	253
miR-208	UGACGGGCGAGCUUUUGGCCCGGGUUAUACCUG AUGCUCACGUAUAAGACGAGCAAAAAGCUUGUU GGUCA	254
miR-210	ACCCGGCAGUGCCUCCAGGCGCAGGGCAGCCCCU GCCCACCGCACACUGCGCUGCCCCAGACCCACUG UGCGUGUGACAGCGGCUGAUCUGUGCCUGGGCA GCGCGACCC	255
miR-211	UCACCUGGCCAUGUGACUUGUGGGCUUCCCUUU GUCAUCCUUCGCCUAGGGCUCUGAGCAGGGCAG GGACAGCAAAGGGGUGCUCAGUUGUCACUUCCC ACAGCACGGAG	256
miR-212	CGGGGCACCCCGCCCGGACAGCGCGCCGGCACCU UGGCUCUAGACUGCUUACUGCCCGGGCCGCCCUC AGUAACAGUCUCCAGUCACGGCCACCGACGCCUG GCCCCGCC	257

Precursor Name	Sequence (5' To 3')*	SEQ ID NO.
miR-213-2	CCUGUGCAGAGAUUAUUUUUUAAAAGGUCACAA UCAACAUUCAUUGCUGUCGGUGGGUUGAACUGU GUGGACAAGCUCACUGAACAAUGAAUGCAACUG UGGCCCCGCUU	258
miR-213	GAGUUUUGAGGUUGCUUCAGUGAACAUUCAACG CUGUCGGUGAGUUUGGAAUUAAAAUCAAAACCA UCGACCGUUGAUUGUACCCUAUGGCUAACCAUC AUCUACUCC	259
miR-214	GGCCUGGCUGGACAGAGUUGUCAUGUGUCUGCC UGUCUACACUUGCUGUGCAGAACAUCCGCUCAC CUGUACAGCAGGCACAGACAGGCAGUCACAUGA CAACCCAGCCU	260
miR-215	AUCAUUCAGAAAUGGUAUACAGGAAAAUGACCU AUGAAUUGACAGACAAUAUAGCUGAGUUUGUCU GUCAUUUCUUUAGGCCAAUAUUCUGUAUGACUG UGCUACUUCAA	261
miR-216	GAUGGCUGUGAGUUGGCUUAAUCUCAGCUGGCA ACUGUGAGAUGUUCAUACAAUCCCUCACAGUGG UCUCUGGGAUUAUGCUAAACAGAGCAAUUUCCU AGCCCUCACGA	262
miR-217	AGUAUAAUUAUUACAUAGUUUUUGAUGUCGCAG AUACUGCAUCAGGAACUGAUUGGAUAAGAAUCA GUCACCAUCAGUUCCUAAUGCAUUGCCUUCAGC AUCUAAACAAG	263
miR-218-1	GUGAUAAUGUAGCGAGAUUUUCUGUUGUGCUUG AUCUAACCAUGUGGUUGCGAGGUAUGAGUAAAA CAUGGUUCCGUCAAGCACCAUGGAACGUCACGC AGCUUUCUACA	264
miR-218-2	GACCAGUCGCUGCGGGGCUUUCCUUUGUGCUUG AUCUAACCAUGUGGUGGAACGAUGGAAACGGAA CAUGGUUCUGUCAAGCACCGCGGAAAGCACCGU GCUCUCCUGCA	265
miR-219	CCGCCCCGGGCCGCGGCUCCUGAUUGUCCAAACG CAAUUCUCGAGUCUAUGGCUCCGGCCGAGAGUU GAGUCUGGACGUCCCGAGCCGCCGCCCCCAAACC UCGAGCGGG	266
miR-219-1	CCGCCCCGGGCCGCGGCUCCUGAUUGUCCAAACG CAAUUCUCGAGUCUAUGGCUCCGGCCGAGAGUU GAGUCUGGACGUCCCGAGCCGCCGCCCCCAAACC UCGAGCGGG	267
miR-219-2	ACUCAGGGGCUUCGCCACUGAUUGUCCAAACGC AAUUCUUGUACGAGUCUGCGGCCAACCGAGAAU UGUGGCUGGACAUCUGUGGCUGAGCUCCGGG	268

Precursor Name	Sequence (5' To 3')*	SEQ ID NO.
miR-220	GACAGUGUGGCAUUGUAGGGCUCCACACCGUAU CUGACACUUUGGGCGAGGGCACCAUGCUGAAGG UGUUCAUGAUGCGGUCUGGGAACUCCUCACGGA UCUUACUGAUG	269
miR-221	UGAACAUCCAGGUCUGGGGCAUGAACCUGGCAU ACAAUGUAGAUUUCUGUGUUCGUUAGGCAACAG CUACAUUGUCUGCUGGGUUUCAGGCUACCUGGA AACAUGUUCUC	270
miR-222	GCUGCUGGAAGGUGUAGGUACCCUCAAUGGCUC AGUAGCCAGUGUAGAUCCUGUCUUUCGUAAUCA GCAGCUACAUCUGGCUACUGGGUCUCUGAUGGC AUCUUCUAGCU	271
miR-223	CCUGGCCUCCUGCAGUGCCACGCUCCGUGUAUUU GACAAGCUGAGUUGGACACUCCAUGUGGUAGAG UGUCAGUUUGUCAAAUACCCCAAGUGCGGCACA UGCUUACCAG	272
miR-224	GGGCUUUCAAGUCACUAGUGGUUCCGUUUAGUA GAUGAUUGUGCAUUGUUUCAAAAUGGUGCCCUA GUGACUACAAAGCCC	273

Precursor Name	Sequence (5' To 3')*	SEQ ID NO.
miR-294-1 (chr16)	CAAUCUUCCUUUAUCAUGGUAUUGAUUUUUCAG UGCUUCCCUUUUGUGUGAGAGAAGAUA	274
miR-296	AGGACCCUUCCAGAGGGCCCCCCCUCAAUCCUGU UGUGCCUAAUUCAGAGGGUUGGGUGGAGGCUCU CCUGAAGGGCUCU	275
miR-299	AAGAAAUGGUUUACCGUCCCACAUACAUUUUGA AUAUGUAUGUGGGAUGGUAAACCGCUUCUU	276
miR-301	ACUGCUAACGAAUGCUCUGACUUUAUUGCACUA CUGUACUUUACAGCUAGCAGUGCAAUAGUAUUG UCAAAGCAUCUGAAAGCAGG	277
miR-302a	CCACCACUUAAACGUGGAUGUACUUGCUUUGAA ACUAAAGAAGUAAGUGCUUCCAUGUUUUGGUGA UGG	278
miR-302b	GCUCCCUUCAACUUUAACAUGGAAGUGCUUUCU GUGACUUUAAAAGUAAGUGCUUCCAUGUUUUAG UAGGAGU	279
miR-302c	CCUUUGCUUUAACAUGGGGGUACCUGCUGUGUG AAACAAAAGUAAGUGCUUCCAUGUUUCAGUGGA GG	280
miR-302d	CCUCUACUUUAACAUGGAGGCACUUGCUGUGAC AUGACAAAAAAUAAGUGCUUCCAUGUUUGAGUGU GG	281

Precursor Name	Sequence (5' To 3')*	SEQ ID NO.
miR-320	GCUUCGCUCCCCUCCGCCUUCUCUUCCCGGUUCU UCCCGGAGUCGGGAAAAGCUGGGUUGAGAGGGC GAAAAAGGAUGAGGU	282
miR-321	UUGGCCUCCUAAGCCAGGGAUUGUGGGUUCGAG UCCCACCCGGGGUAAAGAAAGGCCGA	283
miR-323	UUGGUACUUGGAGAGAGGUGGUCCGUGGCGCGU UCGCUUUAUUUAUGGCGCACAUUACACGGUCGA CCUCUUUGCAGUAUCUAAUC	284
miR-324	CUGACUAUGCCUCCCCGCAUCCCCUAGGGCAUUG GUGUAAAGCUGGAGACCCACUGCCCCAGGUGCU GCUGGGGGUUGUAGUC	285
miR-325	AUACAGUGCUUGGUUCCUAGUAGGUGUCCAGUA AGUGUUUGUGACAUAAUUUGUUUAUUGAGGACC UCCUAUCAAUCAAGCACUGUGCUAGGCUCUGG	286
miR-326	CUCAUCUGUCUGUUGGGCUGGAGGCAGGGCCUU UGUGAAGGCGGGUGGUGCUCAGAUCGCCUCUGG GCCCUUCCUCCAGCCCCGAGGCGGAUUCA	287
miR-328	UGGAGUGGGGGGGCAGGAGGGGCUCAGGGAGAA AGUGCAUACAGCCCCUGGCCCUCUCUGCCCUUCC GUCCCCUG	288
miR-330	CUUUGGCGAUCACUGCCUCUCUGGGCCUGUGUC UUAGGCUCUGCAAGAUCAACCGAGCAAAGCACA CGGCCUGCAGAGAGGCAGCGCUCUGCCC	289
miR-331	GAGUUUGGUUUUGUUUGGGUUUGUUCUAGGUAU GGUCCCAGGGAUCCCAGAUCAAACCAGGCCCCUG GGCCUAUCCUAGAACCAACCUAAGCUC	290
miR-335	UGUUUUGAGCGGGGGUCAAGAGCAAUAACGAAA AAUGUUUGUCAUAAACCGUUUUUCAUUAUUGCU CCUGACCUCCUCUCAUUUGCUAUAUUCA	291
miR-337	GUAGUCAGUAGUUGGGGGGUGGGAACGGCUUCA UACAGGAGUUGAUGCACAGUUAUCCAGCUCCUA UAUGAUGCCUUUCUUCAUCCCCUUCAA	292
miR-338	UCUCCAACAAUAUCCUGGUGCUGAGUGAUGACU CAGGCGACUCCAGCAUCAGUGAUUUUGUUGAAG A	293
miR-339	CGGGGCGGCCGCUCUCCCUGUCCUCCAGGAGCUC ACGUGUGCCUGCCUGUGAGCGCCUCGACGACAG AGCCGGCGCCUGCCCCAGUGUCUGCGC	294
miR-340	UUGUACCUGGUGUGAUUAUAAAGCAAUGAGACU GAUUGUCAUAUGUCGUUUGUGGGAUCCGUCUCA GUUACUUUAUAGCCAUACCUGGUAUCUUA	295
miR-342	GAAACUGGGCUCAAGGUGAGGGGUGCUAUCUGU GAUUGAGGGACAUGGUUAAUGGAAUUGUCUCAC ACAGAAAUCGCACCCGUCACCUUGGCCUACUUA	296

Precursor Name	Sequence (5' To 3')*	SEQ ID NO.
miR-345	ACCCAAACCCUAGGUCUGCUGACUCCUAGUCCAG GGCUCGUGAUGGCUGGUGGGCCCUGAACGAGGG GUCUGGAGGCCUGGGUUUGAAUAUCGACAGC	297
miR-346	GUCUGUCUGCCCGCAUGCCUGCCUCUCUGUUGCU CUGAAGGAGGCAGGGGCUGGGCCUGCAGCUGCC UGGGCAGAGCGGCUCCUGC	298
miR-367	CCAUUACUGUUGCUAAUAUGCAACUCUGUUGAA UAUAAAUUGGAAUUGCACUUUAGCAAUGGUGAU GG	299
miR-368	AAAAGGUGGAUAUUCCUUCUAUGUUUAUGUUAU UUAUGGUUAAACAUAGAGGAAAUUCCACGUUUU	300
miR-369	UUGAAGGGAGAUCGACCGUGUUAUAUUCGCUUU AUUGACUUCGAAUAAUACAUGGUUGAUCUUUUC UCAG	301
miR-370	AGACAGAAGCCAGGUCACGUCUCUGCAGUUA CACAGCUCACGAGUGCCUGCUGGGGUGGAACCU GGUCUGUCU	302
miR-371	GUGGCACUCAAACUGUGGGGCACUUUCUGCUC UCUGGUGAAAGUGCCGCCAUCUUUUGAGUGUUA C	303
miR-372	GUGGGCCUCAAAUGUGGAGCACUAUUCUGAUGU CCAAGUGGAAAGUGCUGCGACAUUUGAGCGUCA C	304
miR-373	GGGAUACUCAAAAUGGGGGCGCUUUCCUUUUUG UCUGUACUGGGAAGUGCUUCGAUUUUGGGGUGU CCC	305
miR-374	UACAUCGGCCAUUAUAAUACAACCUGAUAAGUG UUAUAGCACUUAUCAGAUUGUAUUGUAAUUGUC UGUGUA	306
miR-hes1	AUGGAGCUGCUCACCCUGUGGGCCUCAAAUGUG GAGGAACUAUUCUGAUGUCCAAGUGGAAAGUGC UGCGACAUUUGAGCGUCACCGGUGACGCCCAUA UCA	307
miR-hes2	GCAUCCCCUCAGCCUGUGGCACUCAAACUGUGG GGGCACUUUCUGCUCUCUGGUGAAAGUGCCGCC AUCUUUUGAGUGUUACCGCUUGAGAAGACUCAA CC	308
miR-hes3	CGAGGAGCUCAUACUGGGAUACUCAAAAUGGGG GCGCUUUCCUUUUUGUCUGUUACUGGGAAGUGC UUCGAUUUUGGGGUGUCCCUGUUUGAGUAGGGC AUC	309

* An underlined sequence within a precursor sequence corresponds to a mature processed miR transcript (see Table 1b). Some precursor sequences have two

underlined sequences denoting two different mature miRs that are derived from the same precursor. All sequences are human.

Table 1b- Human Mature microRNA Sequences.

Mature miRNA Name	Mature miRNA Sequence (5' to 3')	SEQ ID NO.	Corresponding precursor microRNA(s); see Table 1a
let-7a	ugagguaguagguuguauaguu	310	*let-7a-1; let-7a-2; let-7a-3; let-7a-4*
let-7b	ugagguaguagguugugugguu	311	*let-7b*
let-7c	ugagguaguagguuguaugguu	312	*let-7c*
let-7d	agagguaguagguugcauagu	313	*let-7d; let-7d-v1*
let-7e	ugagguaggagguuguauagu	314	*let-7e*
let-7f	ugagguaguagauuguauaguu	315	*let-7f-1; let-7f-2-1; let-7f-2-2*
let-7g	ugagguaguaguuuguacagu	316	*let-7g*
let-7i	ugagguaguaguuugugcu	317	*let-7i*
miR-1	uggaauguaaagaaguaugua	318	*miR-1b; miR-1b-1; miR-1b-2*
miR-7	uggaagacuagugauuuuguu	319	*miR-7-1; miR-7-1a; miR-7-2; miR-7-3*
miR-9	ucuuugguuaucuagcuguauga	320	*miR-9-1; miR-9-2; miR-9-3*
*miR-9**	uaaagcuagauaaccgaaagu	321	*miR-9-1; miR-9-2; miR-9-3*
miR-10a	uacccuguagauccgaauuugug	322	*miR-10a*
miR-10b	uacccuguagaaccgaauuugu	323	*miR-10b*
miR-15a	uagcagcacauaaugguuugug	324	*miR-15a; miR-15a-2*
miR-15b	uagcagcacaucaugguuuaca	325	*miR-15b*
miR-16	uagcagcacguaaauauuggcg	326	*miR-16-1; miR-16-2; miR-16-13*
miR-17-5p	caaagugcuuacagugcagguagu	327	*miR-17*
miR-17-3p	acugcagugaaggcacuugu	328	*miR-17*

Mature miRNA Name	Mature miRNA Sequence (5' to 3')	SEQ ID NO.	Corresponding precursor microRNA(s); see Table 1a
miR-18	uaaggugcaucuagugcagaua	329	*miR-18; miR-18-13*
miR-19a	ugugcaaaucuaugcaaaacuga	330	*miR-19a; miR-19a-13*
miR-19b	ugugcaaauccaugcaaaacuga	331	*miR-19b-1; miR-19b-2*
miR-20	uaaagugcuuauagugcaggua	332	*miR-20 (miR-20a)*
miR-21	uagcuuaucagacugauguuga	333	*miR-21; miR-21-17*
miR-22	aagcugccaguugaagaacugu	334	*miR-22*
miR-23a	aucacauugccagggauuucc	335	*miR-23a*
miR-23b	aucacauugccagggauuaccac	336	*miR-23b*
miR-24	uggcucaguucagcaggaacag	337	*miR-24-1; miR-24-2; miR-24-19; miR-24-9*
miR-25	cauugcacuugucucggucuga	338	*miR-25*
miR-26a	uucaaguaauccaggauaggcu	339	*miR-26a; miR-26a-1; miR-26a-2*
miR-26b	uucaaguaauucaggauaggu	340	*miR-26b*
miR-27a	uucacaguggcuaaguuccgcc	341	*miR-27a*
miR-27b	uucacaguggcuaaguucug	342	*miR-27b-1; miR-27b-2*
miR-28	aaggagcucacagucuauugag	343	*miR-28*
miR-29a	cuagcaccaucugaaaucgguu	344	*miR-29a-2; miR-29a*
miR-29b	uagcaccauuugaaaucagu	345	*miR-29b-1; miR-29b-2*
miR-29c	uagcaccauuugaaaucgguua	346	*miR-29c*
miR-30a-5p	uguaaacauccucgacuggaagc	347	*miR-30a*
miR-30a-3p	cuuucagucggauguuugcagc	348	*miR-30a*
miR-30b	uguaaacauccuacacucagc	349	*miR-30b-1; miR-30b-2*
miR-30c	uguaaacauccuacacucucagc	350	*miR-30c*

Mature miRNA Name	Mature miRNA Sequence (5' to 3')	SEQ ID NO.	Corresponding precursor microRNA(s); see Table 1a
miR-30d	uguaaacaucccc gacuggaag	351	*miR-30d*
miR-30e	uguaaacauccuugacugga	352	*miR-30e*
miR-31	ggcaagaugcuggcauagcug	353	*miR-31*
miR-32	uauugcacauuacuaaguugc	354	*miR-32*
miR-33	gugcauuguaguugcauug	355	*miR-33; miR-33b*
miR-34a	uggcagugucuuagcugguugu	356	*miR-34a*
miR-34b	aggcagugucauuagcugauug	357	*miR-34b*
miR-34c	aggcaguguaguuagcugauug	358	*miR-34c*
miR-92	uauugcacuugucccggccugu	359	*miR-92-2; miR-92-1*
miR-93	aaagugcuguucgugcagguag	360	*miR-93-1; miR-93-2*
miR-95	uucaacggguauuuauugagca	361	*miR-95*
miR-96	uuuggcacuagcacauuuuugc	362	*miR-96*
miR-98	ugagguaguaaguuguauuguu	363	*miR-98*
miR-99a	aacccguagauccgaucuugug	364	*miR-99a*
miR-99b	cacccguagaaccgaccuugcg	365	*miR-99b*
miR-100	uacaguacugugauaacugaag	366	*miR-100*
miR-101	uacaguacugugauaacugaag	367	*miR-101-1; miR-101-2*
miR-103	agcagcauuguacagggcuauga	368	*miR-103-1*
miR-105	ucaaaugcucagacuccugu	369	*miR-105*
miR-106-a	aaaagugcuuacagugcagguagc	370	*miR-106-a*
miR-106-b	uaaagugcugacagugcagau	371	*miR-106-b*
miR-107	agcagcauuguacagggcuauca	372	*miR-107*

Mature miRNA Name	Mature miRNA Sequence (5' to 3')	SEQ ID NO.	Corresponding precursor microRNA(s); see Table 1a
miR-122a	uggagugugacaaugguguuugu	373	*miR-122a-1; miR-122a-2*
miR-124a	uuaaggcacgcggugaaugcca	374	*miR-124a-1; miR-124a-2; miR-124a-3*
miR-125a	ucccugagacccuuuaaccugug	375	*miR-125a-1; miR-125a-2*
miR-125b	ucccugagacccuaacuuguga	376	*miR-125b-1; miR-125b-2*
*miR-126**	cauuauuacuuuugguacgcg	377	*miR-126-1; miR-126-2*
miR-126	ucguaccgugaguaauaaugc	378	*miR-126-1; miR-126-2*
miR-127	ucggauccgucugagcuuggcu	379	*miR-127-1; miR-127-2*
miR-128a	ucacagugaaccggucucuuuu	380	*miR-128; miR-128a*
miR-128b	ucacagugaaccggucucuuuc	381	*miR-128b*
miR-129	cuuuuugcggucugggcuugc	382	*miR-129-1; miR-129-2*
miR-130a	cagugcaauguuaaaagggc	383	*miR-130a*
miR-130b	cagugcaaugaugaaagggcau	384	*miR-130b*
miR-132	uaacagucuacagccauggucg	385	*miR-132-1*
miR-133a	uuggucccccuucaaccagcugu	386	*miR-133a-1; miR-133a-2*
miR-133b	uuggucccccuucaaccagcua	387	*miR-133b*
miR-134	ugugacuggguugaccagaggg	388	*miR-134-1; miR-134-2*
miR-135a	uauggcuuuuuauuccuauguga	389	*miR-135a; miR-135a-2 (miR-135-2)*
miR-135b	uauggcuuuucauuccuaugug	390	*miR-135b*
miR-136	acuccauuuguuuugaugaugga	391	*miR-136-1; miR-136-2*
miR-137	uauugcuuaagaauacgcguag	392	*miR-137*
miR-138	agcugguguugugaauc	393	*miR-138-1; miR-138-2*

Mature miRNA Name	Mature miRNA Sequence (5' to 3')	SEQ ID NO.	Corresponding precursor microRNA(s); see Table 1a
miR-139	ucuacagugcacgugucu	394	*miR-139*
miR-140	agugguuuuacccuaugguag	395	*miR-140; miR-140as; miR-140s*
miR-141	aacacugucugguaaagaugg	396	*miR-141-1; miR-141-2*
miR-142-3p	uguaguguuuccuacuuuaugga	397	*miR-142*
miR-142-5p	cauaaaguagaaagcacuac	398	*miR-142*
miR-143	ugagaugaagcacuguagcuca	399	*miR-143-1*
miR-144	uacaguauagaugauguacuag	400	*miR-144-1; miR-144-2*
miR-145	guccaguuuucccaggaaucccuu	401	*miR-145-1; miR-145-2*
miR-146	ugagaacugaauuccauggguu	402	*miR-146-1; miR-146-2*
miR-147	guguguggaaaugcuucugc	403	*miR-147*
miR-148a	ucagugcacuacagaacuuugu	404	*miR-148a (miR-148)*
miR-148b	ucagugcaucacagaacuuugu	405	*miR-148b*
miR-149	ucuggcuccgugucuucacucc	406	*miR-149*
miR-150	ucucccaacccuuguaccagug	407	*miR-150-1; miR-150-2*
miR-151	acuagacugaagcuccuugagg	408	*miR-151*
miR-152	ucagugcaugacagaacuugg	409	*miR-152-1; miR-152-2*
miR-153	uugcauagucacaaaaguga	410	*miR-153-1-1; miR-153-1-2; miR-153-2-1; miR-153-2-2*
miR-154	uaggguauccguguugccuucg	411	*miR-154-1; miR-154-2*
*miR-154**	aaucauacacgguugaccuauu	412	*miR-154-1; miR-154-2*
miR-155	uuaaugcuaaucgugauagggg	413	*miR-155*
miR-181a	aacauucaacgcugucggugagu	414	*miR-181a*

Mature miRNA Name	Mature miRNA Sequence (5' to 3')	SEQ ID NO.	Corresponding precursor microRNA(s); see Table 1a
miR-181b	aacauucauugcugucgguggguu	415	*miR-181b-1; miR-181b-2*
miR-181c	aacauucaaccugucggugagu	416	*miR-181c*
miR-182	uuuggcaaugguagaacucaca	417	*miR-182; miR-182as*
*miR-182**	ugguucuagacuugccaacua	418	*miR-182; miR-182as*
miR-183	uauggcacugguagaauucacug	419	*miR-183*
miR-184	uggacggagaacugauaagggu	420	*miR-184-1; miR-184-2*
miR-185	uggagagaaaggcaguuc	421	*miR-185-1; miR-185-2*
miR-186	caaagaauucuccuuuugggcuu	422	*miR-186-1; miR-186-2*
miR-187	ucgugucuuguguugcagccg	423	*miR-187*
miR-188	caucccuugcaugguggagggu	424	*miR-188*
miR-189	gugccuacugagcugauaucagu	425	*miR-189-1; miR-189-2*
miR-190	ugauauguuugauauauuaggu	426	*miR-190-1; miR-190-2*
miR-191	caacggaaucccaaaagcagcu	427	*miR-191-1; miR-191-2*
miR-192	cugaccuaugaauugacagcc	428	*miR-192*
miR-193	aacuggccuacaaagucccag	429	*miR-193-1; miR-193-2*
miR-194	uguaacagcaacuccaugugga	430	*miR-194-1; miR-194-2*
miR-195	uagcagcacagaaauauuggc	431	*miR-195-1; miR-195-2*
miR-196a	uagguaguuucauguuguugg	432	*miR-196a; miR-196a-2 (miR196-2)*
miR-196b	uagguaguuuccuguuguugg	433	*miR-196b*
miR-197	uucaccaccuucuccacccagc	434	*miR-197*
miR-198	gguccagaggggagauagg	435	*miR-198*
miR-199a	cccaguguucagacuaccuguuc	436	*miR-199a-1; miR-199a-2*

Mature miRNA Name	Mature miRNA Sequence (5' to 3')	SEQ ID NO.	Corresponding precursor microRNA(s); see Table 1a
*miR-199a**	uacaguagucugcacauugguu	437	*miR-199a-1; miR-199a-2; miR-199s; miR-199b*
miR-199b	cccaguguuuagacuaucuguuc	438	*miR-199b*
miR-200a	uaacacugucugguaacgaugu	439	*miR-200a*
miR-200b	cucuaauacugccugguaaugaug	440	*miR-200b*
miR-200c	aauacugccggguaaugaugga	441	*miR-200c*
miR-202	agagguauagggcaugggaaga	442	*miR-202*
miR-203	gugaaauguuuaggaccacuag	443	*miR-203*
miR-204	uucccuuugucauccuaugccu	444	*miR-204*
miR-205	uccuucauuccaccggagucug	445	*miR-205*
miR-206	uggaauguaaggaagugugugg	446	*miR-206-1; miR-206-2*
miR-208	auaagacgagcaaaaagcuugu	447	*miR-208*
miR-210	cugugcgugugacagcggcug	448	*miR-210*
miR-211	uucccuuugucauccuucgccu	449	*miR-211*
miR-212	uaacagucuccagucacggcc	450	*miR-212*
miR-213	accaucgaccguugauuguacc	451	*miR-213*
miR-214	acagcaggcacagacaggcag	452	*miR-214*
miR-215	augaccuaugaauugacagac	453	*miR-215*
miR-216	uaaucucagcuggcaacugug	454	*miR-216*
miR-217	uacugcaucaggaacugauuggau	455	*miR-217*
miR-218	uugugcuugaucuaaccaugu	456	*miR-218-1; miR-218-2*
miR-219	ugauuguccaaacgcaauucu	457	*miR-219; miR-219-1; miR-219-2*
miR-220	ccacaccguaucugacacuuu	458	*miR-220*

Mature miRNA Name	Mature miRNA Sequence (5' to 3')	SEQ ID NO.	Corresponding precursor microRNA(s); see Table 1a
miR-221	agcuacauugucugcuggguuuc	459	*miR-221*
miR-222	agcuacaucuggcuacugggucuc	460	*miR-222*
miR-223	ugucaguuugucaaauacccc	461	*miR-223*
miR-224	caagucacuagugguuccguuua	462	*miR-224*
miR-296	agggccccccucaauccugu	463	*miR-296*
miR-299	ugguuuaccgucccacauacau	464	*miR-299*
miR-301	cagugcaauaguauugucaaagc	465	*miR-301*
miR-302a	uaagugcuuccauguuuugguga	466	*miR-302a*
*miR-302b**	acuuuaacauggaagugcuuucu	467	*miR-302b*
miR-302b	uaagugcuuccauguuuuaguag	468	*miR-302b*
*miR-302c**	uuuaacauggggguaccugcug	469	*miR-302c*
miR-302c	uaagugcuuccauguuucagugg	470	*miR-302c*
miR-302d	uaagugcuuccauguuugagugu	471	*miR-302d*
miR-320	aaaagcuggguugagagggcgaa	472	*miR-320*
miR-321	uaagccagggauuguggguuc	473	*miR-321*
miR-323	gcacauuacacggucgaccucu	474	*miR-323*
miR-324-5p	cgcauccccuagggcauuggugu	475	*miR-324*
miR-324-3p	ccacugccccaggugcugcugg	476	*miR-324*
miR-325	ccuaguagguguccaguaagu	477	*miR-325*
miR-326	ccucuggggcccuuccuccag	478	*miR-326*
miR-328	cuggcccucucugcccuuccgu	479	*miR-328*

Mature miRNA Name	Mature miRNA Sequence (5' to 3')	SEQ ID NO.	Corresponding precursor microRNA(s); see Table 1a
miR-330	gcaaagcacacggccugcagaga	480	*miR-330*
miR-331	gccccugggccuauccuagaa	481	*miR-331*
miR-335	ucaagagcaauaacgaaaaaugu	482	*miR-335*
miR-337	uccagcuccuauaugaugccuuu	483	*miR-337*
miR-338	uccagcaucagugauuuuguguga	484	*miR-338*
miR-339	ucccuguccuccaggagcuca	485	*miR-339*
miR-340	uccgucucaguuacuuuauagcc	486	*miR-340*
miR-342	ucucacacagaaaucgcacccguc	487	*miR-342*
miR-345	ugcugacuccuaguccagggc	488	*miR-345*
miR-346	ugucugcccgcaugccugccucu	489	*miR-346*
miR-367	aauugcacuuuagcaaugguga	490	*miR-367*
miR-368	acauagaggaaauuccacguuu	491	*miR-368*
miR-369	aauaauacaugguugaucuuu	492	*miR-369*
miR-370	gccugcugggguggaaccugg	493	*miR-370*
miR-371	gugccgccaucuuuugagugu	494	*miR-371*
miR-372	aaagugcugcgacauuugagcgu	495	*miR-372*
*miR-373**	acucaaaaugggggcgcuuucc	496	*miR-373*
miR-373	gaagugcuucgauuuuggggugu	497	*miR-373*
miR-374	uuauaauacaaccugauaagug	498	*miR-374*

The present invention encompasses methods of diagnosing whether a subject has, or is at risk for developing, a solid cancer, comprising measuring the level of at

least one miR gene product in a test sample from the subject and comparing the level of the miR gene product in the test sample to the level of a corresponding miR gene product in a control sample. As used herein, a "subject" can be any mammal that has, or is suspected of having, a solid cancer. In a preferred embodiment, the subject is a human who has, or is suspected of having, a solid cancer.

In one embodiment, the at least one miR gene product measured in the test sample is selected from the group consisting of miR-21, miR-17-5p, miR-191, miR-29b-2, miR-223, miR-128b, miR-199a-1, miR-24-1, miR-24-2, miR-146, miR-155, miR-181b-1, miR-20a, miR-107, miR-32, miR-92-2, miR-214, miR-30c, miR-25, miR-221, miR-106a and combinations thereof. In a particular embodiment, the miR gene product is miR-21, miR-191 or miR-17-5p. In another embodiment, the miR gene product is not miR-15a or miR-16-1. In an additional embodiment, the miR gene product is not miR 159-1 or miR-192. In an additional embodiment, the miR gene product is not miR-186, miR-101-1, miR-194, miR-215, miR-106b, miR-25, miR-93, miR-29b, miR-29a, miR-96, miR-182s, miR-182as, miR-183, miR-129-1, let-7a-1, let-7d, let-7f-1, miR-23b, miR-24-1, miR-27b, miR-32, miR-159-1, miR-192, miR-125b-1, let-7a-2, miR-100, miR-196-2, miR-148b, miR-190, miR-21, miR-301, miR-142s, miR-142as, miR-105-1, or miR-175. In a further embodiment, the miR gene product is not miR-21, miR-301, miR-142as, miR-142s, miR-194, miR-215, or miR-32. In another embodiment, the miR gene product is not miR-148, miR-10a, miR-196-1, miR-152, miR-196-2, miR-148b, miR-10b, miR-129-1, miR-153-2, miR-202, miR-139, let-7a, let-7f, or let-7d. In yet another embodiment, the miR gene product is not miR-15a, miR-16-1, miR-182, miR-181, miR-30, miR-15a, miR-16-1, miR-15b, miR-16-2, miR-195, miR-34, miR-153, miR-21, miR-217, miR-205, miR-204, miR-211, miR-143, miR-96, miR-103, miR-107, miR-129, miR-9, miR-137, miR-217, miR-186.

The solid cancer can be any cancer that arises from organs and solid tissues. Such cancers are typically associated with the formation and/or presence of tumor masses and can be carcinomas, sarcomas and lymphomas. Specific examples of solid cancers to be diagnosed by the methods of the invention include, but are not limited to, colon cancer, rectal cancer, stomach (gastric) cancer, pancreatic cancer, breast cancer, lung cancer, prostate cancer, bronchial cancer, testicular cancer, ovarian cancer, uterine cancer, penile cancer, melanoma and other skin cancers, liver cancer, esophogeal

cancer, cancers of the oral cavity and pharynx (e.g., tongue cancer, mouth cancer), cancers of the digestive system (e.g., intestinal cancer, gall bladder cancer), bone and joint cancers, cancers of the endocrine system (e.g., thyroid cancer), brain cancer, eye cancer, cancers of the urinary system (e.g., kidney cancer, urinary bladder cancer), Hodgkin disease and non-Hodgkin lymphoma. In particular embodiments, the solid cancer is not one or more of breast cancer, lung cancer, prostate cancer, pancreatic cancer or gastrointestinal cancer.

In one embodiment, the solid cancer is breast cancer or lung cancer and the at least one miR gene product measured in the test sample is selected from the group consisting of miR-210, miR-213 and a combination thereof.

In a further embodiment, the solid cancer is colon cancer, stomach cancer, prostate cancer or pancreas cancer and the at least one miR gene product measured in the test sample is miR-218-2.

In a certain embodiment of the invention, the solid cancer is breast cancer and the at least one miR gene product measured in the test sample is selected from the group consisting of miR-125b-1, miR-125b-2, miR-145, miR-21 and combinations thereof. In a related embodiment, the solid cancer is breast cancer and the at least one miR gene product in the test sample is selected from the group consisting of miR-21, miR-29b-2, miR-146, miR-125b-2, miR-125b-1, miR-10b, miR-145, miR-181a, miR-140, miR-213, miR-29a prec, miR-181b-1, miR-199b, miR-29b-1, miR-130a, miR-155, let-7a-2, miR-205, miR-29c, miR-224, miR-100, miR-31, miR-30c, miR-17-5p, miR-210, miR-122a, miR-16-2 and combinations thereof. In a related embodiment, the solid cancer is breast cancer and the at least one miR gene product is not miR-15a or miR-16-1. In a further embodiment, the solid cancer is breast cancer and the at least one miR gene product is not miR-145, miR-21, miR-155, miR-10b, miR-125b-1, miR-125b-2, let7a-2, let7a-3, let-7d, miR-122a, miR-191, miR-206, miR-210, let-7i, miR-009-1 (miR131-1), miR-34 (miR-170), miR-102 (miR-29b), miR-123 (miR-126), miR-140-as, miR-125a, miR-194, miR-204, miR-213, let-7f-2, miR-101, miR-128b, miR-136, miR-143, miR-149, miR-191, miR-196-1, miR-196-2, miR-202, miR-103-1, or miR-30c. In another embodiment, the solid cancer is breast cancer and the miR gene product is not miR-21, miR-125b-1, let-7a-2, let-7i, miR-100, let-7g, miR-31, miR-32a-1, miR-33b, miR-34a-2, miR-101-1, miR-135-1, miR-142as, miR-142s, miR-144, miR-

301, miR-29c, miR-30c, miR-106a, or miR-29b-1. In yet another embodiment, the solid cancer is breast cancer and the miR gene product is not miR-159-1 or miR-192. In an additional embodiment, the solid cancer is breast cancer and the miR gene product is not miR-186, miR-101-1, miR-194, miR-215, miR-106b, miR-25, miR-93, miR-29b, miR-29a, miR-96, miR-182s, miR-182as, miR-183, miR-129-1, let-7a-1, let-7d, let-7f-1, miR-23b, miR-24-1, miR-27b, miR-32, miR-159-1, miR-192, miR-125b-1, let-7a-2, miR-100, miR-196-2, miR-148b, miR-190, miR-21, miR-301, miR-142s, miR-142as, miR-105-1, or miR-175. In a further embodiment, the solid cancer is breast cancer and the miR gene product is not miR-21, miR-301, miR-142as, miR-142s, miR-194, miR-215, or miR-32. In another embodiment, the solid cancer is breast cancer and the miR gene product is not miR-148, miR-10a, miR-196-1, miR-152, miR-196-2, miR-148b, miR-10b, miR-129-1, miR-153-2, miR-202, miR-139, let-7a, let-7f, or let-7d. In yet another embodiment, the solid cancer is breast cancer and the miR gene product is not miR-181b, miR-181c, miR-181d, miR-30, miR-15b, miR-16-2, miR-153-1, miR-217, miR-205, miR-204, miR-103, miR-107, miR-129-2, miR-9 or miR-137.

In another embodiment, the solid cancer is colon cancer and the at least one miR gene product in the test sample is selected from the group consisting of miR-24-1, miR-29b-2, miR-20a, miR-10a, miR-32, miR-203, miR-106a, miR-17-5p, miR-30c, miR-223, miR-126*, miR-128b, miR-21, miR-24-2, miR-99b prec, miR-155, miR-213, miR-150, miR-107, miR-191, miR-221, miR-9-3 and combinations thereof. In another embodiment, the solid cancer is colon cancer and the miR gene product is not miR 159-1 or miR-192. In an additional embodiment, the solid cancer is colon cancer and the miR gene product is not miR-186, miR-101-1, miR-194, miR-215, miR-106b, miR-25, miR-93, miR-29b, miR-29a, miR-96, miR-182s, miR-182as, miR-183, miR-129-1, let-7a-1, let-7d, let-7f-1, miR-23b, miR-24-1, miR-27b, miR-32, miR-159-1, miR-192, miR-125b-1, let-7a-2, miR-100, miR-196-2, miR-148b, miR-190, miR-21, miR-301, miR-142s, miR-142as, miR-105-1, or miR-175. In a further embodiment, the solid cancer is colon cancer and the miR gene product is not miR-21, miR-301, miR-142as, miR-142s, miR-194, miR-215, or miR-32. In another embodiment, the solid cancer is colon cancer and the miR gene product is not miR-148, miR-10a, miR-196-1, miR-152, miR-196-2, miR-148b, miR-10b, miR-129-1, miR-153-2, miR-202, miR-139, let-7a,

let-7f, or let-7d. In yet another embodiment, the solid cancer is colon cancer and the miR gene product is not miR-181b, miR-181c, miR-181d, miR-30, miR-15b, miR-16-2, miR-153-1, miR-217, miR-205, miR-204, miR-103, miR-107, miR-129-2, miR-9 or miR-137.

In yet another embodiment, the solid cancer is lung cancer and the miR gene product in the test sample is selected from the group consisting of miR-21, miR-205, miR-200b, miR-9-1, miR-210, miR-148, miR-141, miR-132, miR-215, miR-128b, let-7g, miR-16-2, miR-129-1/2 prec, miR-126*, miR-142-as, miR-30d, miR-30a-5p, miR-7-2, miR-199a-1, miR-127, miR-34a prec, miR-34a, miR-136, miR-202, miR-196-2, miR-199a-2, let-7a-2, miR-124a-1, miR-149, miR-17-5p, miR-196-1 prec, miR-10a, miR-99b prec, miR-196-1, miR-199b, miR-191, miR-195, miR-155 and combinations thereof. In a related embodiment, the solid cancer is lung cancer and the at least one miR gene product is not miR-15a or miR-16-1. In a further embodiment, the solid cancer is lung cancer and the at least one miR gene product is not miR-21, miR-191, miR-126*, miR-210, miR-155, miR-143, miR-205, miR-126, miR-30a-5p, miR-140, miR-214, miR-218-2, miR-145, miR-106a, miR-192, miR-203, miR-150, miR-220, miR-192, miR-224, miR-24-2, miR-212, miR-9, miR-17, miR-124a-1, miR-95, miR-198, miR-216, miR-219-1, miR-197, miR-125a, miR-26a-1, miR-146, miR-199b, let7a-2, miR-27b, miR-32, miR-29b-2, miR-33, miR-181c, miR-101-1, miR-124a-3, miR-125b-1 or let7f-1. In another embodiment, the solid cancer is lung cancer and the at least one miR gene product is not miR-21, miR-182, miR-181, miR-30, miR-15a, miR-143, miR-205, miR-96, miR-103, miR-107, miR-129, miR-137, miR-186, miR-15b, miR-16-2, miR-195, miR-34, miR-153, miR-217, miR-204, miR-211, miR-9, miR-217, let-7a-2 or miR-32. In a further embodiment, the solid cancer is lung cancer and the miR gene product is not let-7c, let-7g, miR-7-3, miR-210, miR-31, miR-34a-1, miR-a-2, miR-99a, miR-100, miR-125b-2, miR-132, miR-135-1, miR-195, miR-34, miR-123, miR-203. In another embodiment, the solid cancer is lung cancer and the miR gene product is not miR 159-1 or miR-192. In an additional embodiment, the solid cancer is lung cancer and the miR gene product is not miR-186, miR-101-1, miR-194, miR-215, miR-106b, miR-25, miR-93, miR-29b, miR-29a, miR-96, miR-182s, miR-182as, miR-183, miR-129-1, let-7a-1, let-7d, let-7f-1, miR-23b, miR-24-1, miR-27b, miR-32, miR-159-1, miR-192, miR-125b-1, let-7a-2, miR-100, miR-196-2, miR-148b, miR-190,

miR-21, miR-301, miR-142s, miR-142as, miR-105-1, or miR-175. In a further embodiment, the solid cancer is lung cancer and the miR gene product is not miR-21, miR-301, miR-142as, miR-142s, miR-194, miR-215, or miR-32. In another embodiment, the solid cancer is lung cancer and the miR gene product is not miR-148, miR-10a, miR-196-1, miR-152, miR-196-2, miR-148b, miR-10b, miR-129-1, miR-153-2, miR-202, miR-139, let-7a, let-7f, or let-7d. In yet another embodiment, the solid cancer is lung cancer and the miR gene product is not miR-181b, miR-181c, miR-181d, miR-30, miR-15b, miR-16-2, miR-153-1, miR-217, miR-205, miR-204, miR-103, miR-107, miR-129-2, miR-9 or miR-137.

In a further embodiment, the solid cancer is pancreatic cancer and the at least one miR gene product measured in the test sample is selected from the group consisting of miR-103-1, miR-103-2, miR-155, miR-204 and combinations thereof. In a related embodiment, the solid cancer is pancreatic cancer and the miR gene product in the test sample is selected from the group consisting of miR-103-2, miR-103-1, miR-24-2, miR-107, miR-100, miR-125b-2, miR-125b-1, miR-24-1, miR-191, miR-23a, miR-26a-1, miR-125a, miR-130a, miR-26b, miR-145, miR-221, miR-126*, miR-16-2, miR-146, miR-214, miR-99b, miR-128b, miR-155, miR-29b-2, miR-29a, miR-25, miR-16-1, miR-99a, miR-224, miR-30d, miR-92-2, miR-199a-1, miR-223, miR-29c, miR-30b, miR-129-1/2, miR-197, miR-17-5p, miR-30c, miR-7-1, miR-93-1, miR-140, miR-30a-5p, miR-132, miR-181b-1, miR-152 prec, miR-23b, miR-20a, miR-222, miR-27a, miR-92-1, miR-21, miR-129-1/2 prec, miR-150, miR-32, miR-106a, miR-29b-1 and combinations thereof. In one embodiment, the solid cancer is pancreatic cancer and the miR gene product is not miR-15a or miR-16-1. In another embodiment, the solid cancer is pancreatic cancer and the miR gene product is not miR 159-1 or miR-192. In an additional embodiment, the solid cancer is pancreatic cancer and the miR gene product is not miR-186, miR-101-1, miR-194, miR-215, miR-106b, miR-25, miR-93, miR-29b, miR-29a, miR-96, miR-182s, miR-182as, miR-183, miR-129-1, let-7a-1, let-7d, let-7f-1, miR-23b, miR-24-1, miR-27b, miR-32, miR-159-1, miR-192, miR-125b-1, let-7a-2, miR-100, miR-196-2, miR-148b, miR-190, miR-21, miR-301, miR-142s, miR-142as, miR-105-1, or miR-175. In a further embodiment, the solid cancer is pancreatic cancer and the miR gene product is not miR-21, miR-301, miR-142as, miR-142s, miR-194, miR-215, or miR-32. In another embodiment, the solid cancer is

pancreatic cancer and the miR gene product is not miR-148, miR-10a, miR-196-1, miR-152, miR-196-2, miR-148b, miR-10b, miR-129-1, miR-153-2, miR-202, miR-139, let-7a, let-7f, or let-7d. In yet another embodiment, the solid cancer is pancreatic cancer and the miR gene product is not miR-181b, miR-181c, miR-181d, miR-30, miR-15b, miR-16-2, miR-153-1, miR-217, miR-205, miR-204, miR-103, miR-107, miR-129-2, miR-9 or miR-137.

In another embodiment, the solid cancer is prostate cancer and the miR gene product in the test sample is selected from the group consisting of let-7d, miR-128a prec, miR-195, miR-203, let-7a-2 prec, miR-34a, miR-20a, miR-218-2, miR-29a, miR-25, miR-95, miR-197, miR-135-2, miR-187, miR-196-1, miR-148, miR-191, miR-21, let-7i, miR-198, miR-199a-2, miR-30c, miR-17-5p, miR-92-2, miR-146, miR-181b-1 prec, miR-32, miR-206, miR-184 prec, miR-29a prec, miR-29b-2, miR-149, miR-181b-1, miR-196-1 prec, miR-93-1, miR-223, miR-16-1, miR-101-1, miR-124a-1, miR-26a-1, miR-214, miR-27a, miR-24-1, miR-106a, miR-199a-1 and combinations thereof. In a related embodiment, the solid cancer is prostate cancer and the miR gene product is not miR-15a or miR-16-1. In another embodiment, the solid cancer is prostate cancer and the miR gene product is not miR 159-1 or miR-192. In an additional embodiment, the solid cancer is prostate cancer and the miR gene product is not miR-186, miR-101-1, miR-194, miR-215, miR-106b, miR-25, miR-93, miR-29b, miR-29a, miR-96, miR-182s, miR-182as, miR-183, miR-129-1, let-7a-1, let-7d, let-7f-1, miR-23b, miR-24-1, miR-27b, miR-32, miR-159-1, miR-192, miR-125b-1, let-7a-2, miR-100, miR-196-2, miR-148b, miR-190, miR-21, miR-301, miR-142s, miR-142as, miR-105-1, or miR-175. In a further embodiment, the solid cancer is prostate cancer and the miR gene product is not miR-21, miR-301, miR-142as, miR-142s, miR-194, miR-215, or miR-32. In another embodiment, the solid cancer is prostate cancer and the miR gene product is not miR-148, miR-10a, miR-196-1, miR-152, miR-196-2, miR-148b, miR-10b, miR-129-1, miR-153-2, miR-202, miR-139, let-7a, let-7f, or let-7d. In yet another embodiment, the solid cancer is prostate cancer and the miR gene product is not miR-181b, miR-181c, miR-181d, miR-30, miR-15b, miR-16-2, miR-153-1, miR-217, miR-205, miR-204, miR-103, miR-107, miR-129-2, miR-9 or miR-137.

In yet another embodiment, the solid cancer is stomach cancer and the miR gene product in the test sample is selected from the group consisting of miR-223, miR-21,

miR-218-2, miR-103-2, miR-92-2, miR-25, miR-136, miR-191, miR-221, miR-125b-2, miR-103-1, miR-214, miR-222, miR-212 prec, miR-125b-1, miR-100, miR-107, miR-92-1, miR-96, miR-192, miR-23a, miR-215, miR-7-2, miR-138-2, miR-24-1, miR-99b, miR-33b, miR-24-2 and combinations thereof. In a related embodiment, the solid cancer is stomach cancer and the miR gene product is not miR-15a or miR-16-1. In another embodiment, the solid cancer is stomach cancer and the miR gene product is not miR 159-1 or miR-192. In an additional embodiment, the solid cancer is stomach cancer and the miR gene product is not miR-186, miR-101-1, miR-194, miR-215, miR-106b, miR-25, miR-93, miR-29b, miR-29a, miR-96, miR-182s, miR-182as, miR-183, miR-129-1, let-7a-1, let-7d, let-7f-1, miR-23b, miR-24-1, miR-27b, miR-32, miR-159-1, miR-192, miR-125b-1, let-7a-2, miR-100, miR-196-2, miR-148b, miR-190, miR-21, miR-301, miR-142s, miR-142as, miR-105-1, or miR-175. In a further embodiment, the solid cancer is stomach cancer and the miR gene product is not miR-21, miR-301, miR-142as, miR-142s, miR-194, miR-215, or miR-32. In another embodiment, the solid cancer is stomach cancer and the miR gene product is not miR-148, miR-10a, miR-196-1, miR-152, miR-196-2, miR-148b, miR-10b, miR-129-1, miR-153-2, miR-202, miR-139, let-7a, let-7f, or let-7d. In yet another embodiment, the solid cancer is stomach cancer and the miR gene product is not miR-181b, miR-181c, miR-181d, miR-30, miR-15b, miR-16-2, miR-153-1, miR-217, miR-205, miR-204, miR-103, miR-107, miR-129-2, miR-9 or miR-137.

The level of at least one miR gene product can be measured in a biological sample (e.g., cells, tissues) obtained from the subject. For example, a tissue sample (e.g., from a tumor) can be removed from a subject suspected of having a solid cancer by conventional biopsy techniques. In another embodiment, a blood sample can be removed from the subject, and blood cells (e.g., white blood cells) can be isolated for DNA extraction by standard techniques. The blood or tissue sample is preferably obtained from the subject prior to initiation of radiotherapy, chemotherapy or other therapeutic treatment. A corresponding control tissue or blood sample can be obtained from unaffected tissues of the subject, from a normal human individual or population of normal individuals, or from cultured cells corresponding to the majority of cells in the subject's sample. The control tissue or blood sample is then processed along with the sample from the subject, so that the levels of miR gene product produced from a given

miR gene in cells from the subject's sample can be compared to the corresponding miR gene product levels from cells of the control sample. A reference miR expression standard for the biological sample can also be used as a control.

An alteration (*e.g.*, an increase or decrease) in the level of a miR gene product in the sample obtained from the subject, relative to the level of a corresponding miR gene product in a control sample, is indicative of the presence of a solid cancer in the subject. In one embodiment, the level of the at least one miR gene product in the test sample is greater than the level of the corresponding miR gene product in the control sample (i.e., expression of the miR gene product is "up-regulated"). As used herein, expression of a miR gene product is "up-regulated" when the amount of miR gene product in a cell or tissue sample from a subject is greater than the amount of the same gene product in a control cell or tissue sample. In another embodiment, the level of the at least one miR gene product in the test sample is less than the level of the corresponding miR gene product in the control sample (i.e., expression of the miR gene product is "down-regulated"). As used herein, expression of a miR gene is "down-regulated" when the amount of miR gene product produced from that gene in a cell or tissue sample from a subject is less than the amount produced from the same gene in a control cell or tissue sample. The relative miR gene expression in the control and normal samples can be determined with respect to one or more RNA expression standards. The standards can comprise, for example, a zero miR gene expression level, the miR gene expression level in a standard cell line, the miR gene expression level in unaffected tissues of the subject, or the average level of miR gene expression previously obtained for a population of normal human controls.

The level of a miR gene product in a sample can be measured using any technique that is suitable for detecting RNA expression levels in a biological sample. Suitable techniques (e.g., Northern blot analysis, RT-PCR, *in situ* hybridization) for determining RNA expression levels in a biological sample (e.g., cells, tissues) are well known to those of skill in the art. In a particular embodiment, the level of at least one miR gene product is detected using Northern blot analysis. For example, total cellular RNA can be purified from cells by homogenization in the presence of nucleic acid extraction buffer, followed by centrifugation. Nucleic acids are precipitated, and DNA is removed by treatment with DNase and precipitation. The RNA molecules are then

separated by gel electrophoresis on agarose gels according to standard techniques, and transferred to nitrocellulose filters. The RNA is then immobilized on the filters by heating. Detection and quantification of specific RNA is accomplished using appropriately labeled DNA or RNA probes complementary to the RNA in question. See, for example, Molecular Cloning: A Laboratory Manual, J. Sambrook *et al.*, eds., 2nd edition, Cold Spring Harbor Laboratory Press, 1989, Chapter 7.

Suitable probes for Northern blot hybridization of a given miR gene product can be produced from the nucleic acid sequences provided in Table 1a and Table 1b and include, but are not limited to, probes having at least about 70%, 75%, 80%, 85%, 90%, 95%, 98%, 99% or complete complementarity to a miR gene product of interest. Methods for preparation of labeled DNA and RNA probes, and the conditions for hybridization thereof to target nucleotide sequences, are described in Molecular Cloning: A Laboratory Manual, J. Sambrook *et al.*, eds., 2nd edition, Cold Spring Harbor Laboratory Press, 1989, Chapters 10 and 11.

For example, the nucleic acid probe can be labeled with, *e.g.*, a radionuclide, such as ^{3}H, ^{32}P, ^{33}P, ^{14}C, or ^{35}S; a heavy metal; a ligand capable of functioning as a specific binding pair member for a labeled ligand (*e.g.*, biotin, avidin or an antibody); a fluorescent molecule; a chemiluminescent molecule; an enzyme or the like.

Probes can be labeled to high specific activity by either the nick translation method of Rigby *et al.* (1977), *J. Mol. Biol.* 113:237-251 or by the random priming method of Fienberg *et al.* (1983), *Anal. Biochem.* 132: 6-13. The latter is the method of choice for synthesizing ^{32}P-labeled probes of high specific activity from single-stranded DNA or from RNA templates. For example, by replacing preexisting nucleotides with highly radioactive nucleotides according to the nick translation method, it is possible to prepare ^{32}P-labeled nucleic acid probes with a specific activity well in excess of 10^{8} cpm/microgram. Autoradiographic detection of hybridization can then be performed by exposing hybridized filters to photographic film. Densitometric scanning of the photographic films exposed by the hybridized filters provides an accurate measurement of miR gene transcript levels. Using another approach, miR gene transcript levels can

be quantified by computerized imaging systems, such as the Molecular Dynamics 400-B 2D Phosphorimager available from Amersham Biosciences, Piscataway, NJ.

Where radionuclide labeling of DNA or RNA probes is not practical, the random-primer method can be used to incorporate an analogue, for example, the dTTP analogue 5-(N-(N-biotinyl-epsilon-aminocaproyl)-3-aminoallyl)deoxyuridine triphosphate, into the probe molecule. The biotinylated probe oligonucleotide can be detected by reaction with biotin-binding proteins, such as avidin, streptavidin, and antibodies (e.g., anti-biotin antibodies) coupled to fluorescent dyes or enzymes that produce color reactions.

In addition to Northern and other RNA hybridization techniques, determining the levels of RNA transcripts can be accomplished using the technique of *in situ* hybridization. This technique requires fewer cells than the Northern blotting technique, and involves depositing whole cells onto a microscope cover slip and probing the nucleic acid content of the cell with a solution containing radioactive or otherwise labeled nucleic acid (e.g., cDNA or RNA) probes. This technique is particularly well-suited for analyzing tissue biopsy samples from subjects. The practice of the *in situ* hybridization technique is described in more detail in U.S. Pat. No. 5,427,916. Suitable probes for *in situ* hybridization of a given miR gene product can be produced from the nucleic acid sequences provided in Table 1a and Table 1b, and include, but are not limited to, probes having at least about 70%, 75%, 80%, 85%, 90%, 95%, 98%, 99% or complete complementarity to a miR gene product of interest, as described above.

The relative number of miR gene transcripts in cells can also be determined by reverse transcription of miR gene transcripts, followed by amplification of the reverse-transcribed transcripts by polymerase chain reaction (RT-PCR). The levels of miR gene transcripts can be quantified in comparison with an internal standard, for example, the level of mRNA from a "housekeeping" gene present in the same sample. A suitable "housekeeping" gene for use as an internal standard includes, e.g., myosin or glyceraldehyde-3-phosphate dehydrogenase (G3PDH). Methods for performing quantitative and semi-quantitative RT-PCR, and variations thereof, are well known to those of skill in the art.

In some instances, it may be desirable to simultaneously determine the expression level of a plurality of different miR gene products in a sample. In other instances, it may be desirable to determine the expression level of the transcripts of all known miR genes correlated with a cancer. Assessing cancer-specific expression levels for hundreds of miR genes or gene products is time consuming and requires a large amount of total RNA (e.g., at least 20 μg for each Northern blot) and autoradiographic techniques that require radioactive isotopes.

To overcome these limitations, an oligolibrary, in microchip format (i.e., a microarray), may be constructed containing a set of oligonucleotide (e.g., oligodeoxynucleotides) probes that are specific for a set of miR genes. Using such a microarray, the expression level of multiple microRNAs in a biological sample can be determined by reverse transcribing the RNAs to generate a set of target oligodeoxynucleotides, and hybridizing them to probe the oligonucleotides on the microarray to generate a hybridization, or expression, profile. The hybridization profile of the test sample can then be compared to that of a control sample to determine which microRNAs have an altered expression level in solid cancer cells. As used herein, "probe oligonucleotide" or "probe oligodeoxynucleotide" refers to an oligonucleotide that is capable of hybridizing to a target oligonucleotide. "Target oligonucleotide" or "target oligodeoxynucleotide" refers to a molecule to be detected (e.g., via hybridization). By "miR-specific probe oligonucleotide" or "probe oligonucleotide specific for a miR" is meant a probe oligonucleotide that has a sequence selected to hybridize to a specific miR gene product, or to a reverse transcript of the specific miR gene product.

An "expression profile" or "hybridization profile" of a particular sample is essentially a fingerprint of the state of the sample; while two states may have any particular gene similarly expressed, the evaluation of a number of genes simultaneously allows the generation of a gene expression profile that is unique to the state of the cell. That is, normal tissue may be distinguished from cancerous (e.g., tumor) tissue, and within cancerous tissue, different prognosis states (for example, good or poor long term survival prospects) may be determined. By comparing expression profiles of solid cancer tissue in different states, information regarding which genes are important (including both up- and down-regulation of genes) in each of these states is obtained.

The identification of sequences that are differentially expressed in solid cancer tissue, as well as differential expression resulting in different prognostic outcomes, allows the use of this information in a number of ways. For example, a particular treatment regime may be evaluated (e.g., to determine whether a chemotherapeutic drug acts to improve the long-term prognosis in a particular patient). Similarly, diagnosis may be done or confirmed by comparing patient samples with known expression profiles. Furthermore, these gene expression profiles (or individual genes) allow screening of drug candidates that suppress the solid cancer expression profile or convert a poor prognosis profile to a better prognosis profile.

Accordingly, the invention provides methods of diagnosing whether a subject has, or is at risk for developing, a solid cancer, comprising reverse transcribing RNA from a test sample obtained from the subject to provide a set of target oligodeoxynucleotides, hybridizing the target oligodeoxynucleotides to a microarray comprising miRNA-specific probe oligonucleotides to provide a hybridization profile for the test sample, and comparing the test sample hybridization profile to a hybridization profile generated from a control sample or reference standard, wherein an alteration in the signal of at least one miRNA is indicative of the subject either having, or being at risk for developing, a solid cancer. In one embodiment, the microarray comprises miRNA-specific probe oligonucleotides for a substantial portion of all known human miRNAs. In a particular embodiment, the microarray comprises miRNA-specific probe oligonucleotides for one or more miRNAs selected from the group consisting of miR-21, miR-17-5p, miR-191, miR-29b-2, miR-223, miR-128b, miR-199a-1, miR-24-1, miR-24-2, miR-146, miR-155, miR-181b-1, miR-20a, miR-107, miR-32, miR-92-2, miR-214, miR-30c, miR-25, miR-221, miR-106a and combinations thereof.

The microarray can be prepared from gene-specific oligonucleotide probes generated from known miRNA sequences. The array may contain two different oligonucleotide probes for each miRNA, one containing the active, mature sequence and the other being specific for the precursor of the miRNA. The array may also contain controls, such as one or more mouse sequences differing from human orthologs by only a few bases, which can serve as controls for hybridization stringency conditions. tRNAs or other RNAs (e.g., rRNAs, mRNAs) from both species may also

be printed on the microchip, providing an internal, relatively stable, positive control for specific hybridization. One or more appropriate controls for non-specific hybridization may also be included on the microchip. For this purpose, sequences are selected based upon the absence of any homology with any known miRNAs.

The microarray may be fabricated using techniques known in the art. For example, probe oligonucleotides of an appropriate length, e.g., 40 nucleotides, are 5'-amine modified at position C6 and printed using commercially available microarray systems, e.g., the GeneMachine OmniGrid™ 100 Microarrayer and Amersham CodeLink™ activated slides. Labeled cDNA oligomer corresponding to the target RNAs is prepared by reverse transcribing the target RNA with labeled primer. Following first strand synthesis, the RNA/DNA hybrids are denatured to degrade the RNA templates. The labeled target cDNAs thus prepared are then hybridized to the microarray chip under hybridizing conditions, e.g., 6X SSPE/30% formamide at 25°C for 18 hours, followed by washing in 0.75X TNT (Tris HC1/NaC1/Tween™ 20) at 37°C for 40 minutes. At positions on the array where the immobilized probe DNA recognizes a complementary target cDNA in the sample, hybridization occurs. The labeled target cDNA marks the exact position on the array where binding occurs, allowing automatic detection and quantification. The output consists of a list of hybridization events, indicating the relative abundance of specific cDNA sequences, and therefore the relative abundance of the corresponding complementary miRs, in the patient sample. According to one embodiment, the labeled cDNA oligomer is a biotin-labeled cDNA, prepared from a biotin-labeled primer. The microarray is then processed by direct detection of the biotin-containing transcripts using, e.g., Streptavidin-Alexa647 conjugate, and scanned utilizing conventional scanning methods. Image intensities of each spot on the array are proportional to the abundance of the corresponding miR in the patient sample.

The use of the array has several advantages for miRNA expression detection. First, the global expression of several hundred genes can be identified in the same sample at one time point. Second, through careful design of the oligonucleotide probes, expression of both mature and precursor molecules can be identified. Third, in comparison with Northern blot analysis, the chip requires a small amount of RNA, and provides reproducible results using 2.5 μg of total RNA. The relatively limited number

of miRNAs (a few hundred per species) allows the construction of a common microarray for several species, with distinct oligonucleotide probes for each. Such a tool would allow for analysis of trans-species expression for each known miR under various conditions.

In addition to use for quantitative expression level assays of specific miRs, a microchip containing miRNA-specific probe oligonucleotides corresponding to a substantial portion of the miRNome, preferably the entire miRNome, may be employed to carry out miR gene expression profiling, for analysis of miR expression patterns. Distinct miR signatures can be associated with established disease markers, or directly with a disease state.

According to the expression profiling methods described herein, total RNA from a sample from a subject suspected of having a cancer (e.g., a solid cancer) is quantitatively reverse transcribed to provide a set of labeled target oligodeoxynucleotides complementary to the RNA in the sample. The target oligodeoxynucleotides are then hybridized to a microarray comprising miRNA-specific probe oligonucleotides to provide a hybridization profile for the sample. The result is a hybridization profile for the sample representing the expression pattern of miRNA in the sample. The hybridization profile comprises the signal from the binding of the target oligodeoxynucleotides from the sample to the miRNA-specific probe oligonucleotides in the microarray. The profile may be recorded as the presence or absence of binding (signal vs. zero signal). More preferably, the profile recorded includes the intensity of the signal from each hybridization. The profile is compared to the hybridization profile generated from a normal, i.e., noncancerous, control sample. An alteration in the signal is indicative of the presence of, or propensity to develop, cancer in the subject.

Other techniques for measuring miR gene expression are also within the skill in the art, and include various techniques for measuring rates of RNA transcription and degradation.

The invention also provides methods of determining the prognosis of a subject with a solid cancer, comprising measuring the level of at least one miR gene product, which is associated with a particular prognosis in a solid cancer (e.g., a good or positive prognosis, a poor or adverse prognosis), in a test sample from the subject. According to

these methods, an alteration in the level of a miR gene product that is associated with a particular prognosis in the test sample, as compared to the level of a corresponding miR gene product in a control sample, is indicative of the subject having a solid cancer with a particular prognosis. In one embodiment, the miR gene product is associated with an adverse (i.e., poor) prognosis. Examples of an adverse prognosis include, but are not limited to, low survival rate and rapid disease progression. In certain embodiments, the level of the at least one miR gene product is measured by reverse transcribing RNA from a test sample obtained from the subject to provide a set of target oligodeoxynucleotides, hybridizing the target oligodeoxynucleotides to a microarray that comprises miRNA-specific probe oligonucleotides to provide a hybridization profile for the test sample, and comparing the test sample hybridization profile to a hybridization profile generated from a control sample.

Without wishing to be bound by any one theory, it is believed that alterations in the level of one or more miR gene products in cells can result in the deregulation of one or more intended targets for these miRs, which can lead to the formation of solid cancers. Therefore, altering the level of the miR gene product (e.g., by decreasing the level of a miR gene product that is up-regulated in solid cancer cells, by increasing the level of a miR gene product that is down-regulated in solid cancer cells) may successfully treat the solid cancer.

Accordingly, the present invention encompasses methods of inhibiting tumorigenesis in a subject who has, or is suspected of having, a solid cancer wherein at least one miR gene product is deregulated (e.g., down-regulated, up-regulated) in the cancer cells of the subject. When the at least one isolated miR gene product is down-regulated in the cancer cells (e.g., miR-145, miR-155, miR-218-2), the method comprises administering an effective amount of the at least one isolated miR gene product, or an isolated variant or biologically-active fragment thereof, such that proliferation of cancer cells in the subject is inhibited. In one embodiment, the isolated miR gene product that is administered is not miR-15a or miR-16-1. In another embodiment, the miR gene product is not miR 159-1 or miR-192. In an additional embodiment, the miR gene product is not miR-186, miR-101-1, miR-194, miR-215, miR-106b, miR-25, miR-93, miR-29b, miR-29a, miR-96, miR-182s, miR-182as, miR-183, miR-129-1, let-7a-1, let-7d, let-7f-1, miR-23b, miR-24-1, miR-27b, miR-32, miR-

159-1, miR-192, miR-125b-1, let-7a-2, miR-100, miR-196-2, miR-148b, miR-190, miR-21, miR-301, miR-142s, miR-142as, miR-105-1, or miR-175. In a further embodiment, the miR gene product is not miR-21, miR-301, miR-142as, miR-142s, miR-194, miR-215, or miR-32. In another embodiment, the miR gene product is not miR-148, miR-10a, miR-196-1, miR-152, miR-196-2, miR-148b, miR-10b, miR-129-1, miR-153-2, miR-202, miR-139, let-7a, let-7f, or let-7d. In yet another embodiment, the miR gene product is not miR-30, miR-15b, miR-16-2, miR-217, miR-205, miR-204, miR-103, miR-107, miR-9, and miR-137. In a further embodiment, the miR gene product is not miR-145, miR-21, miR-155, miR-10b, miR-125b-1, miR-125b-2, let7a-2, let7a-3, let-7d, miR-122a, miR-191, miR-206, miR-210, let-7i, miR-009-1 (miR131-1), miR-34 (miR-170), miR-102 (miR-29b), miR-123 (miR-126), miR-140-as, miR-125a, miR-194, miR-204, miR-213, let-7f-2, miR-101, miR-128b, miR-136, miR-143, miR-149, miR-191, miR-196-1, miR-196-2, miR-202, miR-103-1, or miR-30c. In another embodiment, the miR gene product is not miR-21, miR-125b-1, let-7a-2, let-7i, miR-100, let-7g, miR-31, miR-32a-1, miR-33b, miR-34a-2, miR-101-1, miR-135-1, miR-142as, miR-142s, miR-144, miR-301, miR-29c, miR-30c, miR-106a, or miR-29b-1.

For example, when a miR gene product is down-regulated in a cancer cell in a subject, administering an effective amount of an isolated miR gene product to the subject can inhibit proliferation of the cancer cell. The isolated miR gene product that is administered to the subject can be identical to the endogenous wild-type miR gene product (e.g., a miR gene product shown in Table 1a or Table 1b) that is down-regulated in the cancer cell or it can be a variant or biologically-active fragment thereof. As defined herein, a "variant" of a miR gene product refers to a miRNA that has less than 100% identity to a corresponding wild-type miR gene product and possesses one or more biological activities of the corresponding wild-type miR gene product. Examples of such biological activities include, but are not limited to, inhibition of expression of a target RNA molecule (e.g., inhibiting translation of a target RNA molecule, modulating the stability of a target RNA molecule, inhibiting processing of a target RNA molecule) and inhibition of a cellular process associated with solid cancer (e.g., cell differentiation, cell growth, cell death). These variants include species variants and variants that are the consequence of one or more mutations

(e.g., a substitution, a deletion, an insertion) in a miR gene. In certain embodiments, the variant is at least about 70%, 75%, 80%, 85%, 90%, 95%, 98%, or 99% identical to a corresponding wild-type miR gene product.

As defined herein, a "biologically-active fragment" of a miR gene product refers to an RNA fragment of a miR gene product that possesses one or more biological activities of a corresponding wild-type miR gene product. As described above, examples of such biological activities include, but are not limited to, inhibition of expression of a target RNA molecule and inhibition of a cellular process associated with solid cancer. In certain embodiments, the biologically-active fragment is at least about 5, 7, 10, 12, 15, or 17 nucleotides in length. In a particular embodiment, an isolated miR gene product can be administered to a subject in combination with one or more additional anti-cancer treatments. Suitable anti-cancer treatments include, but are not limited to, chemotherapy, radiation therapy and combinations thereof (e.g., chemoradiation).

When the at least one isolated miR gene product is up-regulated in the cancer cells, the method comprises administering to the subject an effective amount of at least one compound for inhibiting expression of the at least one miR gene product, referred to herein as miR gene expression-inhibition compounds, such that proliferation of solid cancer cells is inhibited. In a particular embodiment, the at least one miR expression-inhibition compound is specific for a miR gene product selected from the group consisting of miR-21, miR-17-5p, miR-191, miR-29b-2, miR-223, miR-128b, miR-199a-1, miR-24-1, miR-24-2, miR-146, miR-155, miR-181b-1, miR-20a, miR-107, miR-32, miR-92-2, miR-214, miR-30c, miR-25, miR-221, miR-106a and combinations thereof. A miR gene expression-inhibiting compound can be administered to a subject in combination with one or more additional anti-cancer treatments. Suitable anti-cancer treatments include, but are not limited to, chemotherapy, radiation therapy and combinations thereof (e.g., chemoradiation).

The terms "treat", "treating" and "treatment", as used herein, refer to ameliorating symptoms associated with a disease or condition, for example, a solid cancer, including preventing or delaying the onset of the disease symptoms, and/or lessening the severity or frequency of symptoms of the disease or condition. The terms "subject", "patient" and "individual" are defined herein to include animals, such as

mammals, including, but not limited to, primates, cows, sheep, goats, horses, dogs, cats, rabbits, guinea pigs, rats, mice or other bovine, ovine, equine, canine, feline, rodent, or murine species. In a preferred embodiment, the animal is a human.

As used herein, an "effective amount" of an isolated miR gene product is an amount sufficient to inhibit proliferation of a cancer cell in a subject suffering from a solid cancer. One skilled in the art can readily determine an effective amount of a miR gene product to be administered to a given subject, by taking into account factors, such as the size and weight of the subject; the extent of disease penetration; the age, health and sex of the subject; the route of administration; and whether the administration is regional or systemic.

For example, an effective amount of an isolated miR gene product can be based on the approximate weight of a tumor mass to be treated. The approximate weight of a tumor mass can be determined by calculating the approximate volume of the mass, wherein one cubic centimeter of volume is roughly equivalent to one gram. An effective amount of the isolated miR gene product based on the weight of a tumor mass can be in the range of about 10-500 micrograms/gram of tumor mass. In certain embodiments, the tumor mass can be at least about 10 micrograms/gram of tumor mass, at least about 60 micrograms/gram of tumor mass or at least about 100 micrograms/gram of tumor mass.

An effective amount of an isolated miR gene product can also be based on the approximate or estimated body weight of a subject to be treated. Preferably, such effective amounts are administered parenterally or enterally, as described herein. For example, an effective amount of the isolated miR gene product is administered to a subject can range from about 5~3000 micrograms/kg of body weight, from about 700 - 1000 micrograms/kg of body weight, or greater than about 1000 micrograms/kg of body weight.

One skilled in the art can also readily determine an appropriate dosage regimen for the administration of an isolated miR gene product to a given subject. For example, a miR gene product can be administered to the subject once (*e.g.*, as a single injection or deposition). Alternatively, a miR gene product can be administered once or twice daily to a subject for a period of from about three to about twenty-eight days, more particularly from about seven to about ten days. In a particular dosage regimen, a miR

gene product is administered once a day for seven days. Where a dosage regimen comprises multiple administrations, it is understood that the effective amount of the miR gene product administered to the subject can comprise the total amount of gene product administered over the entire dosage regimen.

As used herein, an "isolated" miR gene product is one that is synthesized, or altered or removed from the natural state through human intervention. For example, a synthetic miR gene product, or a miR gene product partially or completely separated from the coexisting materials of its natural state, is considered to be "isolated." An isolated miR gene product can exist in substantially-purified form, or can exist in a cell into which the miR gene product has been delivered. Thus, a miR gene product that is deliberately delivered to, or expressed in, a cell is considered an "isolated" miR gene product. A miR gene product produced inside a cell from a miR precursor molecule is also considered to be an "isolated" molecule. According to the invention, the isolated miR gene products described herein can be used for the manufacture of a medicament for treating a solid cancer in a subject (e.g., a human).

Isolated miR gene products can be obtained using a number of standard techniques. For example, the miR gene products can be chemically synthesized or recombinantly produced using methods known in the art. In one embodiment, miR gene products are chemically synthesized using appropriately protected ribonucleoside phosphoramidites and a conventional DNA/RNA synthesizer. Commercial suppliers of synthetic RNA molecules or synthesis reagents include, e.g., Proligo (Hamburg, Germany), Dharmacon Research (Lafayette, CO, U.S.A.), Pierce Chemical (part of Perbio Science, Rockford, IL, U.S.A.), Glen Research (Sterling, VA, U.S.A.), ChemGenes (Ashland, MA, U.S.A.) and Cruachem (Glasgow, UK).

Alternatively, the miR gene products can be expressed from recombinant circular or linear DNA plasmids using any suitable promoter. Suitable promoters for expressing RNA from a plasmid include, e.g., the U6 or H1 RNA pol III promoter sequences, or the cytomegalovirus promoters. Selection of other suitable promoters is within the skill in the art. The recombinant plasmids of the invention can also comprise inducible or regulatable promoters for expression of the miR gene products in cancer cells.

The miR gene products that are expressed from recombinant plasmids can be isolated from cultured cell expression systems by standard techniques. The miR gene products that are expressed from recombinant plasmids can also be delivered to, and expressed directly in, the cancer cells. The use of recombinant plasmids to deliver the miR gene products to cancer cells is discussed in more detail below.

The miR gene products can be expressed from a separate recombinant plasmid, or they can be expressed from the same recombinant plasmid. In one embodiment, the miR gene products are expressed as RNA precursor molecules from a single plasmid, and the precursor molecules are processed into the functional miR gene product by a suitable processing system, including, but not limited to, processing systems extant within a cancer cell. Other suitable processing systems include, e.g., the *in vitro* Drosophila cell lysate system (e.g., as described in U.S. Published Patent Application No. 2002/0086356 to Tuschl *et al.*) and the *E. Coli* RNAse III system (e.g., as described in U.S. Published Patent Application No. 2004/0014113 to Yang *et al.*).

Selection of plasmids suitable for expressing the miR gene products, methods for inserting nucleic acid sequences into the plasmid to express the gene products, and methods of delivering the recombinant plasmid to the cells of interest are within the skill in the art. See, for example, Zeng *et al.* (2002), *Molecular Cell* 9:1327-1333; Tuschl (2002), *Nat. Biotechnol*, 20:446-448; Brummelkamp *et al.* (2002), *Science* 296:550-553; Miyagishi *et al.* (2002), *Nat. Biotechnol.* 20:497-500; Paddison *et al.* (2002), *Genes Dev.* 16:948-958; Lee *et al.* (2002), *Nat. Biotechnol.* 20:500-505; and Paul *et al.* (2002), *Nat. Biotechnol.* 20:505-508.

In one embodiment, a plasmid expressing the miR gene products comprises a sequence encoding a miR precursor RNA under the control of the CMV intermediate-early promoter. As used herein, "under the control" of a promoter means that the nucleic acid sequences encoding the miR gene product are located 3' of the promoter, so that the promoter can initiate transcription of the miR gene product coding sequences.

The miR gene products can also be expressed from recombinant viral vectors. It is contemplated that the miR gene products can be expressed from two separate recombinant viral vectors, or from the same viral vector. The RNA expressed from the recombinant viral vectors can either be isolated from cultured cell expression systems by standard techniques, or can be expressed directly in cancer cells. The use of recombinant viral vectors to deliver the miR gene products to cancer cells is discussed in more detail below.

The recombinant viral vectors of the invention comprise sequences encoding the miR gene products and any suitable promoter for expressing the RNA sequences. Suitable promoters include, but are not limited to, the U6 or H1 RNA pol III promoter sequences, or the cytomegalovirus promoters. Selection of other suitable promoters is within the skill in the art. The recombinant viral vectors of the invention can also comprise inducible or regulatable promoters for expression of the miR gene products in a cancer cell.

Any viral vector capable of accepting the coding sequences for the miR gene products can be used; for example, vectors derived from adenovirus (AV); adeno-associated virus (AAV); retroviruses (*e.g.*, lentiviruses (LV), Rhabdoviruses, murine leukemia virus); herpes virus, and the like. The tropism of the viral vectors can be modified by pseudotyping the vectors with envelope proteins or other surface antigens from other viruses, or by substituting different viral capsid proteins, as appropriate.

For example, lentiviral vectors of the invention can be pseudotyped with surface proteins from vesicular stomatitis virus (VSV), rabies, Ebola, Mokola, and the like. AAV vectors of the invention can be made to target different cells by engineering the vectors to express different capsid protein serotypes. For example, an AAV vector expressing a serotype 2 capsid on a serotype 2 genome is called AAV 2/2. This serotype 2 capsid gene in the AAV 2/2 vector can be replaced by a serotype 5 capsid gene to produce an AAV 2/5 vector. Techniques for constructing AAV vectors that express different capsid protein serotypes are within the skill in the art; *see, e.g.*, Rabinowitz, J.E., *et al.* (2002), *J. Virol.* 76:791-801.

Selection of recombinant viral vectors suitable for use in the invention, methods for inserting nucleic acid sequences for expressing RNA into the vector, methods of

delivering the viral vector to the cells of interest, and recovery of the expressed RNA products are within the skill in the art. See, for example, Dornburg (1995), *Gene Therapy* 2:301-310; Eglitis (1988), *Biotechniques* 6:608-614; Miller (1990), *Hum. Gene Therapy* 1:5-14; and Anderson (1998), *Nature* 392:25-30.

Particularly suitable viral vectors are those derived from AV and AAV. A suitable AV vector for expressing the miR gene products, a method for constructing the recombinant AV vector, and a method for delivering the vector into target cells, are described in Xia *et al.* (2002), *Nat. Biotech.* 20:1006-1010. Suitable AAV vectors for expressing the miR gene products, methods for constructing the recombinant AAV vector, and methods for delivering the vectors into target cells are described in Samulski *et al.* (1987),J. *Virol.* 61:3096-3101; Fisher *et al.* (1996), *J. Virol*, 70:520-532; Samulski *et al.* (1989), *J. Virol.* 63:3822-3826; U.S. Pat. No. 5,252,479; U.S. Pat. No. 5,139,941; International Patent Application No. WO 94/13788; and International Patent Application No. WO 93/24641. In one embodiment, the miR gene products are expressed from a single recombinant AAV vector comprising the CMV intermediate early promoter.

In a certain embodiment, a recombinant AAV viral vector of the invention comprises a nucleic acid sequence encoding a miR precursor RNA in operable connection with a polyT termination sequence under the control of a human U6 RNA promoter. As used herein, "in operable connection with a polyT termination sequence" means that the nucleic acid sequences encoding the sense or antisense strands are immediately adjacent to the polyT termination signal in the 5' direction. During transcription of the miR sequences from the vector, the polyT termination signals act to terminate transcription.

In other embodiments of the treatment methods of the invention, an effective amount of at least one compound that inhibits miR expression can be administered to the subject. As used herein, "inhibiting miR expression" means that the production of the precursor and/or active, mature form of miR gene product after treatment is less than the amount produced prior to treatment. One skilled in the art can readily determine whether miR expression has been inhibited in a cancer cell, using, for

example, the techniques for determining miR transcript level discussed above for the diagnostic method. Inhibition can occur at the level of gene expression (i.e., by inhibiting transcription of a miR gene encoding the miR gene product) or at the level of processing (e.g., by inhibiting processing of a miR precursor into a mature, active miR).

As used herein, an "effective amount" of a compound that inhibits miR expression is an amount sufficient to inhibit proliferation of a cancer cell in a subject suffering from a cancer (e.g., a solid cancer). One skilled in the art can readily determine an effective amount of a miR expression-inhibition compound to be administered to a given subject, by taking into account factors, such as the size and weight of the subject; the extent of disease penetration; the age, health and sex of the subject; the route of administration; and whether the administration is regional or systemic.

For example, an effective amount of the expression-inhibition compound can be based on the approximate weight of a tumor mass to be treated, as described herein. An effective amount of a compound that inhibits miR expression can also be based on the approximate or estimated body weight of a subject to be treated, as described herein.

One skilled in the art can also readily determine an appropriate dosage regimen for administering a compound that inhibits miR expression to a given subject.

Suitable compounds for inhibiting miR gene expression include double-stranded RNA (such as short- or small-interfering RNA or "siRNA"), antisense nucleic acids, and enzymatic RNA molecules, such as ribozymes. Each of these compounds can be targeted to a given miR gene product and interfere with the expression of (e.g., inhibit translation of, induce cleavage or destruction of) the target miR gene product.

For example, expression of a given miR gene can be inhibited by inducing RNA interference of the miR gene with an isolated double-stranded RNA ("dsRNA") molecule which has at least 90%, for example at least 95%, at least 98%, at least 99%, or 100%, sequence homology with at least a portion of the miR gene product. In a particular embodiment, the dsRNA molecule is a "short or small interfering RNA" or "siRNA."

siRNA useful in the present methods comprise short double-stranded RNA from about 17 nucleotides to about 29 nucleotides in length, preferably from about 19 to about 25 nucleotides in length. The siRNA comprise a sense RNA strand and a complementary antisense RNA strand annealed together by standard Watson-Crick

base-pairing interactions (hereinafter "base-paired"). The sense strand comprises a nucleic acid sequence that is substantially identical to a nucleic acid sequence contained within the target miR gene product.

As used herein, a nucleic acid sequence in an siRNA which is "substantially identical" to a target sequence contained within the target mRNA is a nucleic acid sequence that is identical to the target sequence, or that differs from the target sequence by one or two nucleotides. The sense and antisense strands of the siRNA can comprise two complementary, single-stranded RNA molecules, or can comprise a single molecule in which two complementary portions are base-paired and are covalently linked by a single-stranded "hairpin" area.

The siRNA can also be altered RNA that differs from naturally-occurring RNA by the addition, deletion, substitution and/or alteration of one or more nucleotides. Such alterations can include addition of non-nucleotide material, such as to the end(s) of the siRNA or to one or more internal nucleotides of the siRNA, or modifications that make the siRNA resistant to nuclease digestion, or the substitution of one or more nucleotides in the siRNA with deoxyribonucleotides.

One or both strands of the siRNA can also comprise a 3' overhang. As used herein, a "3' overhang" refers to at least one unpaired nucleotide extending from the 3'-end of a duplexed RNA strand. Thus, in certain embodiments, the siRNA comprises at least one 3' overhang of from 1 to about 6 nucleotides (which includes ribonucleotides or deoxyribonucleotides) in length, from 1 to about 5 nucleotides in length, from 1 to about 4 nucleotides in length, or from about 2 to about 4 nucleotides in length. In a particular embodiment, the 3' overhang is present on both strands of the siRNA, and is 2 nucleotides in length. For example, each strand of the siRNA can comprise 3' overhangs of dithymidylic acid ("TT") or diuridylic acid ("uu").

The siRNA can be produced chemically or biologically, or can be expressed from a recombinant plasmid or viral vector, as described above for the isolated miR gene products. Exemplary methods for producing and testing dsRNA or siRNA molecules are described in U.S. Published Patent Application No. 2002/0173478 to Gewirtz and in U.S. Published Patent Application No. 2004/0018176 to *Reich et al.*

Expression of a given miR gene can also be inhibited by an antisense nucleic acid. As used herein, an "antisense nucleic acid" refers to a nucleic acid molecule that binds to target RNA by means of RNA-RNA, RNA-DNA or RNA-peptide nucleic acid interactions, which alters the activity of the target RNA. Antisense nucleic acids suitable for use in the present methods are single-stranded nucleic acids (*e.g.*, RNA, DNA, RNA-DNA chimeras, peptide nucleic acid (PNA)) that generally comprise a nucleic acid sequence complementary to a contiguous nucleic acid sequence in a miR gene product. The antisense nucleic acid can comprise a nucleic acid sequence that is 50-100% complementary, 75-100% complementary, or 95-100% complementary to a contiguous nucleic acid sequence in a miR gene product. Nucleic acid sequences for the miR gene products are provided in Tables 1a and 1b. Without wishing to be bound by any theory, it is believed that the antisense nucleic acids activate RNase H or another cellular nuclease that digests the miR gene product/antisense nucleic acid duplex.

Antisense nucleic acids can also contain modifications to the nucleic acid backbone or to the sugar and base moieties (or their equivalent) to enhance target specificity, nuclease resistance, delivery or other properties related to efficacy of the molecule. Such modifications include cholesterol moieties, duplex intercalators, such as acridine, or one or more nuclease-resistant groups.

Antisense nucleic acids can be produced chemically or biologically, or can be expressed from a recombinant plasmid or viral vector, as described above for the isolated miR gene products. Exemplary methods for producing and testing are within the skill in the art; *see, e.g.*, Stein and Cheng (1993), *Science* 261:1004 and U.S. Pat. No. 5,849,902 to Woolf *et al.*

Expression of a given miR gene can also be inhibited by an enzymatic nucleic acid. As used herein, an "enzymatic nucleic acid" refers to a nucleic acid comprising a substrate binding region that has complementarity to a contiguous nucleic acid sequence of a miR gene product, and which is able to specifically cleave the miR gene product. The enzymatic nucleic acid substrate binding region can be, for example, 50-100% complementary, 75-100% complementary, or 95-100% complementary to a contiguous nucleic acid sequence in a miR gene product. The enzymatic nucleic acids

can also comprise modifications at the base, sugar, and/or phosphate groups. An exemplary enzymatic nucleic acid for use in the present methods is a ribozyme.

The enzymatic nucleic acids can be produced chemically or biologically, or can be expressed from a recombinant plasmid or viral vector, as described above for the isolated miR gene products. Exemplary methods for producing and testing dsRNA or siRNA molecules are described in Werner and Uhlenbeck (1995), *Nucl. Acids Res.* 23:2092-96; Hammann *et al.* (1999), *Antisense and Nucleic Acid Drug Dev.* 9:25-31; and U.S. Pat. No. 4,987,071 to Cech *et al.*

Administration of at least one miR gene product, or at least one compound for inhibiting miR expression, will inhibit the proliferation of cancer cells in a subject who has a solid cancer. As used herein, to "inhibit the proliferation of a cancer cell" means to kill the cell, or permanently or temporarily arrest or slow the growth of the cell. Inhibition of cancer cell proliferation can be inferred if the number of such cells in the subject remains constant or decreases after administration of the miR gene products or miR gene expression-inhibition compounds. An inhibition of cancer cell proliferation can also be inferred if the absolute number of such cells increases, but the rate of tumor growth decreases.

The number of cancer cells in the body of a subject can be determined by direct measurement, or by estimation from the size of primary or metastatic tumor masses. For example, the number of cancer cells in a subject can be measured by immunohistological methods, flow cytometry, or other techniques designed to detect characteristic surface markers of cancer cells.

The size of a tumor mass can be ascertained by direct visual observation, or by diagnostic imaging methods, such as X-ray, magnetic resonance imaging, ultrasound, and scintigraphy. Diagnostic imaging methods used to ascertain size of the tumor mass can be employed with or without contrast agents, as is known in the art. The size of a tumor mass can also be ascertained by physical means, such as palpation of the tissue mass or measurement of the tissue mass with a measuring instrument, such as a caliper.

The miR gene products or miR gene expression-inhibition compounds can be administered to a subject by any means suitable for delivering these compounds to cancer cells of the subject. For example, the miR gene products or miR expression-

inhibition compounds can be administered by methods suitable to transfect cells of the subject with these compounds, or with nucleic acids comprising sequences encoding these compounds. In one embodiment, the cells are transfected with a plasmid or viral vector comprising sequences encoding at least one miR gene product or miR gene expression-inhibition compound.

Transfection methods for eukaryotic cells are well known in the art, and include, e.g., direct injection of the nucleic acid into the nucleus or pronucleus of a cell; electroporation; liposome transfer or transfer mediated by lipophilic materials; receptor-mediated nucleic acid delivery, bioballistic or particle acceleration; calcium phosphate precipitation, and transfection mediated by viral vectors.

For example, cells can be transfected with a liposomal transfer compound, e.g., DOTAP (N-[1-(2,3-dioleoyloxy)propyl]-N,N,N-trimethyl-ammonium methylsulfate, Boehringer-Mannheim) or an equivalent, such as LIPOFECTIN. The amount of nucleic acid used is not critical to the practice of the invention; acceptable results may be achieved with 0.1-100 micrograms of nucleic acid/10^5 cells. For example, a ratio of about 0.5 micrograms of plasmid vector in 3 micrograms of DOTAP per 10^5 cells can be used.

A miR gene product or miR gene expression-inhibition compound can also be administered to a subject by any suitable enteral or parenteral administration route. Suitable enteral administration routes for the present methods include, e.g., oral, rectal, or intranasal delivery. Suitable parenteral administration routes include, e.g., intravascular administration (e.g., intravenous bolus injection, intravenous infusion, intra-arterial bolus injection, intra-arterial infusion and catheter instillation into the vasculature); peri- and intra-tissue injection (e.g., peri-tumoral and intra-tumoral injection, intra-retinal injection, or subretinal injection); subcutaneous injection or deposition, including subcutaneous infusion (such as by osmotic pumps); direct application to the tissue of interest, for example by a catheter or other placement device (e.g., a retinal pellet or a suppository or an implant comprising a porous, non-porous, or gelatinous material); and inhalation. Particularly suitable administration routes are injection, infusion and direct injection into the tumor.

In the present methods, a miR gene product or miR gene product expression-inhibition compound can be administered to the subject either as naked RNA, in

combination with a delivery reagent, or as a nucleic acid (*e.g.*, a recombinant plasmid or viral vector) comprising sequences that express the miR gene product or miR gene product expression-inhibition compound. Suitable delivery reagents include, e.g., the Mirus Transit TKO lipophilic reagent; lipofectin; lipofectamine; cellfectin; polycations (*e.g.*, polylysine), and liposomes.

Recombinant plasmids and viral vectors comprising sequences that express the miR gene products or miR gene expression-inhibition compounds, and techniques for delivering such plasmids and vectors to cancer cells, are discussed herein and/or are well known in the art.

In a particular embodiment, liposomes are used to deliver a miR gene product or miR gene expression-inhibition compound (or nucleic acids comprising sequences encoding them) to a subject. Liposomes can also increase the blood half-life of the gene products or nucleic acids. Suitable liposomes for use in the invention can be formed from standard vesicle-forming lipids, which generally include neutral or negatively charged phospholipids and a sterol, such as cholesterol. The selection of lipids is generally guided by consideration of factors, such as the desired liposome size and half-life of the liposomes in the blood stream. A variety of methods are known for preparing liposomes, for example, as described in Szoka *et al.* (1980), *Ann. Rev. Biophys. Bioeng.* 9:467; and U.S. Pat. Nos. 4,235,871, 4,501,728, 4,837,028, and 5,019,369.

The liposomes for use in the present methods can comprise a ligand molecule that targets the liposome to cancer cells. Ligands that bind to receptors prevalent in cancer cells, such as monoclonal antibodies that bind to tumor cell antigens, are preferred.

The liposomes for use in the present methods can also be modified so as to avoid clearance by the mononuclear macrophage system ("MMS") and reticuloendothelial system ("RES"). Such modified liposomes have opsonization-inhibition moieties on the surface or incorporated into the liposome structure. In a particularly preferred embodiment, a liposome of the invention can comprise both an opsonization-inhibition moiety and a ligand.

Opsonization-inhibiting moieties for use in preparing the liposomes of the invention are typically large hydrophilic polymers that are bound to the liposome

membrane. As used herein, an opsonization-inhibiting moiety is "bound" to a liposome membrane when it is chemically or physically attached to the membrane, e.g., by the intercalation of a lipid-soluble anchor into the membrane itself, or by binding directly to active groups of membrane lipids. These opsonization-inhibiting hydrophilic polymers form a protective surface layer that significantly decreases the uptake of the liposomes by the MMS and RES; e.g., as described in U.S. Pat. No. 4,920,016.

Opsonization-inhibiting moieties suitable for modifying liposomes are preferably water-soluble polymers with a number-average molecular weight from about 500 to about 40,000 daltons, and more preferably from about 2,000 to about 20,000 daltons. Such polymers include polyethylene glycol (PEG) or polypropylene glycol (PPG) derivatives; e.g., methoxy PEG or PPG, and PEG or PPG stearate; synthetic polymers, such as polyacrylamide or poly N-vinyl pyrrolidone; linear, branched, or dendrimeric polyamidoamines; polyacrylic acids; polyalcohols, e.g., polyvinylalcohol and polyxylitol to which carboxylic or amino groups are chemically linked, as well as gangliosides, such as ganglioside GM1. Copolymers of PEG, methoxy PEG, or methoxy PPG, or derivatives thereof, are also suitable. In addition, the opsonization-inhibiting polymer can be a block copolymer of PEG and either a polyamino acid, polysaccharide, polyamidoamine, polyethyleneamine, or polynucleotide. The opsonization-inhibiting polymers can also be natural polysaccharides containing amino acids or carboxylic acids, e.g., galacturonic acid, glucuronic acid, mannuronic acid, hyaluronic acid, pectic acid, neuraminic acid, alginic acid, carrageenan; aminated polysaccharides or oligosaccharides (linear or branched); or carboxylated polysaccharides or oligosaccharides, e.g., reacted with derivatives of carbonic acids with resultant linking of carboxylic groups. Preferably, the opsonization-inhibiting moiety is a PEG, PPG, or a derivative thereof. Liposomes modified with PEG or PEG-derivatives are sometimes called "PEGylated liposomes."

The opsonization-inhibiting moiety can be bound to the liposome membrane by any one of numerous well-known techniques. For example, an N-hydroxysuccinimide ester of PEG can be bound to a phosphatidyl-ethanolamine lipid-soluble anchor, and then bound to a membrane. Similarly, a dextran polymer can be derivatized with a

stearylamine lipid-soluble anchor via reductive amination using $Na(CN)BH_3$ and a solvent mixture, such as tetrahydrofuran and water in a 30:12 ratio at 60 °C.

Liposomes modified with opsonization-inhibition moieties remain in the circulation much longer than unmodified liposomes. For this reason, such liposomes are sometimes called "stealth" liposomes. Stealth liposomes are known to accumulate in tissues fed by porous or "leaky" microvasculature. Thus, tissue characterized by such microvasculature defects, for example, solid tumors, will efficiently accumulate these liposomes; see Gabizon, *et al.* (1988), Proc. Natl. Acad. Sci., U.S.A., 18:6949-53. In addition, the reduced uptake by the RES lowers the toxicity of stealth liposomes by preventing significant accumulation of the liposomes in the liver and spleen. Thus, liposomes that are modified with opsonization-inhibition moieties are particularly suited to deliver the miR gene products or miR gene expression-inhibition compounds (or nucleic acids comprising sequences encoding them) to tumor cells.

The miR gene products or miR gene expression-inhibition compounds can be formulated as pharmaceutical compositions, sometimes called "medicaments," prior to administering them to a subject, according to techniques known in the art. Accordingly, the invention encompasses pharmaceutical compositions for treating a solid cancer. In one embodiment, the pharmaceutical composition comprises at least one isolated miR gene product, or an isolated variant or biologically-active fragment thereof, and a pharmaceutically-acceptable carrier. In a particular embodiment, the at least one miR gene product corresponds to a miR gene product that has a decreased level of expression in solid cancer cells relative to suitable control cells. In certain embodiments the isolated miR gene product is selected from the group consisting of miR-145, miR-155, miR-218-2 combinations thereof.

In other embodiments, the pharmaceutical compositions of the invention comprise at least one miR expression-inhibition compound. In a particular embodiment, the at least one miR gene expression-inhibition compound is specific for a miR gene whose expression is greater in solid cancer cells than control cells. In certain embodiments, the miR gene expression-inhibition compound is specific for one or more miR gene products selected from the group consisting of miR-21, miR-17-5p, miR-191, miR-29b-2, miR-223, miR-128b, miR-199a-1, miR-24-1, miR-24-2, miR-146,

miR-155, miR-181b-1, miR-20a, miR-107, miR-32, miR-92-2, miR-214, miR-30c, miR-25, miR-221, miR-106a and combinations thereof.

Pharmaceutical compositions of the present invention are characterized as being at least sterile and pyrogen-free. As used herein, "pharmaceutical compositions" include formulations for human and veterinary use. Methods for preparing pharmaceutical compositions of the invention are within the skill in the art, for example as described in Remington's Pharmaceutical Science, 17th ed., Mack Publishing Company, Easton, Pa. (1985).

The present pharmaceutical compositions comprise at least one miR gene product or miR gene expression-inhibition compound (or at least one nucleic acid comprising sequences encoding them) (e.g., 0.1 to 90% by weight), or a physiologically-acceptable salt thereof, mixed with a pharmaceutically-acceptable carrier. In certain embodiments, the pharmaceutical compositions of the invention additionally comprise one or more anti-cancer agents (e.g., chemotherapeutic agents).The pharmaceutical formulations of the invention can also comprise at least one miR gene product or miR gene expression-inhibition compound (or at least one nucleic acid comprising sequences encoding them), which are encapsulated by liposomes and a pharmaceutically-acceptable carrier. In one embodiment, the pharmaceutical composition comprises a miR gene or gene product that is not miR-15 and/or miR-16.

Especially suitable pharmaceutically-acceptable carriers are water, buffered water, normal saline, 0.4% saline, 0.3% glycine, hyaluronic acid and the like.

In a particular embodiment, the pharmaceutical compositions of the invention comprise at least one miR gene product or miR gene expression-inhibition compound (or at least one nucleic acid comprising sequences encoding them) that is resistant to degradation by nucleases. One skilled in the art can readily synthesize nucleic acids that are nuclease resistant, for example, by incorporating one or more ribonucleotides that is modified at the 2'-position into the miR gene product. Suitable 2'-modified ribonucleotides include those modified at the 2'-position with fluoro, amino, alkyl, alkoxy, and O-allyl.

Pharmaceutical compositions of the invention can also comprise conventional pharmaceutical excipients and/or additives. Suitable pharmaceutical excipients include

stabilizers, antioxidants, osmolality adjusting agents, buffers, and pH adjusting agents. Suitable additives include, e.g., physiologically biocompatible buffers (e.g., tromethamine hydrochloride), additions of chelants (such as, for example, DTPA or DTPA-bisamide) or calcium chelate complexes (such as, for example, calcium DTPA, CaNaDTPA-bisamide), or, optionally, additions of calcium or sodium salts (for example, calcium chloride, calcium ascorbate, calcium gluconate or calcium lactate). Pharmaceutical compositions of the invention can be packaged for use in liquid form, or can be lyophilized.

For solid pharmaceutical compositions of the invention, conventional nontoxic solid pharmaceutically-acceptable carriers can be used; for example, pharmaceutical grades of mannitol, lactose, starch, magnesium stearate, sodium saccharin, talcum, cellulose, glucose, sucrose, magnesium carbonate, and the like.

For example, a solid pharmaceutical composition for oral administration can comprise any of the carriers and excipients listed above and 10-95%, preferably 25%-75%, of the at least one miR gene product or miR gene expression-inhibition compound (or at least one nucleic acid comprising sequences encoding them). A pharmaceutical composition for aerosol (inhalational) administration can comprise 0.01-20% by weight, preferably 1%-10% by weight, of the at least one miR gene product or miR gene expression-inhibition compound (or at least one nucleic acid comprising sequences encoding them) encapsulated in a liposome as described above, and a propellant. A carrier can also be included as desired; e.g., lecithin for intranasal delivery.

The pharmaceutical compositions of the invention can further comprise one or more anti-cancer agents. In a particular embodiment, the compositions comprise at least one miR gene product or miR gene expression-inhibition compound (or at least one nucleic acid comprising sequences encoding them) and at least one chemotherapeutic agent. Chemotherapeutic agents that are suitable for the methods of the invention include, but are not limited to, DNA-alkylating agents, anti-tumor antibiotic agents, anti-metabolic agents, tubulin stabilizing agents, tubulin destabilizing agents, hormone antagonist agents, topoisomerase inhibitors, protein kinase inhibitors, HMG-CoA inhibitors, CDK inhibitors, cyclin inhibitors, caspase inhibitors, metalloproteinase inhibitors, antisense nucleic acids, triple-helix DNAs, nucleic acid

aptamers, and molecularly-modified viral, bacterial and exotoxic agents. Examples of suitable agents for the compositions of the present invention include, but are not limited to, cytidine arabinoside, methotrexate, vincristine, etoposide (VP-16), doxorubicin (adriamycin), cisplatin (CDDP), dexamethasone, arglabin, cyclophosphamide, sarcolysin, methylnitrosourea, fluorouracil, 5-fluorouracil (5FU), vinblastine, camptothecin, actinomycin-D, mitomycin C, hydrogen peroxide, oxaliplatin, irinotecan, topotecan, leucovorin, carmustine, streptozocin, CPT-11, taxol, tamoxifen, dacarbazine, rituximab, daunorubicin, 1-β-D-arabinofuranosylcytosine, imatinib, fludarabine, docetaxel, FOLFOX4.

The invention also encompasses methods of identifying an inhibitor of tumorigenesis, comprising providing a test agent to a cell and measuring the level of at least one miR gene product in the cell. In one embodiment, the method comprises providing a test agent to a cell and measuring the level of at least one miR gene product associated with decreased expression levels in cancer cells. An increase in the level of the miR gene product in the cell after the agent is provided, relative to a suitable control cell (e.g., agent is not provided), is indicative of the test agent being an inhibitor of tumorigenesis. In a particular embodiment, at least one miR gene product associated with decreased expression levels in cancer cells is selected from the group consisting of miR-145, miR-155, miR-218-2 and combinations thereof.

In other embodiments the method comprises providing a test agent to a cell and measuring the level of at least one miR gene product associated with increased expression levels in cancer cells. A decrease in the level of the miR gene product in the cell after the agent is provided, relative to a suitable control cell(e.g., agent is not provided), is indicative of the test agent being an inhibitor of tumorigenesis. In a particular embodiment, at least one miR gene product associated with increased expression levels in cancer cells is selected from the group consisting of miR-21, miR-17-5p, miR-191, miR-29b-2, miR-223, miR-128b, miR-199a-1, miR-24-1, miR-24-2, miR-146, miR-155, miR-181b-1, miR-20a, miR-107, miR-32, miR-92-2, miR-214, miR-30c, miR-25, miR-221, miR-106a.

Suitable agents include, but are not limited to drugs (e.g., small molecules, peptides), and biological macromolecules (e.g., proteins, nucleic acids). The agent can be produced recombinantly, synthetically, or it may be isolated (i.e., purified) from a

natural source. Various methods for providing such agents to a cell (e.g., transfection) are well known in the art, and several of such methods are described hereinabove. Methods for detecting the expression of at least one miR gene product (e.g., Northern blotting, *in situ* hybridization, RT-PCR, expression profiling) are also well known in the art. Several of these methods are also described hereinabove.

The invention will now be illustrated by the following non-limiting examples.

EXEMPLIFICATION

The following Materials and Methods were used in the Examples:

Samples

A total of 540 samples, including 363 primary tumor samples and 177 normal tissues, were used in this study (Table 2). The following solid cancers were represented: lung carcinoma, breast carcinoma, prostate carcinoma, stomach carcinoma, colon carcinoma and pancreatic endocrine tumors. All samples were obtained with informed consent from each patient and were confirmed histologically. Normal samples were paired with samples from individuals affected with lung and stomach carcinoma, and from normal individuals for the remaining tissues. All normal breast samples were obtained by pooling 5 unrelated normal tissues. Total RNA was isolated from tissues using TRIzol™ reagent (Invitrogen), according to manufacturer's instructions.

MicroRNA microarrays.

Microarray analysis was performed as previously described (Liu, C.-G., et al., *Proc. Natl. Acad. Sci. USA 101*: 11755-11760 (2004)). Briefly, 5 μg of total RNA was used for hybridization on miRNA microarray chips. These chips contain gene-specific 40-mer oligonucleotide probes, spotted by contacting technologies and covalently attached to a polymeric matrix. The microarrays were hybridized in 6× SSPE (0.9 M NaCl/60 mM $NaH_2PO_4 \cdot H_2O$/8 mM EDTA, pH 7.4)/30% formamide at 25°C for 18 hr, washed in 0.75× TNT (Tris-HCl/NaCl/ Tween 20) at 37°C for 40 min, and processed using direct detection of the biotin-labeled transcripts by streptavidin-Alexa647 (Molecular Probes) conjugate. Processed slides were scanned using a microarray scanner (GenePix Pro, Axon), with the laser set to 635 nm, at fixed PMT setting and a scan resolution of 10 mm. The data were confirmed by Northern blotting as described (Calin, G.A., et al., *Proc. Natl. Acad. Sci. USA 101*:11755-11760 (2004); Iorio, M.V., et al., *Cancer Res. 65*: 7065-7070 (2005)).

Table 2. Samples used in the study (tumors and corresponding normals).

Tumour type	Cancer Samples	Normal Samples
Lung carcinoma	123	123
Breast carcinoma	79	6*
Colon carcinoma	46	8
Gastric carcinoma	20	21
Endocrine pancreatic tumours	39	12
Prostate cancer	56	7
All tissues (527)	363	177

* Pools of 5 unrelated normal breast tissues per sample (for a total of 30 unrelated individuals).

Computational analysis.

Microarray images were analyzed using GenePix Pro (Axon). Average values of the replicate spots of each miRNA were background-subtracted, normalized and subjected to further analysis. Normalization was performed by using a per chip median normalization method, using the median array as a reference. Finally, miRNAs measured as present in at least the smallest of the two classes in a dataset were selected. Absent calls were thresholded to 4.5 prior to statistical analysis. This level is the average minimum intensity level detected in the experiments. MicroRNA nomenclature was according to the Genome Browser and the microRNA database at Sanger Center (Griffiths-Jones, S., *Nucleic Acids Res 32:* D109-11(2004)); in case of discrepancies we followed the microRNA database. Differentially-expressed microRNAs were identified by using the t test procedure within significance analysis of microarrays (SAM) (Tusher, V. G., et al., *Proc Natl Acad Sci U S A 98:* 5116-21 (2001). SAM calculates a score for each gene on the basis of the change in expression relative to the standard deviation of all measurements. Within SAM, t test was used.

The microRNA signatures were determined by applying nearest shrunken centroids method. This method identifies a subgroup of genes that best characterizes each solid cancer from its respective normal counterpart. The prediction error was calculated by means of 10-fold cross validation, and for each cancer, we obtained the miR signature that resulted in the minimal prediction error. A resampling test was performed by random permutation analysis to compute the p-value of the shared signature.

Example 1: Identification of a microRNA expression signature in human solid cancers

Statistics

The combined cancers/normal tissue comparison was conducted using a reduced number of lung samples (80 cancer and 40 normal samples), in order to balance the different tissues numerically, yielding a total of 404 samples. For statistical analysis, 137 miRs, whose expression values were above 256 (threshold value) in at least 50% of the samples, were retained from the 228 that were measured. A T test was used to identify differentially-expressed microRNAs (Table 3). The p-values of the T test were corrected for multiple testing procedures and to control Type I error rates. Adjusted p–values were obtained by performing resampling with 500,000 permutations (Jung, S. H., et al. *Biostatistics 6*: 157-69 (2005)). This analysis was performed in order to evaluate the results by using the same method as Lu and co-workers (Lu, J., et al., *Nature 435*: 834-8(2005)).

As an alternative to T test, significance analysis of microarrays (SAM) was used to identify differentially-expressed microRNAs. This procedure allows for the control of false detection rate (FDR). The delta was chosen to result in an FDR less than or equal to 0.01. microRNA subsets which result in the best tumor classification, *i.e.*, which best predict the two classes (cancer and normal), were then identified using the method of the nearest shrunken centroids, as implemented in PAM (prediction analysis of microarray). The prediction error was calculated by means of 10-fold cross validation. The microRNAs were selected yielding the minimum misclassification error after cross-validation.

Results

By T-test, 43 differentially-expressed miRs with an adjusted p-value below 0.05 were obtained (Table 3). Twenty six miRs were overexpressed and 17 were under-expressed relative to corresponding normal tissues when the six solid cancers are

grouped together (breast, colon, lung, pancreas, prostate, stomach). These results indicated that the spectrum of expressed miRNAs in solid cancers is very different from that of normal cells (43 out of 137 miRNAs, 31%). Using SAM, 49 miRNAs were identified as differentially-expressed, of which 34 were up-regulated (Table 4). Using PAM, 36 over-expressed miRNAs in cancer (indicated by positive cancer scores) and 21 down-regulated miRs (indicated by negative cancer scores) were identified as differentially-expressed (Table 5). However, these analyses are not tailored to identify alterations in miR expression that consistently result in transformation, because miR expression is heavily tissue-specific (He, L., et al. *Nature 435*: 828-833 (2005); also see FIG. 1 and FIG. 2).

The clustering of miRs based on expression profiles derived from 363 solid cancer and 177 normal samples using 228 miRs is shown in FIG. 1. The tree, which shows a very good separation between the different tissues, was constructed using 137 different miRNAs that were expressed in at least 50% of the samples used in the study.

Table 3. Differentially regulated miRs in 6 solid cancer types vs. normal tissues (T test stats.) *.

miR	ID	Cancer Mean	Normal Mean	Test stat	Raw p	Adj p
miR-21	#47	11.538663	9.648338	7.861136	2.00E-06	2.00E-06
miR-141	#137	9.024091	7.905398	6.238014	2.00E-06	2.00E-06
miR-212	#208	13.540651	14.33617	-6.57942	2.00E-06	2.00E-06
miR-128a prec	#113	12.32588	13.522675	-6.76388	2.00E-06	2.00E-06
miR-138-2	#133	11.739557	13.144746	-7.01204	2.00E-06	2.00E-06
miR-218-2	#221	11.279787	12.539366	-7.40557	2.00E-06	2.00E-06
miR-23b	#51	14.169748	15.949736	-8.37744	2.00E-06	2.00E-06
miR-195	#184	10.343991	9.172985	5.763262	2.00E-06	1.00E-05
miR-212 prec	#209	12.686966	13.661763	-5.83132	4.00E-06	1.00E-05
miR-29b-2	#95	11.27556	9.940731	5.660854	2.00E-06	1.40E-05
miR-199a-1	#191	10.032008	8.920183	5.528849	2.00E-06	3.00E-05
miR-9-3	#28	11.461922	12.570412	-5.43006	2.00E-06	4.60E-05
miR-128a	#114	13.024235	13.856624	-5.35102	6.00E-06	7.20E-05
let-7a-1	#1	12.616569	13.455246	-5.35346	2.00E-06	7.20E-05
let-7b	#5	13.42636	14.068521	-5.17701	1.00E-05	0.000146
miR-16-2	#39	10.460707	9.305895	5.048375	4.00E-06	0.000224
miR-199a-2	#192	9.714225	8.759237	4.862553	1.00E-05	0.000494
miR-152 prec	#151	11.388676	12.357529	-4.83716	2.00E-06	0.00053
miR-16-1	#38	10.443169	9.338182	4.755258	1.00E-05	0.00071
miR-30d	#72	13.982017	14.775206	-4.5707	1.20E-05	0.001476
miR-34a	#78	10.675566	9.63769	4.467301	2.60E-05	0.00217
miR-17-5p	#41	11.567244	10.281468	4.341834	3.80E-05	0.0034
miR-128b	#115	10.930395	9.947746	4.304764	3.80E-05	0.003912
miR-20a	#46	11.409852	10.19284	4.304678	3.20E-05	0.003912
miR-181b-1 prec	#211	9.577504	8.804294	4.285968	4.80E-05	0.004126
miR-132	#121	9.599947	8.775966	4.284737	5.60E-05	0.004126
miR-200b	#195	9.475221	8.527243	4.221511	4.00E-05	0.0052
let-7a-3	#4	10.436089	9.511546	4.08952	0.000104	0.008242
miR-138-1	#132	8.299613	9.200253	-4.05204	5.60E-05	0.00931
miR-29c	#65	11.291005	10.326912	4.019385	0.000144	0.010312
miR-29a	#62	11.381359	10.461075	4.013697	0.00015	0.010398
miR-96	#86	11.37218	12.136636	-3.94825	0.000138	0.012962
miR-191	#177	13.498207	12.729872	3.817228	0.000158	0.02015
miR-27a	#59	10.399338	9.548582	3.715048	0.000344	0.028096
let-7g	#15	10.819688	10.01157	3.653239	0.000426	0.033874
miR-9-1	#24	10.102819	9.212988	3.651886	0.000388	0.033874
miR-125a	#107	10.960998	10.005312	3.651356	0.000452	0.033874
miR-95	#84	9.435733	8.751331	3.59406	0.000478	0.039594
miR-155	#157	12.505359	13.231221	-3.58369	0.000614	0.040394
miR-199b	#194	9.755066	9.082751	3.55934	0.000588	0.04314
miR-24-2	#54	12.611696	11.612557	3.518774	0.00087	0.048278
let-7e	#11	12.497795	13.055093	-3.51589	0.00054	0.048354
miR-92-1	#81	16.081074	16.592426	-3.50446	0.000928	0.049828

* - Forty-three miRs have an adjusted p-value lower than 0.05. Twenty-six miRs are overexpressed and 17 down-regulated in breast, colon, lung, pancreas, prostate, stomach
carcinomas.

Table 4. Differentially regulated miRs in 6 solid cancer types vs. normal tissues (SAM, significance analysis of microarrays) *.

miR	ID	d.value	stdev	p.value	q.value	R.fold
miR-21	#47	3.156	0.24	0	0	2.593
miR-23b	#51	-3.117	0.212	0	0	0.443
miR-138-2	#133	-2.514	0.2	0	0	0.402
miR-218-2	#221	-2.383	0.17	0	0	0.384
miR-29b-2	#95	2.246	0.236	0	0	1.868
miR-128a prec	#113	-2.235	0.177	0	0	0.368
miR-195	#184	2.085	0.203	0	0	1.695
miR-141	#137	2.08	0.179	0	0	2.459
miR-199a-1	#191	1.987	0.201	0	0	1.945
miR-9-3	#28	-1.97	0.204	0	0	0.433
miR-16-2	#39	1.966	0.229	0	0	1.788
miR-17-5p	#41	1.964	0.296	0	0	0.725
miR-20a	#46	1.898	0.283	0	0	0.969
miR-16-1	#38	1.87	0.232	0	0	1.447
miR-212 prec	#209	-1.854	0.167	0	0	0.509
miR-34a	#78	1.756	0.232	0	0	1.219
miR-152 prec	#151	-1.734	0.2	0	0	0.46
miR-199a-2	#192	1.721	0.196	0	0	1.838
miR-128b	#115	1.674	0.238	0	0	1.266
miR-212	#208	-1.659	0.121	0	0	0.627
let-7a-1	#1	-1.628	0.157	0	0	0.461
miR-200b	#195	1.626	0.225	0	0	1.432
miR-128a	#114	-1.619	0.156	0	0	0.511
miR-29c	#65	1.611	0.24	0	0	1.225
let-7a-3	#4	1.581	0.226	0	0	1.109
miR-29a	#62	1.565	0.229	0	0	1.706
miR-24-2	#54	1.555	0.284	0	0	0.831
miR-138-1	#132	-1.551	0.222	0	0	0.432
miR-125a	#107	1.541	0.262	0	0	1.164
miR-106a	#99	1.514	0.275	0	0	0.952
miR-132	#121	1.496	0.192	0	0	2.158
miR-30d	#72	-1.491	0.174	0	0	0.424
miR-9-1	#24	1.478	0.244	0	0	0.763
miR-27a	#59	1.448	0.229	0	0	1.174
miR-181b-1 prec	#211	1.435	0.18	0	0	1.525
let-7g	#15	1.394	0.221	0	0	1.072
miR-96	#86	-1.384	0.194	0	0	0.519
miR-191	#177	1.372	0.201	0	0	1.165
miR-93-1	#83	1.363	0.266	0	0	0.775
miR-136	#130	-1.355	0.267	0	0	0.364
miR-205	#201	1.343	0.309	0	0	1.281
miR-185	#170	1.287	0.222	0.001	0.001	0.609
miR-125b-1	#109	1.262	0.283	0.001	0.001	1.215
miR-10a	#30	1.252	0.227	0.001	0.001	1.643
miR-95	#84	1.247	0.19	0.001	0.001	1.509
miR-199b	#194	1.228	0.189	0.001	0.001	1.246
miR-10b	#32	1.219	0.232	0.002	0.001	1.342
let-7i	#10	1.216	0.203	0.002	0.001	1.026
miR-210	#205	1.213	0.237	0.002	0.001	1.088

* - Thirty five miRs are over-expressed and 14 are down-regulated in breast, colon, lung, pancreas, prostate, stomach carcinomas (Delta = 0.9, FDR=0.001).

Table 5. MicroRNAs selected by PAM (prediction analysis of microarray) in 6 solid cancer types vs. normal tissues

miR	ID	Solid cancer score	Normal tissues score
miR-21	#47	0.0801	-0.2643
miR-138-2	#133	-0.055	0.1815
miR-218-2	#221	-0.0535	0.1765
miR-23b	#51	-0.0516	0.17
miR-128a prec	#113	-0.0498	0.1642
miR-29b-2	#95	0.0457	-0.1508
miR-195	#184	0.0404	-0.1333
miR-17-5p	#41	0.0383	-0.1263
miR-9-3	#28	-0.0357	0.1176
miR-212 prec	#209	-0.0342	0.1129
miR-20a	#46	0.0322	-0.1061
miR-141	#137	0.0322	-0.1061
miR-199a-1	#191	0.0319	-0.1053
miR-16-2	#39	0.0315	-0.1037
miR-152 prec	#151	-0.0283	0.0933
miR-16-1	#38	0.0277	-0.0913
miR-34a	#78	0.0269	-0.0886
miR-212	#208	-0.0265	0.0875
let-7a-1	#1	-0.0264	0.0872
miR-128a	#114	-0.0259	0.0855
miR-128b	#115	0.0254	-0.0839
miR-24-2	#54	0.0244	-0.0803
miR-29c	#65	0.0224	-0.0738
miR-199a-2	#192	0.0223	-0.0736
let-7a-3	#4	0.0221	-0.073
miR-191	#177	0.0188	-0.062
miR-125a	#107	0.0186	-0.0613
miR-30d	#72	-0.0185	0.061
miR-29a	#62	0.0184	-0.0608
miR-106a	#99	0.0177	-0.0584
miR-93-1	#83	0.0163	-0.0537
miR-200b	#195	0.0159	-0.0524
let-7g	#15	0.0158	-0.0521
miR-27a	#59	0.0157	-0.0518
miR-96	#86	-0.0156	0.0514
let-7b	#5	-0.0152	0.0501
miR-138-1	#132	-0.0151	0.0499
miR-9-1	#24	0.0136	-0.0448
miR-181b-1 prec	#211	0.0134	-0.0442
miR-155	#157	-0.0128	0.0423
miR-132	#121	0.0127	-0.0418
miR-136	#130	-0.0112	0.037
let-7i	#10	0.0103	-0.034
miR-210	#205	0.0074	-0.0245
miR-205	#201	0.0073	-0.024
* miR-185	#170	0.0071	-0.0234

miR-24-1	#52	0.007	-0.023
miR-199b	#194	0.0064	-0.021
miR-125b-1	#109	0.006	-0.0199
miR-206 prec	#203	-0.005	0.0166
miR-10a	#30	0.0045	-0.015
miR-95	#84	0.0045	-0.0149
let-7c	#11	-0.0039	0.013
miR-124a-3	#106	-0.0028	0.0091
miR-10b	#32	0.002	-0.0066
miR-185 prec	#171	-0.0014	0.0047
miR-92-1	#81	-2.00E-04	5.00E-04

* - T=1.5 and misclassification error = 0.176. Thirty six over-expressed miRs in cancer are indicated by positive cancer scores; 21 down-regulated miRs are indicated by negative cancer scores.

Example 2: Identification of microRNA expression signatures associated with various human solid cancers.

Results

To identify microRNAs that are prognostic for cancer status associated with solid tumors, without incurring bias due to tissue specificity, an alternative approach was used. First, six tissue-specific signatures, one for each cancer histotype, were obtained by performing independent PAM tests (summarized in Tables 6 and 7) Specific signatures for each cancer are shown in Tables 8-13: e.g., breast-Table 8; colon-Table 9; lung-Table 10; pancreas-Table 11; prostate-Table 12; stomach-Table 13. Using these data, deregulated microRNAs that were shared among the different histotype miRNA signatures were identified (Table 14). In order to compute the p-values for this comparative analysis, a re-sampling test with 1,000,000 random permutations on the miRNA identity was performed. The p-value was defined as the relative frequency of simulation scores exceeding the real score. Twenty-one misregulated microRNAs that were common to at least 3 types of solid cancers (p-value = 2.5×10^{-3}) were identified (Table 14).

Table 6. MicroRNAs used to classify human cancers and normal tissues*.

Cancer	Up-regulated miRs	Down-regulated miRs	Misclassification error after 10 fold cross validation
Breast	15	12	0.08
Colon	21	1	0.09
Lung	35	3	0.31
Pancreas	55	2	0.02
Prostate	39	6	0.11
Stomach	22	6	0.19

* - Median normalization was performed and the method of the nearest shrunken centroids was used to select predictive miRNAs.

Table 7. Deregulated microRNAs in solid common cancers*.

Cancer	PAM Up-regulated	SAM Up-regulated	PAM Down-regulated	SAM Down-regulated
Breast	15	3 (FDR=0.33)	12	47
Colon	21	42 (FDR<=0.06)	1	5
Lung	35	38 (FDR<=0.01)	3	3
Pancreas	55	50 (FDR<=0.01)	2	8
Stomach	22	22 (FDR=0.06)	6	4
Prostate	39	49 (FDR=0.06)	6	3

* - Prediction analysis of microarrays (PAM) identifies those genes which best characterize cancers and normal tissues, whilst significance analysis of microarrays (SAM) identifies all those which have differential expression in the two classes. False detection rates (FDR) computed in SAM are indicated in parenthesis.

Table 8. MicroRNAs selected by prediction analysis of microarray (PAM) in breast cancer (cancer vs. normal tissues) *.

miR	Cancer score	Normal score
miR-21 (#47)	0.0331	-0.4364
miR-29b-2 (#95)	0.0263	-0.3467
miR-146 (#144)	0.0182	-0.2391
miR-125b-2 (#111)	-0.0174	0.2286
miR-125b-1 (#109)	-0.0169	0.222
miR-10b (#32)	-0.0164	0.2166
miR-145 (#143)	-0.0158	0.2076
miR-181a (#158)	0.0153	-0.201
miR-140 (#136)	-0.0122	0.1613
miR-213 (#160)	0.0116	-0.1527
miR-29a prec (#63)	0.0109	-0.1441
miR-181b-1 (#210)	0.0098	-0.1284
miR-199b (#194)	0.0089	-0.1172
miR-29b-1 (#64)	0.0084	-0.1111
miR-130a (#120)	-0.0076	0.1001
miR-155 (#157)	0.0072	-0.0951
let-7a-2 (#3)	-0.0042	0.0554
miR-205 (#201)	-0.004	0.0533
miR-29c (#65)	0.0032	-0.0423
miR-224 (#228)	-0.003	0.0399
miR-100 (#91)	-0.0021	0.0283
miR-31 (#73)	0.0017	-0.022
miR-30c (#70)	-7.00E-04	0.009
miR-17-5p (#41)	7.00E-04	-0.0089
miR-210 (#205)	4.00E-04	-0.0057
miR-122a (#101)	4.00E-04	-0.005
miR-16-2 (#39)	-1.00E-04	0.0013

* 27 miRs selected, misclassification error after cross validation of 0.008. Seventeen overexpressed miRs in cancer are indicated by positive cancer scores; 12 down-regulated miRs are indicated by negative cancer scores.

Table 9. MicroRNAs selected by prediction analysis of microarray (PAM) in colon (cancer vs. normal tissues) *.

miR	Cancer score	Normal score
miR-24-1 (#52)	0.0972	-0.5589
miR-29b-2 (#95)	0.0669	-0.3845
miR-20a (#46)	0.0596	-0.3424
miR-10a (#30)	0.0511	-0.2938
miR-32 (#75)	0.0401	-0.2306
miR-203 (#197)	0.0391	-0.2251
miR-106a (#99)	0.0364	-0.2094
miR-17-5p (#41)	0.0349	-0.2005
miR-30c (#70)	0.0328	-0.1888
miR-223 (#227)	0.0302	-0.1736
miR-126* (#102)	0.0199	-0.1144
miR-128b (#115)	0.0177	-0.102
miR-21 (#47)	0.0162	-0.0929
miR-24-2 (#54)	0.0145	-0.0835
miR-99b prec (#88)	0.0125	-0.0721
miR-155 (#157)	0.0092	-0.0528
miR-213 (#160)	0.0091	-0.0522
miR-150 (#148)	0.0042	-0.0243
miR-107 (#100)	0.003	-0.0173
miR-191 (#177)	0.0028	-0.0159
miR-221 (#224)	0.002	-0.0116
miR-9-3 (#28)	-0.0014	0.0083

* 22 miRs selected, misclassification error after cross validation of 0.09. Twenty-one over-expressed miRs in cancer are indicated by positive cancer scores; 1 down-regulated miR is indicated by a negative cancer score.

Table 10. MicroRNAs selected by prediction analysis of microarray (PAM) in lung cancer (cancer vs. normal tissues) *.

miR	Cancer score	Normal score
miR-21 (#47)	0.175	-0.175
miR-205 (#201)	0.1317	-0.1317
miR-200b (#195)	0.1127	-0.1127
miR-9-1 (#24)	0.1014	-0.1014
miR-210 (#205)	0.0994	-0.0994
miR-148 (#146)	0.0737	-0.0737
miR-141 (#137)	0.0631	-0.0631
miR-132 (#121)	0.0586	-0.0586
miR-215 (#213)	0.0575	-0.0575
miR-128b (#115)	0.0559	-0.0559
let-7g (#15)	0.0557	-0.0557
miR-16-2 (#39)	0.0547	-0.0547
miR-129-1/2 prec (#118)	0.0515	-0.0515
miR-126* (#102)	-0.0406	0.0406
miR-142-as (#139)	0.0366	-0.0366
miR-30d (#72)	-0.0313	0.0313
miR-30a-5p (#66)	-0.0297	0.0297
miR-7-2 (#21)	0.0273	-0.0273
miR-199a-1 (#191)	0.0256	-0.0256
miR-127 (#112)	0.0254	-0.0254
miR-34a prec (#79)	0.0214	-0.0214
miR-34a (#78)	0.0188	-0.0188
miR-136 (#130)	0.0174	-0.0174
miR-202 (#196)	0.0165	-0.0165
miR-196-2 (#188)	0.0134	-0.0134
miR-199a-2 (#192)	0.0126	-0.0126
let-7a-2 (#3)	0.0109	-0.0109
miR-124a-1 (#104)	0.0081	-0.0081
miR-149 (#147)	0.0079	-0.0079
miR-17-5p (#41)	0.0061	-0.0061
miR-196-1 prec (#186)	0.0053	-0.0053
miR-10a (#30)	0.0049	-0.0049
miR-99b prec (#88)	0.0045	-0.0045
miR-196-1 (#185)	0.0044	-0.0044
miR-199b (#194)	0.0039	-0.0039
miR-191 (#177)	0.0032	-0.0032
miR-195 (#184)	7.00E-04	-7.00E-04
miR-155 (#157)	7.00E-04	-7.00E-04

* 38 miRs selected, misclassification error after cross validation of 0.31. Thirty-five over-expressed miRs in cancer are indicated by positive cancer scores; 3 down-regulated miRs are indicated by negative cancer scores.

Table 11. MicroRNAs selected by prediction analysis of microarray (PAM) in pancreatic cancer (cancer vs. normal tissues) *.

miR	Cancer score	Normal score
miR-103-2 (#96)	0.4746	-1.582
miR-103-1 (#97)	0.4089	-1.3631
miR-24-2 (#54)	0.4059	-1.3529
miR-107 (#100)	0.3701	-1.2336
miR-100 (#91)	0.3546	-1.182
miR-125b-2 (#111)	0.3147	-1.0489
miR-125b-1 (#109)	0.3071	-1.0237
miR-24-1 (#52)	0.2846	-0.9488
miR-191 (#177)	0.2661	-0.887
miR-23a (#50)	0.2586	-0.8619
miR-26a-1 (#56)	0.2081	-0.6937
miR-125a (#107)	0.1932	-0.644
miR-130a (#120)	0.1891	-0.6303
miR-26b (#58)	0.1861	-0.6203
miR-145 (#143)	0.1847	-0.6158
miR-221 (#224)	0.177	-0.59
miR-126* (#102)	0.1732	-0.5772
miR-16-2 (#39)	0.1698	-0.5659
miR-146 (#144)	0.1656	-0.552
miR-214 (#212)	0.1642	-0.5472
miR-99b (#89)	0.1636	-0.5454
miR-128b (#115)	0.1536	-0.512
miR-155 (#157)	-0.1529	0.5098
miR-29b-2 (#95)	0.1487	-0.4956
miR-29a (#62)	0.1454	-0.4848

Table 11 (continued). MicroRNAs selected by prediction analysis of microarray (PAM) in pancreatic cancer (cancer vs. normal tissues) *.

miR	Cancer score	Normal score
miR-25 (#55)	0.1432	-0.4775
miR-16-1 (#38)	0.1424	-0.4746
miR-99a (#90)	0.1374	-0.4581
miR-224 (#228)	0.1365	-0.4549
miR-30d (#72)	0.1301	-0.4336
miR-92-2 (#82)	0.116	-0.3865
miR-199a-1 (#191)	0.1158	-0.3861
miR-223 (#227)	0.1141	-0.3803
miR-29c (#65)	0.113	-0.3768
miR-30b (#68)	0.1008	-0.3361
miR-129-1/2 (#117)	0.1001	-0.3337
miR-197 (#189)	0.0975	-0.325
miR-17-5p (#41)	0.0955	-0.3185
miR-30c (#70)	0.0948	-0.316
miR-7-1 (#19)	0.0933	-0.311
miR-93-1 (#83)	0.0918	-0.3061
miR-140 (#136)	0.0904	-0.3015
miR-30a-5p (#66)	0.077	-0.2568
miR-132 (#121)	0.0654	-0.2179
miR-181b-1 (#210)	0.0576	-0.1918
miR-152 prec (#151)	-0.0477	0.1591
miR-23b (#51)	0.0469	-0.1562
miR-20a (#46)	0.0452	-0.1507
miR-222 (#225)	0.0416	-0.1385
miR-27a (#59)	0.0405	-0.1351
miR-92-1 (#81)	0.0332	-0.1106
miR-21 (#47)	0.0288	-0.0959
miR-129-1/2 prec (#118)	0.0282	-0.0939
miR-150 (#148)	0.0173	-0.0578
miR-32 (#75)	0.0167	-0.0558
miR-106a (#99)	0.0142	-0.0473
miR-29b-1 (#64)	0.0084	-0.028

* 57 miRs selected, misclassification error after cross validation of 0.02. Fifty-seven miRs are over-expressed and 2 are down-regulated in cancer (indicated by positive and negative scores, respectively).

Table 12. MicroRNAs selected by prediction analysis of microarray (PAM) in prostate cancer (cancer vs. normal tissues) *.

miR	Cancer score	Normal score
let-7d (#8)	0.0528	-0.4227
miR-128a prec (#113)	-0.0412	0.3298
miR-195 (#184)	0.04	-0.3199
miR-203 (#197)	0.0356	-0.2851
let-7a-2 prec (#2)	-0.0313	0.2504
miR-34a (#78)	0.0303	-0.2428
miR-20a (#46)	0.029	-0.2319
miR-218-2 (#221)	-0.0252	0.2018
miR-29a (#62)	0.0247	-0.1978
miR-25 (#55)	0.0233	-0.1861
miR-95 (#84)	0.0233	-0.1861
miR-197 (#189)	0.0198	-0.1587
miR-135-2 (#128)	0.0198	-0.1582
miR-187 (#173)	0.0192	-0.1535
miR-196-1 (#185)	0.0176	-0.1411
miR-148 (#146)	0.0175	-0.1401
miR-191 (#177)	0.017	-0.136
miR-21 (#47)	0.0169	-0.1351
let-7i (#10)	0.0163	-0.1303
miR-198 (#190)	0.0145	-0.1161
miR-199a-2 (#192)	0.0136	-0.1088
miR-30c (#70)	0.0133	-0.1062
miR-17-5p (#41)	0.0132	-0.1053
miR-92-2 (#82)	0.012	-0.0961
miR-146 (#144)	0.0113	-0.0908
miR-181b-1 prec (#211)	0.011	-0.0878
miR-32 (#75)	0.0109	-0.0873

Table 12 (continued). MicroRNAs selected by prediction analysis of microarray (PAM) in prostate cancer (cancer vs. normal tissues) *.

miR	Cancer score	Normal score
miR-206 (#202)	0.0104	-0.083
miR-184 prec (#169)	0.0096	-0.0764
miR-29a prec (#63)	-0.0095	0.076
miR-29b-2 (#95)	0.0092	-0.0739
miR-149 (#147)	-0.0084	0.0676
miR-181b-1 (#210)	0.0049	-0.0392
miR-196-1 prec (#186)	0.0042	-0.0335
miR-93-1 (#83)	0.0039	-0.0312
miR-223 (#227)	0.0038	-0.0308
miR-16-1 (#38)	0.0028	-0.0226
miR-101-1 prec (#92)	0.0015	-0.0123
miR-124a-1 (#104)	0.0015	-0.0119
miR-26a-1 (#56)	0.0015	-0.0119
miR-214 (#212)	0.0013	-0.0105
miR-27a (#59)	0.0011	-0.0091
miR-24-1 (#53)	-8.00E-04	0.0067
miR-106a (#99)	7.00E-04	-0.0057
miR-199a-1 (#191)	4.00E-04	-0.0029

* - T=1, 45 miRs selected, misclassification error after cross validation of 0.11. Thirty-nine over-expressed miRs in cancer are indicated by positive cancer scores; 6 downregulated miRs are indicated by negative cancer scores.

Table 13. MicroRNAs selected by prediction analysis of microarray (PAM) in stomach cancer (cancer vs. normal tissues) *.

miR	Cancer score	Normal score
miR-223 (#227)	0.1896	-0.1806
miR-21 (#47)	0.1872	-0.1783
miR-218-2 (#221)	-0.1552	0.1478
miR-103-2 (#96)	0.1206	-0.1148
miR-92-2 (#82)	0.1142	-0.1088
miR-25 (#55)	0.1097	-0.1045
miR-136 (#130)	-0.1097	0.1045
miR-191 (#177)	0.0946	-0.0901
miR-221 (#224)	0.0919	-0.0876
miR-125b-2 (#111)	0.0913	-0.0869
miR-103-1 (#97)	0.0837	-0.0797
miR-214 (#212)	0.0749	-0.0713
miR-222 (#225)	0.0749	-0.0713
miR-212 prec (#209)	-0.054	0.0514
miR-125b-1 (#109)	0.0528	-0.0503
miR-100 (#91)	0.0526	-0.0501
miR-107 (#100)	0.0388	-0.0369
miR-92-1 (#81)	0.0369	-0.0351
miR-96 (#86)	-0.0306	0.0291
miR-192 (#178)	0.0236	-0.0224
miR-23a (#50)	0.022	-0.021
miR-215 (#213)	0.0204	-0.0194
miR-7-2 (#21)	0.0189	-0.018
miR-138-2 (#133)	-0.0185	0.0176
miR-24-1 (#52)	0.0151	-0.0144
miR-99b (#89)	0.0098	-0.0093
miR-33b (#76)	-0.0049	0.0046
miR-24-2 (#54)	0.0041	-0.0039

* - T=1, 28 miRs selected, misclassification error after cross validation of 0.19. Twenty-two over-expressed miRs in cancer are indicated by positive cancer scores; 6 down-regulated miRs are indicated by negative cancer scores.

Table 14. The microRNAs shared by the signatures of the 6 solid cancers*.

miR	N	Tumor Type
miR-21	6	Breast Colon Lung Pancreas Prostate Stomach
miR-17-5p	5	Breast Colon Lung Pancreas Prostate
miR-191	5	Colon Lung Pancreas Prostate Stomach
miR-29b-2	4	Breast Colon Pancreas Prostate
miR-223	4	Colon Pancreas Prostate Stomach
miR-128b	3	Colon Lung Pancreas
miR-199a-1	3	Lung Pancreas Prostate
miR-24-1	3	Colon Pancreas Stomach
miR-24-2	3	Colon Pancreas Stomach
miR-146	3	Breast Pancreas Prostate
miR-155	3	Breast Colon Lung
miR-181b-1	3	Breast Pancreas Prostate
miR-20a	3	Colon Pancreas Prostate
miR-107	3	Colon Pancreas Stomach
miR-32	3	Colon Pancreas Prostate
miR-92-2	3	Pancreas Prostate Stomach
miR-214	3	Pancreas Prostate Stomach
miR-30c	3	Colon Pancreas Prostate
miR-25	3	Pancreas Prostate Stomach
miR-221	3	Colon Pancreas Stomach
miR-106a	3	Colon Pancreas Prostate

* - The list includes 21 commonly up-regulated microRNAs in 3 or more (N) types of solid cancers (p-value = 2.5×10^{-3}).

To maximize concision, the mean absolute expression levels of the deregulated miRs for the 6 cancer/normal pairs were computed. Using the expression level of miRs in the comprehensive subset, the different tissues were correctly classified, irrespective of the disease status (FIG. 3).

FIG. 4 shows differential expression of the common microRNAs across the different tumor tissues, in relation to the normal tissues. The tree displays the different cancer types according to fold changes in the miRNA subset. Prostate, colon, stomach and pancreatic tissues are most similar among them, while lung and breast tissues were represented by a fairly different signature (FIG. 4). This tree clearly shows which miRNAs are associated with a particular cancer histotype.

Strikingly, miR-21, miR-191 and miR-17-5p are significantly over-expressed in all, or in 5 out of 6, of the tumor types that were considered. miR-21 was reported to be over-expressed in glioblastoma and to have anti-apoptotic properties (Chan, J.A., et al.,

Cancer Res. 65: 6029-6033 (2005)). Lung cancer shares a portion of its signature with breast cancer and a portion with the other solid tumors, including miR-17/20/92, all three of which are members of the microRNA cluster that actively cooperates with c-Myc to accelerate lymphomagenesis (He, L., et al., *Nature 435*: 828-833 (2005)). The identification of these microRNAs as being over-expressed is an excellent confirmation of our approach. A second miRNA group that is activated includes miR-210 and miR-213, together with miR-155, which was already reported to be amplified in large cell lymphomas (Eis, P.S., et al., *Proc. Natl. Acad. Sci. USA 102*: 3627-3632 (2005)), children with Burkitt lymphoma (Metzler, M., et al., *Genes Chromosomes Cancer 39*:167-169 (2004)) and various B cell lymphomas (Kluiver, J, et al., *J. Pathol.*, e-published online, July 22, 2005). These microRNAs are the only ones up-regulated in breast and lung cancer. miR-218-2 is consistently down-regulated in colon, stomach, prostate and pancreas cancers, but not in lung and breast carcinomas.

Several observations strengthen these results. First, in this study, the expression levels of both the precursor pre-miRNA and the mature miRNA were determined for the majority of genes. Of note, with the exception of miR-212 and miR-128a, in all other instances, the abnormally-expressed region was that corresponding to the active gene product. Second, as shown in FIG. 3, the expression variation of the miRNAs in the comprehensive subset was often univocal (namely, down- or up-regulation) across the different types of cancers, suggesting a common mechanism in human tumorigenesis. Third, the microarray data were validated by solution hybridization for 12 breast samples (miR-125b, miR-145 and miR-21; Iorio, M.V., *et al., Cancer Res. 65*: 7065-7070 (2005)) and 17 endocrine pancreatic and normal samples (miR-103, miR-155 and miR-204; data not shown), strongly confirming the accuracy of the microarray data.

Example 3: Identification of predicted targets for microRNAs that are deregulated in solid tumors.

Materials and Methods:

Tumor suppressor and oncogene target predictions

The most recent TargetScan predictions (April 2005) were used to identify putative microRNA targets. These include essentially the 3'UTR targets reported by Lewis et al. (Lewis, B.P., et al, *Cell 120*: 15-20 (2005)), with a few changes arising

from updated gene boundary definitions from the April 2005 UCSC Genome Browser mapping of RefSeq mRNAs to the hg17 human genome assembly. Among the putative targets, known cancer genes (tumor suppressors and oncogenes) were specified according to their identification in the Cancer Gene Census, or as reported by OMIM.

Target in vitro assays

For luciferase reporter experiments, 3' UTR segments of Rb1, TGFBR2 and Plag1 that are predicted to interact with specific cancer-associated microRNAs were amplified by PCR from human genomic DNA and inserted into the pGL3 control vector (Promega) using the XbaI site immediately downstream from the stop codon of luciferase. The human megakaryocytic cell line, MEG-01, was grown in 10% FBS in RPMI medium 1640, supplemented with 1x nonessential amino acid and 1 mmol sodium pyruvate at 37° C in a humidified atmosphere of 5% CO_2. The cells were co-transfected in 12-well plates by using siPORT neoFX (Ambion, Austin, TX), according to the manufacturer's protocol, with 0.4 μg of the firefly luciferase reporter vector and 0.08 μg of the control vector containing Renilla luciferase, pRL-TK (Promega). For each well, microRNA oligonucleotides (Dharmacon Research, Lafayette, CO) and anti-sense or scrambled oligonucleotides (Ambion) were used at a concentration of 10 nM. Firefly and Renilla luciferase activities were measured consecutively at 24 h post transfection using dual-luciferase assays (Promega).

Western blotting for RB1

Levels of RB1 protein were quantified using a mouse monoclonal anti-RB1 antibody (Santa Cruz, CA) using standard procedures for Western blotting. The normalization was performed with mouse monoclonal anti-Actin antibody (Sigma).

Results

The functional significance of microRNA deregulation in cancer needs to be understood. In solid tumors, it appears that the most common microRNA event is gain of expression, while loss of expression in cancer is a more limited event, and more tissue specific. We used a three-step consequential approach in the following order: first, "*in silico*" prediction of targets, then luciferase assay for first validation of cancer relevant targets and finally, *ex vivo* tumor correlation between miRNA expression (by

microarray) and target protein expression (by Western blotting) for a specific miRNA:mRNA interactor pair. Relevant targets for cancer miRNAs could be either recessive (e.g., tumor suppressors) or dominant (e.g., oncogenes) cancer genes. To test the hypothesis that microRNAs that are deregulated in solid tumors target known oncogenes or tumor suppressors, the predicted targets for these miRNAs were determined using TargetScan, a database of conserved 3' UTR microRNA targets (Lewis, B.P., *et al, Cell 120*: 15-20 (2005)). TargetScan contained 5,121 predictions for 18 miRNAs that are dysregulated in solid tumors, in the total 22,402 (26.5%) predictions. One hundred fifteen out of 263 (44%) well-known cancer genes were predicted as targets for these 18 miRNAs (Table 15). Because a high percentage of cancer genes are targeted by miRs that are deregulated in solid tumors, it is unlikely that these predictions are due to chance ($P<0.0001$ at Fisher exact-test).

In silico predictions for three different cancer genes, Retinoblastoma (Rb), TGF-beta-2 receptor (TGFBR2), and pleiomorphic adenoma gene 1 (PLAG1), were confirmed experimentally by in vitro assays. Using a luciferase reporter assay, three microRNAs tested (miR-106a, miR-20a and miR-26a-1) caused a significant reduction of protein translation relative to the scrambled control oligoRNAs in transfected MEG-01 cells (FIG. 6). Retinoblastoma 3'UTR, for example, was found to interact functionally with miR-106a. The biological significance of this miRNA:mRNA interaction is reinforced by previous reports showing that the Rb1 gene is normally transcribed in colon cancers, whilst various fractions of cells do not express Rb1 protein (Ali, A.A., *et al.*, *FASEB J.* 7:931-937 (1993)). This finding suggests the existence of a post-transcriptional mechanism for regulating Rb1 that could be explained by concomitant miR-106a over-expression in colon carcinoma (FIG. 4). Furthermore, mir-20a is down-regulated in breast cancer (FIG. 4) and TFGBR2 protein is expressed in the epithelium of breast cancer cells (Buck, M.B., *et al., Clin. Cancer Res. 10*:491-498 (2004)). Conversely, the over-expression of mir-20a in colon cancer may represent a novel mechanism for down-regulating TGFBR2, in addition to mutational inactivation (Biswas, S., *et al., Cancer Res. 64*:687-692 (2004)).

Finally, a set of patient samples was tested to verify whether RB1 protein expression correlates with *miR-106a* expression (FIG. 5 and FIG. 6B). As expected, in gastric, prostate and lung tumor samples RB1 was down-regulated (in respect to the

paired normal) and *miR-106a* was found to be over-expressed, while in breast tumor samples, where *miR-106a* is slightly down-regulated (FIG. 5 and FIG. 6B), RB1 is expressed at slightly higher levels then in the paired normal control.

These experimental proofs reinforce the hypothesis that key cancer genes are regulated by aberrant expression of miRs in solid cancers. These data add novel examples to the list of microRNA with important cancer gene targets, as previously shown by Johnsson et al. (Johnson, S.M., *et al., Cell 120:* 635-647 (2005)) for the let-7:Ras interaction, O'Donnell et al. (O'Donnell, K.A., *et al., Nature 435*:839-843 (2005)) for the miR-17-5p:cMyc interaction, and Cimmino et al. (Cimmino, A., *et al., Proc. Natl. Acad. Sci. USA 102*:13944-13949 (2005)) for the mir-16:Bcl2 interaction. Notably, miR-17-5p and miR-16 are members of the miRNA solid cancer signature described herein.

Table 15. Oncogenes and tumor suppressor genes predicted by TargetScanS as targets of microRNAs from the comprehensive cancer subset.*

miRNA gene	Gene Name	Gene description
miR-26a, miR-146	ABL2	v-abl Abelson murine leukemia viral oncogene homolog 2 (arg, Abelson-related gene)
miR-107	AF5q31	ALL1 fused gene from 5q31
miR-20, miR-125b	AKT3	v-akt murine thymoma viral oncogene homolog 3
miR-26a, miR-155 miR-125b	APC	adenomatosis polyposis coli
miR-26a, miR-218	ARHGEF12	RHO guanine nucleotide exchange factor (GEF) 12 (LARG)
miR-107, miR-221	ARNT	aryl hydrocarbon receptor nuclear translocator
miR-192	ATF1	activating transcription factor 1
miR-26a	ATM	Ataxia telangiectasia mutated (includes complementation groups A, C and D)
miR-24	AXL	AXL receptor tyrosine kinase
miR-26a, miR-107, miR-146, miR-155 miR-138, miR-92	BCL11A	B-cell CLL/lymphoma 11A
miR-20	BCL11B	B-cell CLL/lymphoma 11B (CTIP2)
miR-21	BCL2	B-cell CLL/lymphoma 2
miR-26a, miR-26a miR-20,	BCL6	B-cell CLL/lymphoma 6 (zinc finger protein 51)
miR-92	BCL9	B-cell CLL/lymphoma 9
miR-26a, miR-223 miR-221, miR-125b	CBFB	core-binding factor, beta subunit
miR-218	CCDC6	coiled-coil domain containing 6
miR-20	CCND1	cyclin D1 (PRAD1: parathyroid adenomatosis 1)
miR-26a, miR-20	CCND2	cyclin D2
miR-26a, miR-107, miR-92	CDK6	cyclin-dependent kinase 6

miR-20	CDKN1A	cyclin-dependent kinase inhibitor 1A (p21, Cip1)
miR-221, miR-92	CDKN1C	cyclin-dependent kinase inhibitor 1C (p57, Kip2)
miR-24	CDX2	caudal type homeo box transcription factor 2
miR-92	CEBPA	CCAAT/enhancer binding protein (C/EBP), alpha
miR-26a	CLTC	clathrin, heavy polypeptide (Hc)
miR-218	COL1A1	collagen, type I, alpha 1
miR-26a	CREBBP	CREB binding protein (CBP)
miR-20	CRK	v-crk avian sarcoma virus CT10 oncogene homolog
miR-20	CSF1	colony stimulating factor 1 (macrophage)
miR-221, miR-192	DDX6	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 6 (RNA helicase, 54kD)
miR-138	DEK	DEK oncogene (DNA binding)
miR-20	E2F1	E2F transcription factor 1
miR-20	ELK3	ELK3, ETS-domain protein (SRF accessory protein 2)
miR-24	ELL	ELL gene (11-19 lysine-rich leukemia gene)
miR-26a, miR-138	ERBB4	v-erb-a avian erythroblastic leukemia viral oncogene homolog-like 4
miR-221, miR-155, miR-125b	ETS1	v-ets avian erythroblastosis virus E26 oncogene homolog 1
miR-20	ETV1	ets variant gene 1
miR-125b	ETV6	ets variant gene 6 (TEL oncogene)
miR-223	FAT	FAT tumor suppressor (Drosophila) homolog
miR-223, miR-125b, miR-218	FGFR2	fibroblast growth factor receptor 2
miR-92	FLI1	Friend leukemia virus integration 1
miR-24, miR-20	FLT1	fms-related tyrosine kinase 1 (vascular endothelial growth factor/vascular permeability factor receptor)
miR-221	FOS	v-fos FBJ murine osteosarcoma viral oncogene homolog

miR-92	FOXG1B	forkhead box G1B
miR-223	FOXO3A	forkhead box O3A
miR-125b	GOLGA5	golgi autoantigen, golgin subfamily a, 5 (PTC5)
miR-138	GPHN	gephyrin (GPH)
miR-107, miR-223, miR-20, miR-218	HLF	hepatic leukemia factor
miR-26a, miR-107	HMGA1	high mobility group AT-hook 1
miR-20	HOXA13	homeo box A13
miR-92	HOXA9	homeo box A9
miR-125b	IRF4	interferon regulatory factor 4
miR-146, miR-20, miR-138	JAZF1	juxtaposed with another zinc finger gene 1
miR-92	JUN	v-jun avian sarcoma virus 17 oncogene homolog
miR-155	KRAS	v-Ki-ras2 Kirsten rat sarcoma 2 viral oncogene homolog
miR-218	LASP1	LIM and SH3 protein 1
miR-218	LHFP	lipoma HMGIC fusion partner
miR-125b, miR-218	LIFR	leukemia inhibitory factor receptor
miR-223	LMO2	LIM domain only 2 (rhombotin-like 1) (RBTN2)
miR-223, miR-155, miR-125b, miR-92	MAF	v-maf musculoaponeurotic fibrosarcoma (avian) oncogene homolog
miR-92	MAP2K4	mitogen-activated protein kinase kinase 4
miR-146, miR-20	MAP3K8	mitogen-activated protein kinase kinase kinase 8
miR-125b	MAX	MAX protein
miR-218	MCC	mutated in colorectal cancers
miR-24	MEN1	multiple endocrine neoplasia I
miR-138	MLLT6	myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila); translocated to, 6 (AF17)

miR-192	MSN	moesin
miR-24	MYB	v-myb avian myeloblastosis viral oncogene homolog
miR-107, miR-223, miR-146, miR-221, miR-155, miR-218	MYBL1	v-myb avian myeloblastosis viral oncogene homolog-like 1
miR-107, miR-20	MYCN	v-myc avian myelocytomatosis viral related oncogene, neuroblastoma derived
miR-107, miR-92	MYH9	myosin, heavy polypeptide 9, non-muscle
miR-24	MYST4	MYST histone acetyltransferase (monocytic leukemia) 4 (MORF)
miR-20	NBL1	neuroblastoma, suppression of tumorigenicity 1
miR-125b	NIN	ninein (GSK3B interacting protein)
miR-26a, miR-107	NKTR	natural killer-tumor recognition sequence
miR-92	NOTCH1	Notch homolog 1, translocation-associated (Drosophila) (TAN1)
miR-24	NTRK3	neurotrophic tyrosine kinase, receptor, type 3
miR-125b	PCSK7	proprotein convertase subtilisin/kexin type 7
miR-24, miR-146	PER1	period homolog 1 (Drosophila)
miR-146, miR-125b, miR-138,	PHOX2B	paired-like homeobox 2b
miR-155	PICALM	phosphatidylinositol binding clathrin assembly protein (CALM)
miR-24, miR-26a	PIM1	pim-1 oncogene
miR-24, miR-26a, miR-21, miR-107, miR-20, miR-155	PLAG1	pleiomorphic adenoma gene 1
miR-218	RAB8A	RAB8A, member RAS oncogene family
miR-24, miR-221	RALA	v-ral simian leukemia viral oncogene homolog A (ras related)
miR-138	RARA	retinoic acid receptor, alpha
miR-20, miR-192	RB1	retinoblastoma 1 (including osteosarcoma)
miR-20,	RBL1	retinoblastoma-like 1 (p107)
miR-20	RBL2	retinoblastoma-like 2 (p130)

miR-155, miR-138	REL	v-rel avian reticuloendotheliosis viral oncogene homolog
miR-20, miR-138	RHOC	ras homolog gene family, member C
miR-20, miR-192	RUNX1	runt-related transcription factor 1 (AML1)
miR-107, miR-223	SEPT6	septin 6
miR-146, miR-20, miR-125b	SET	SET translocation
miR-21, miR-20, miR-155, miR-218	SKI	v-ski avian sarcoma viral oncogene homolog
miR-26a, miR-146	SMAD4	SMAD, mothers against DPP homolog 4 (Drosophila)
miR-155	SPI1	spleen focus forming virus (SFFV) proviral integration oncogene spi1
miR-125b	SS18	synovial sarcoma translocation, chromosome 18
miR-107, miR-155	SUFU	suppressor of fused homolog (Drosophila)
miR-92	TAF15	TAF15 RNA polymerase II, TATA box binding protein (TBP)-associated factor, 68kDa
miR-26a, miR-221, miR-138	TCF12	transcription factor 12 (HTF4, helix-loop-helix transcription factors 4)
miR-21, miR-20	TGFBR2	transforming growth factor, beta receptor II (70-80kD)
miR-24, miR-26a, miR-92	TOP1	topoisomerase (DNA) I
miR-138	TPM4	tropomyosin 4
miR-20	TRIP11	thyroid hormone receptor interactor 11
miR-92	TSC1	Tuberous sclerosis 1
miR-20	TSG101	Tumor susceptibility gene 101
miR-20	TUSC2	Tumor suppressor candidate 2
miR-24	VAV1	vav 1 oncogene
miR-125b	VAV2	vav 2 oncogene
miR-107	WHSC1	Wolf-Hirschhorn syndrome candidate 1(MMSET)
miR-138	WHSC1L1	Wolf-Hirschhorn syndrome candidate 1-like 1 (NSD3)

miR-26a	WNT5A	wingless-type MMTV integration site family, member 5A
miR-26a, miR-20, miR-125b	YES1	v-yes-1 Yamaguchi sarcoma viral oncogene homolog 1
miR-107, miR-221	ZNF198	zinc finger protein 198
miR-218	ZNFN1A1	zinc finger protein, subfamily 1A, 1 (Ikaros)

* - Known cancer genes (e.g., tumor suppressors, oncogenes) comprise those identified in the Cancer Gene Census or reported by OMIM.

While this invention has been particularly shown and described with references to preferred embodiments thereof, it will be understood by those skilled in the art that various changes in form and details may be made therein without departing from the scope of the invention encompassed by the appended claims.

THE EMBODIMENTS OF THE PRESENT INVENTION IN WHICH AN EXCLUSIVE PROPERTY OR PRIVILEGE IS CLAIMSED ARE DEFINED AS FOLLOWS:

1. A method of diagnosing whether a subject has, or is at risk for developing, a solid cancer, comprising:

measuring in a test sample from the subject a level of miR-21 gene product and miR-191 gene product,

comparing the level of miR-21 gene product and miR-191 gene product in the test sample to a control level of miR-21 gene product and miR-191 gene product; and

diagnosing whether a subject has, or is at risk for developing, a solid cancer selected from the group consisting of colon, lung, pancreas, prostate and stomach cancer,

wherein an increase in the level of miR-21 gene product and miR-191 gene product in the test sample, relative to the control level of a corresponding miR-21 gene product and miR-191 gene product is indicative of the subject either having, or being at risk for developing, a solid cancer selected from the group consisting of: colon, lung, pancreatic, prostate and stomach cancer.

2. A method of diagnosing whether a subject has, or is at risk for developing, a solid cancer, comprising:

a. reverse transcribing miR-21 RNA and miR-191 RNA from a test sample obtained from the subject to provide miR-21 and miR-191 oligodeoxynucleotides;

b. hybridizing the miR-21 and miR-191 oligodeoxynucleotides to a microarray comprising miRNA-specific probe oligonucleotides, which includes miR-21 and miR-191 specific probe oligonucleotides, to provide a hybridization profile for the test sample; and

c. comparing the test sample hybridization profile to a control hybridization profile,

wherein, if a signal of miR-21 RNA and miR-191 RNA in the test sample hybridization profile is greater than the signal of miR-21 RNA and miR-191 RNA in the control hybridization profile, than the subject either has, or is at risk for developing, a solid cancer selected from the group consisting of colon, lung, pancreas, prostate and stomach cancer.

3. The method of Claim 1 or 2, further comprising measuring at least one miR in addition to measuring miR-21 and miR-191.

4. The method of Claim 3, wherein the at least one miR is miR-17-5p.

Application number/numéro de demande: 2.633.754

Figures: 1, 2, 3, 4, 5, 6A, 6B, 7, 8

Pages: No page #'s

Unscannable Item(s) received with this application.

To inquire is you can order a copy of the unscannable item(s),

Please visit the CIPO Website at HTTP://CIPO.GC.CA

Item(s) ne pouvant être balayés.

Documents reçus avec cette demande ne pouvant être balayes.

Pour vous reseigner si vous pouvez commander

une copie des items ne pouvant être balayes,

veuillez visiter le site web de l' OPIC au HTTP://CIPO.GC.CA