Innovation, Sciences et Développement économique Canada
Office de la Propriété Intellectuelle du Canada

Innovation, Science and Economic Development Canada
Canadian Intellectual Property Office

CA 2875710 C 2021/06/29

(11)(21) 2 875 710

(12) BREVET CANADIEN
CANADIAN PATENT

(13) C

(86) Date de dépôt PCT/PCT Filing Date: 2013/06/24
(87) Date publication PCT/PCT Publication Date: 2013/12/27
(45) Date de délivrance/Issue Date: 2021/06/29
(85) Entrée phase nationale/National Entry: 2014/12/03
(86) N° demande PCT/PCT Application No.: US 2013/047354
(87) N° publication PCT/PCT Publication No.: 2013/192616
(30) Priorité/Priority: 2012/06/22 (US61/663,428)

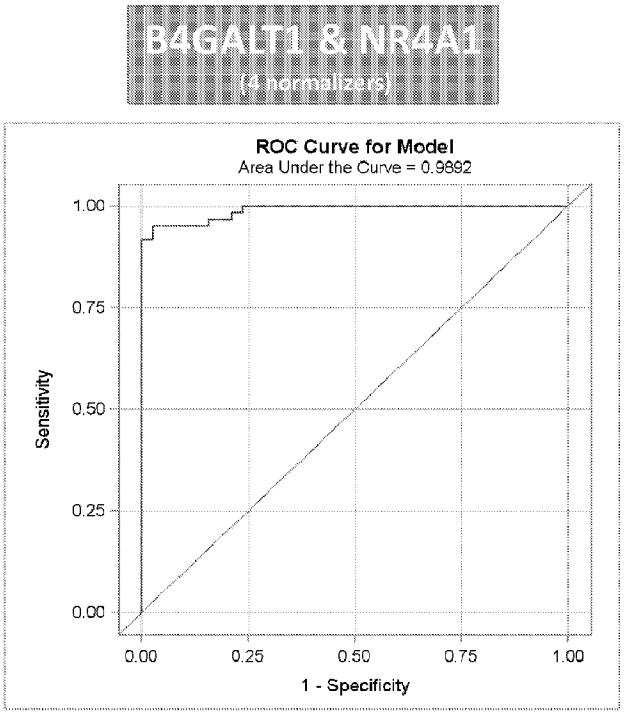
(51) Cl.Int./Int.Cl. *C12Q 1/6809* (2018.01), *C12Q 1/6837* (2018.01), *C12Q 1/6886* (2018.01), *C40B 40/06* (2006.01)

(72) Inventeurs/Inventors:
WANG, HUI, US;
ROBERTS, CHRISTOPHER, US;
MADDULA, KRISHNA, US;
LU, ZHENQUIANG, US;
BROWN, JUSTIN, US;
...

(73) Propriétaires/Owners:
JOHN WAYNE CANCER INSTITUTE, US;
HTG MOLECULAR DIAGNOSTICS, INC., US

(74) Agent: LAVERY, DE BILLY, LLP

(54) Titre : MALIGNITE MOLECULAIRE DANS DES LESIONS MELANOCYTIQUES
(54) Title: MOLECULAR MALIGNANCY IN MELANOCYTIC LESIONS




(57) Abrégé/Abstract:
Disclosed are methods for determining whether a melanocyte-containing sample (such as a nevus or other pigmented lesion) is benign or a primary melanoma. These methods can include detecting (at the molecular level, e.g., mRNA, miRNA, or protein) the

CA 2875710 C 2021/06/29
(11)(21) 2 875 710
(13) C

(72) Inventeurs(suite)/Inventors(continued): VASICEK, TOM, US; KERNS, B.J., US; SELIGMANN, BRUCE, US; HOON, DAVE S.B., US

(57) Abrégé(suite)/Abstract(continued):
expression of at least two disclosed genes in a biological sample obtained from a subject. Also provided are arrays and kits that can be used with the methods.

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
27 December 2013 (27.12.2013)

WIPO | PCT

(10) International Publication Number
WO 2013/192616 A1

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

(21) International Application Number:
PCT/US2013/047354

(22) International Filing Date:
24 June 2013 (24.06.2013)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/663,428 22 June 2012 (22.06.2012) US

(71) Applicants: **HTG MOLECULAR DIAGNOSTICS, INC.** [US/US]; 3430 E. Global Loop, Tucson, AZ 85706 (US). **JOHN WAYNE CANCER INSTITUTE** [US/US]; 2200 Santa Monica Blvd., Santa Monica, CA 90404 (US).

(72) Inventors: **WANG, Hui**; 3430 E. Global Loop, Tucson, AZ 85706 (US). **ROBERTS, Christopher**; 3430 E. Global Loop, Tucson, AZ 85706 (US). **MADDULA, Krishna**; 3430 E. Global Loop, Tucson, AZ 85706 (US). **LU, Zhenquiang**; 3430 E. Global Loop, Tucson, AZ 85706 (US). **BROWN, Justin**; 3430 E. Global Loop, Tucson, AZ 85706 (US). **VASICEK, Tom**; 3430 E. Global Loop, Tucson, AZ 85706 (US). **KERNS, B.J.**; 3430 E. Global Loop, Tucson, AZ 85706 (US). **SELIGMANN, Bruce**; 3430 E. Global Loop, Tucson, AZ 85706 (US). **HOON, Dave S.B.**; 2200 Santa Monica Blvd., Santa Monica, CA 90404 (US).

(74) Agent: **RYBAK, Sheree Lynn**; Klarquist Sparkman, LLP, One World Trade Center, Suite 1600, 121 SW Salmon Street, Portland, OR 97204 (US).

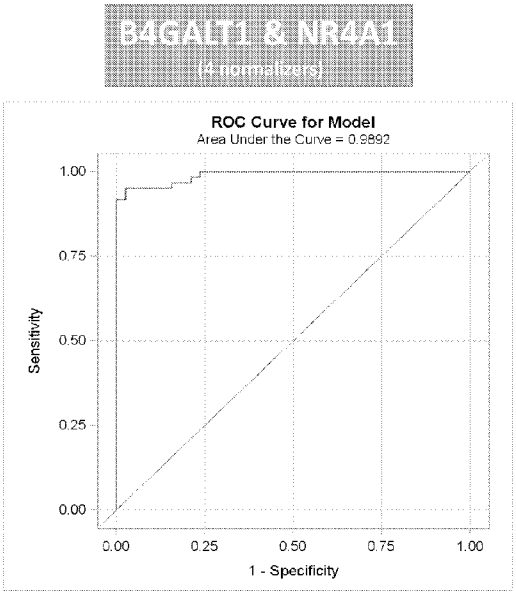
(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE,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

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK,

*[Continued on next page]*

(54) Title: MOLECULAR MALIGNANCY IN MELANOCYTIC LESIONS

FIG. 4




(57) Abstract: Disclosed are methods for determining whether a melanocyte-containing sample (such as a nevus or other pigmented lesion) is benign or a primary melanoma. These methods can include detecting (at the molecular level, e.g., mRNA, miRNA, or protein) the expression of at least two disclosed genes in a biological sample obtained from a subject. Also provided are arrays and kits that can be used with the methods.

WO 2013/192616 A1

WO 2013/192616 A1

EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

**Declarations under Rule 4.17**:

— *of inventorship (Rule 4.17(iv))*

**Published**:

— *with international search report (Art. 21(3))*

— *with sequence listing part of description (Rule 5.2(a))*

# MOLECULAR MALIGNANCY IN MELANOCYTIC LESIONS

## CROSS-REFERENCE TO RELATED APPLICATION

This application claims priority to U.S. Provisional Application No. 61/663,428 filed June 22, 2012.

## FIELD

This disclosure concerns biomarkers for characterizing melanocytic lesions as benign or malignant. In particular, this disclosure concerns the identification of biomarkers (including mRNA and/or miRNA) that are significantly differentially expressed in nevi and primary melanoma samples, clinically predictive algorithms based on the expression of such biomarkers, and methods of and compositions for their use.

## PARTIES TO JOINT RESEARCH AGREEMENT

HTG Molecular Diagnostics and the John Wayne Cancer Institute are parties to a joint research agreement governing inventions disclosed herein.

## BACKGROUND

Skin cancer is the most common of all cancers in the United States. Melanoma, a cancer originating in melanocytes, accounts for a relatively small percentage of skin cancers. However, melanoma causes the most skin cancer deaths making it one of the most dangerous types of skin cancer. In 2012, melanoma will account for more than 75,000 skin cancer cases.

Melanocytes also are found in organs other than skin, including the eye (*e.g.*, in or on the uvea, ciliary body, conjunctiva, eyelid, iris, or orbit), the inner ear, meninges, bones, and heart. Ocular melanoma is the most common type of eye tumor in adults and the second most common type of primary malignant melanoma in the body. Ocular melanoma has an incidence of about five cases per one-million people in the United States.

To diagnose melanoma, suspect tissue is biopsied and examined under a microscope by a pathologist, preferably (but often not) one who is specially trained to identify melanoma in tissue biopsies. If the pathologist reports finding a melanoma, a number of factors (including the depth of the tumor in millimeters, the presence or absence of ulceration, the mitotic rate, and/or whether the tumor has spread) are used in determining a person's prognosis and course of treatment(s). When the tumor has not spread, a wider local excision is often performed to ensure that the entire lesion

was removed along with a clear margin of normal tissue around the melanoma. If more extreme treatments are indicated, the patient also may receive lymphadenectomy, immunotherapy, chemotherapy, or radiation therapy.

Melanoma is almost always curable when it is found in its very early stages. Unfortunately, misdiagnoses of this disease are common (Piepkorn *et al.*, *J. Am. Acad. Dermatol.*, 30:707, 1994; Farmer *et al.*, *Hum. Pathol.*, 27:528, 1996; Corona *et al.*, *J. Clin. Oncol.* 14:1218, 1996; Barnhill *et al.*, *Hum. Pathol.*, 30:513, 1990; Brochez *et al.*, *J. Pathol.* 196:459, 2002). Diagnostic errors have a number of root causes (*e.g.*, see Ruiter *et al.*, *Sem. Cutaneous Med. Surg.*, 22:33, 2003), including difficulties in differentiating between benign melanocytic nevi and early melanoma and between atypical and dysplastic nevi.

Mistakes in melanoma diagnosis have a significant adverse impact on the patients, their families, and society in general. Patients mistakenly diagnosed with a melanoma may undergo inappropriate and potentially dangerous therapy(ies), may live a life in constant fear of relapse, and may not be able to obtain life or health insurance. On the other hand, patients mistakenly diagnosed with a nevus instead of a melanoma are deprived of appropriate therapy for their malignancy, and may have their lives prematurely cut short. Finally, the societal toll of this problem is demonstrated by the fact that misdiagnosis of melanoma is the second only to misdiagnosis of breast cancer as the most common reason for cancer-based medical malpractice claims in the United States (McDonald *et al.*, *Internet J. Fam. Practice*, 7(2), 2009; *Troxel, Am. J. Surg. Pathol.*, 27:1278, 200).

Given the limitations of histopathology alone, it is of critical importance in medical science to have additional tools for the proper diagnosis of melanoma. In particular, tools are needed to determine which biopsies (*e.g.*, dysplastic or indeterminate nevi) may, in fact, be misdiagnosed melanoma, and/or which biopsies (*e.g.*, nevi) may demonstrate molecular characteristics of melanoma or progression to melanoma.

## SUMMARY

Disclosed are methods for characterizing a melanocyte-containing sample, for example determining whether a sample is a benign nevi or a malignant melanoma. In some examples, these methods include characterizing a melanocyte-containing sample by determining an expression level (such as a nucleic acid or protein level) for (i) at least two of the biomarkers selected from MAGEA2, PRAME, PDIA4, NR4A1, PDLIM7, B4GALT1, SAT1, RUNX1, SOCS3 and those in Table 13 and (ii) at least one normalization biomarker(s), in the melanocyte-containing sample

obtained from a subject (such as a nevi sample), thereby generating raw expression values for each of the at least two biomarkers and the at least one normalization biomarker(s). The raw expression values for each of the at least two biomarkers are normalized to the raw expression values for the at least one normalization biomarker(s) to generate normalized expression values for each of the at least two biomarkers. The normalized expression values are used in a regression or machine learning algorithm to generate an output value. The resulting output value is compared to a cut-off value, which can be derived from normalized expression values for the at least two biomarkers in a plurality of melanocyte-containing samples known in advance to be benign or malignant. The melanocyte-containing sample obtained from the subject is then characterized, for example as benign if the output value is on the same side of the cut-off value as the plurality of known benign samples or as malignant if the output value is on the same side of the cut-off value as the plurality of known malignant samples.

Also provided are methods for determining malignancy in a melanocyte-containing sample. Such a method can include determining an expression level (such as a nucleic acid expression level) for at least two biomarkers selected from: B4GALT1, BAX, MAGEA2, NR4A1, PDIA4, PRAME, RUNX1, SOCS3, SAT1, PDLIM7, BIRC5, MET, MAGEC2, POLR2J3, ZFYVE16, and BEST1 in a melanocyte-containing sample obtained from a subject. The method can also include calculating an output from an algorithm that uses the expression levels of the at least two biomarkers as an input and determining from the algorithm output that the sample is or is not malignant by comparing the output to a reference standard from known malignant melanocyte-containing samples. The method can further include normalizing the expression levels of the at least two selected biomarkers to the expression level of at least one normalization biomarker, such as at least one of those in Table 3.

Also disclosed are arrays and kits for diagnosing a biological sample (such as a melanocyte-containing sample) as a benign nevi or a primary melanoma. For example, an array can include at least three addressable locations, each location having immobilized capture probes with the same specificity, and each location having capture probes with a specificity that is different from the capture probes at each other location, wherein the capture probes at two of the at least three locations are capable of directly or indirectly specifically hybridizing a biomarker that includes two or more of MAGEA2, PRAME, PDIA4, NR4A1, PDLIM7, B4GALT1, SAT1, RUNX1, SOCS3 and those in Table 13, and the capture probes at one of the at least three locations is capable of directly or indirectly specifically hybridizing to a normalization biomarker listed in Table 3, and wherein the specificity of each capture probe is identifiable by the addressable location the array.

Kits are provided that include one or more arrays provided herein, as well as one or more of: a container containing lysis buffer; a container containing a nuclease specific for single-stranded nucleic acids; a container containing a plurality of nucleic acid programming linkers; a container containing a plurality of NPPs; a container containing a plurality of the bifunctional detection linker; a container containing a detection probe that specifically binds the bifunctional detection linkers; and a container containing a detection reagent.

Also provided is a method of characterizing a melanocyte-containing sample, comprising:

determining an expression level for

(i) biomarkers MAGEA2, PRAME, PDIA4, NR4A1, PDLIM7, B4GALT1, SAT1, RUNX1, and SOCS3, and

(ii) at least one normalization biomarker(s), in a melanocyte-containing sample obtained from a subject, thereby generating raw expression values for each of the biomarkers and the at least one normalization biomarker(s);

normalizing the raw expression values for each of the biomarkers to the raw expression values for the at least one normalization biomarker(s) to generate normalized expression values for each of the biomarkers;

using the normalized expression values in a regression or machine learning algorithm to generate an output value;

comparing the output value to a cut-off value, wherein the cut-off value was derived from normalized expression values for the biomarkers in a plurality of melanocyte-containing samples known in advance to be benign or malignant; and

characterizing the sample as benign if the output value is on the same side of the cut-off value as the plurality of known benign samples or characterizing the sample as malignant if the output value is on the same side of the cut-off value as the plurality of known malignant samples.

Also provided is a method of determining gene expression in a melanocyte-containing sample, comprising:

determining in the sample the expression levels of a plurality of genes comprising biomarkers MAGEA2, PRAME, PDIA4, NR4A1, PDLIM7, B4GALT1, SAT1, RUNX1, and SOCS3; and

providing a report of the plurality of genes expression levels in the sample or a characterization of the sample as a nevus or melanoma based on the expression levels of the plurality of genes.

Also provided is a method of determining malignancy in a melanocyte-containing sample, comprising:

determining, in a melanocyte-containing sample obtained from a subject, an expression level of biomarkers B4GALT1, BAX, MAGEA2, NR4A1, PDIA4, PRAME, RUNX1, SOCS3, SAT1, PDLIM7, BIRC5, MET, MAGEC2, POLR2J3, ZFYVE16, and BEST1;

calculating an output from an algorithm that uses the expression levels of the biomarkers as an input; and

determining from the algorithm output that the sample is or is not malignant by comparing the output to a reference standard from known malignant melanocyte-containing samples.

Also provided is an array, comprising:

at least three addressable locations, each location comprising immobilized capture probes having the same specificity, and each location comprising capture probes having specificity different than capture probes at each other location,

wherein the capture probes at two of the at least three locations are capable of directly or indirectly specifically hybridizing biomarker MAGEA2, PRAME, PDIA4, NR4A1, PDLIM7, B4GALT1, SAT1, RUNX1, and SOCS3, and the capture probes at one of the at least three locations is capable of directly or indirectly specifically hybridizing a normalization biomarker listed in Table 3; and

wherein the specificity of each capture probe is identifiable by the addressable location the array.

Also provided is a kit, comprising:

an array described herein, and

one or more of:

a container containing lysis buffer;

a container containing a nuclease specific for single-stranded nucleic acids;

a container containing a plurality of nucleic acid programming linkers;

a container containing a plurality of NPPs;

a container containing a plurality of the bifunctional detection linkers;

a container containing a detection probe that specifically binds the bifunctional detection linkers; and

a container containing a detection reagent.

The foregoing and other features of this disclosure will become more apparent from the following detailed description of a several embodiments which proceeds with reference to the accompanying figures.

## BRIEF DESCRIPTION OF THE DRAWINGS

**FIG. 1** is a flow diagram showing how embodiments of a diagnostic test disclosed herein (as indicated by the flowchart elements (in gray shaded) emanating from the arrow downward from the "Biopsy" point) fit into current Nation Comprehensive Cancer Network (NCCN) clinical recommendations for melanoma diagnosis.

**FIGS. 2A** and **2B** show box plots (top), mean plots (middle) and SAS diffograms (bottom) for the representative normalization genes indicated above the respective graphs (*i.e.*, MFI2, RAP2B, BMP1 and NCOR2). Collectively, these results show that there were no statistically significant differences between nevi and primary melanoma samples for each normalizer gene, and that each such gene produced consistent results with low standard deviations.

**FIG. 3** shows SAS output demonstrating the statistical significance of the representative B4GALT1 and NR4A1 (4-normalizer) model. Collectively, the output demonstrate that the model converged on a solution and, thus, that the results of the model were reliable. The model fit and test of global null hypotheses show that the overall model was statistically significant or that the probability that the observed results were far less likely than could be attributed to chance alone, Wald Chi-Square = 15.856, 2df, p=0.0004. The Hosmer and Lemeshow test tests the null hypothesis that there is no lack of fit to the model; or the model accurately reproduces the data. No significance was found using the Hosmer and Lemeshow test further supporting the value of the model. It is noted that a significant Hosmer and Lemeshow p-value (*e.g.*, less than 0.05) would suggest that there was some lack of fit to the model or that the proposed model, in some capacity, failed to fit the experimental data adequately.

**FIG. 4** shows the ROC curve for the representative B4GALT1 and NR4A1 (4-normalizer) model. The ROC curve illustrates the very high sensitivity and specificity for the model.

Sensitivity represents the true positive rate (*i.e.*, if a person has a disease, how often will the test be positive; or, sensitivity = (true positives/(true positive + false negative)). Specificity represents the true negative rate (*i.e.*, if a person does not have the disease how often will the test be negative; or, specificity = (true negatives/(true negative + false positives). The area under the curve (AUC = 0.9892) illustrates the ability of the model to differentiate between the two populations, i.e., nevi and primary melanoma, with very high accuracy.

**FIG. 5** shows the classification results after SAS cross validation for the representative B4GALT1 and NR4A1 (4-normalizer) model at different thresholds. The probability level is the probability of calling a test sample a primary melanoma. By raising the threshold (cut-off value) for calling a sample a primary melanoma the model obtained very high specificity and good sensitivity. These results further demonstrated that very high specificity and good sensitivity was obtained using this model over a wide range of threshold values.

**FIG. 6** shows a continuation of the FIG. 5 classification table. These continued results show that lowering the cut-off threshold resulted in higher sensitivity with a minor tradeoff in specificity while still maintaining very high overall classification accuracy.

**FIG. 7** shows that the representative B4GALT1 and NR4A1 (4-normalizer) model was highly significant even under multiple different estimation routines. One common assumption in regression-based models is equal variances. Unequal variances, especially when sample sizes are unequal, can cause standard estimation practices to give incorrect results. Although the Brown-Forsythe test for equality of variances showed no significant difference between the population variances (not shown), an Empirical Covariance "Sandwich" Estimator test, which is used when there may be unequal variances or some other violation of common assumption, was run. The Sandwich Estimator test (left box) confirmed that the original results obtained under the standard Fisher Scoring method were not due to violation of model assumptions. Similarly, the Firth bias reduction penalized likelihood model (right box) provided additional confirmation that the results were not sensitive to estimation procedure.

**FIG. 8** shows that the B4GALT1 and NR4A1 (2-normalizer) model fit, as indicated by the Wald Chi-Square, was also highly significant. The ROC curve demonstrates that this model also had very high sensitivity and specificity. The very small change in the area under the curves (i.e., $\Delta = 0.0125$) for the B4GALT1 and NR4A1 (2-normalizer) and B4GALT1 and NR4A1 (4-normalizer) models shows that the two models are very similar with respect to their abilities to correctly differentiate between nevi and primary melanoma samples.

**FIG. 9** shows a probability classification table for the B4GALT1 and NR4A1 (2-normalizer) model. These results demonstrate that the model maintained very high sensitivity and specificity. Compared to the B4GALT1 and NR4A1 (4-normalizer) model, the overall specificity of the 2-normalizer model was somewhat attenuated across the range of the model; however, there is always a tradeoff between sensitivity and specificity. The overall sensitivity for thresholds of 0.34 and below showed that the model provided moderately higher sensitivity while maintaining good specificity. Given that the clinical implications are far worse for misdiagnosing a sample, trading some specificity for sensitivity is an acceptable outcome. The B4GALT1 and NR4A1 (2-normalizer) model had overall correct classification of 88.9% or better for approximately 50% of the thresholds.

**FIG. 10A** shows three scatter plots, each showing the result of a univariate statistical test (AUC (top), fold change (fch; middle), and FDR-adjusted p-value (bottom)) for each gene (as measured by mRNA expression) listed on the x-axis. The dotted line in each scatterplot shows the selected cut off for statistical significance. The result is considered significant if above the AUC cut off (also boxed), below the fold change cut off (also boxed), or below the FDR-adjusted p-value cut off. The symbol representing each gene represents on which ArrayPlate (AP) the expression data was measured.

**FIG. 10B** shows similar results as FIG. 10A for each indicated miRNA (x-axis), except that the cut off for fold change is positive 1 (vs. negative 1) and the fold change result is considered significant if above the line. The expression value for each miRNA was (+) or was not (*) normalized.

**FIG. 11** shows the classification accuracy (based on AUC) of exemplary two (bottom left) to 40 (top right) gene nevus/melanoma classifiers built on the expression data from ArrayPlate No. 3. In each case, the AUC equals or exceed 0.9 indicating good accuracy regardless of the number of genes in the classifier and increasing classifier accuracy until approximately 18-gene classifiers whereafter the AUC is relatively stable at approximately 0.95.

**FIG. 12** is a composite of four line graphs, each showing the misclassification rate (y-axis) of two to 40 gene (x-axis) AUC, T-test, Random Forest, or LIMMA classification models based on expression data collected from ArrayPlate Nos. 3-6, as indicated.

## SEQUENCES

The nucleic acid sequences listed herein are shown using standard letter abbreviations for nucleotide bases, as defined in 37 C.F.R. 1.822. Only one strand of each nucleic acid sequence is

shown, but the complementary strand is understood as included by any reference to the displayed strand.

The Sequence Listing is submitted as an ASCII text file in the form of the file named "Sequence.txt" (~371 kb), which was created on June 24, 2013, which is incorporated by reference herein.

In the provided sequences:

SEQ ID NOs. 1-36, 123, and 124 are representative nuclease protection probe (NPP) sequences.

SEQ ID NOs. 47-119 are GenBank mRNA RefSeqs for the genes disclosed as differentially expressed in nevi and primary melanoma.

SEQ ID NOs. 37-46, 120, and 121 are GenBank mRNA RefSeqs for disclosed normalizers.

SEQ ID NO. 122 is the GenBank mRNA RefSeq for a disclosed negative control plant gene (ANT).

SEQ ID NOs. 125-144 are representative NPP sequences for disclosed mRNA targets.

SEQ ID NOs. 145-164 are representative NPP sequences for disclosed miRNA targets.

## DETAILED DESCRIPTION

Unless otherwise noted, technical terms are used according to conventional usage. Definitions of common terms in molecular biology may be found in Benjamin Lewin, Genes IX, published by Jones and Bartlet, 2008 (ISBN 0763752223); Kendrew *et al.* (eds.), The Encyclopedia of Molecular Biology, published by Blackwell Science Ltd., 1994 (ISBN 0632021829); and Robert A. Meyers (ed.), Molecular Biology and Biotechnology: a Comprehensive Desk Reference, published by VCH Publishers, Inc., 1995 (ISBN 9780471185710).

The singular terms "a," "an," and "the" include plural referents unless context clearly indicates otherwise. Similarly, the word "or" is intended to include "and" unless the context clearly indicates otherwise. The term "comprises" means "includes." In case of conflict, the present specification, including explanations of terms, will control.

Suitable methods and materials for the practice or testing of this disclosure are described below. Such methods and materials are illustrative only and are not intended to be limiting. Other methods and materials similar or equivalent to those described herein can be used. For example, conventional methods well known in the art to which a disclosed invention pertains are described in various general and more specific references, including, for example, Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual, 2d ed.*, Cold Spring Harbor Laboratory Press, 1989; Sambrook *et*

al., *Molecular Cloning: A Laboratory Manual, 3d ed.*, Cold Spring Harbor Press, 2001; Ausubel *et al.*, *Current Protocols in Molecular Biology*, Greene Publishing Associates, 1992 (and Supplements to 2000); Ausubel *et al.*, *Short Protocols in Molecular Biology: A Compendium of Methods from Current Protocols in Molecular Biology*, 4th ed., Wiley & Sons, 1999; Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, 1990; and Harlow and Lane, *Using Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, 1999. In addition, the materials, methods, and examples are illustrative only and not intended to be limiting.

Genbank Numbers are referred to herein for the sequence available on June 22, 2012.

To facilitate review of the various embodiments of this disclosure, the following explanations of specific terms are provided:

**Antibody:** A polypeptide ligand comprising at least a light chain or heavy chain immunoglobulin variable region which specifically recognizes and binds an epitope of an antigen or a fragment thereof, for example an epitope a biomarker shown in Table 3, 4, 11, or 13. The term antibody includes intact immunoglobulins and the variants and portions of them well known in the art, such as Fab' fragments, $F(ab)'_2$ fragments, single chain Fv proteins ("scFv"), and disulfide stabilized Fv proteins ("dsFv"). The term also includes genetically engineered forms such as chimeric antibodies, heteroconjugate antibodies (such as, bispecific antibodies). The term antibody includes both polyclonal and monoclonal antibodies. The preparation of polyclonal and monoclonal antibodies, molecularly engineered antibodies and antibody fragments is well known to those of ordinary skill in the art (see, *e.g.*, Green *et al.*, "Production of Polyclonal Antisera," in: *Immunochemical Protocols* pages 1-5, Manson, ed., Humana Press 1992; and Harlow *et al.*, in: *Antibodies: a Laboratory Manual*, page 726, Cold Spring Harbor Pub., 1988).

**Binding or stable binding (of an oligonucleotide):** An oligonucleotide binds or stably binds to a target nucleic acid (such as a biomarker shown in Table 3, 4, 11, or 13) if a sufficient amount of the oligonucleotide forms base pairs or is hybridized to its target nucleic acid, for example the binding of an oligonucleotide, such as an probe or primer to the nucleic acid sequence of a gene shown in Table 3, 4, 11, or 13. Binding between a target and an oligonucleotide can be detected by any procedure known to one skilled in the art, including both functional (for example reduction in expression and/or activity) and physical binding assays.

**Contacting:** Placement in direct physical association including in solid and/or liquid form, for example contacting a sample (*e.g.*, a sample suspended in buffer) with a nucleic acid probe, such as a probe specific for one of the biomarkers shown in Table 3, 4, 11, or 13. Contacting can occur *in vitro*, for example in a diagnostic assay, or, in other examples, ex situ.

**Conditions sufficient to detect:** Any environment that permits the desired activity, for example, that permits an antibody to bind an antigen (such as a biomarker shown in Table 3, 4, 11 or 13), and the interaction to be detected. In other examples, it is the detection of a nucleic acid, such as a biomarker shown in Table 3, 4, 11 or 13, for example by detecting hybridization of the biomarker to a nucleic acid probe.

**Degenerate variant:** A polynucleotide encoding a protein of interest (such as a biomarker shown in Table 3, 4, or 11) that includes a sequence that is degenerate as a result of the genetic code. There are 20 natural amino acids, most of which are specified by more than one codon. Therefore, all degenerate nucleotide sequences are included as long as the amino acid sequence of the polypeptide encoded by the nucleotide sequence is unchanged.

**Detect**: To determine if an agent (such as a signal or particular nucleic acid, nucleic acid probe, or protein, for example one of those in Table 3, 4, 11 or 13) is present or absent. In some examples, this can further include quantification, for example the quantification of the amount of the gene or protein, or a fraction of a sample, such as a particular cell or cells within a tissue.

**Diagnostic:** Identifying the presence or nature of a pathologic condition, such as, but not limited to cancer, such as melanoma. Diagnostic methods differ in their sensitivity and specificity. The "sensitivity" of a diagnostic assay is the percentage of diseased individuals who test positive (percent of true positives). The "specificity" of a diagnostic assay is 1 minus the false positive rate, where the false positive rate is defined as the proportion of those without the disease who test positive. While a particular diagnostic method may not provide a definitive diagnosis of a condition, it suffices if the method provides information (*e.g.*, a positive indication) that aids in diagnosis.

**Hybridization:** Oligonucleotides and their analogs hybridize by hydrogen bonding, which includes Watson-Crick, Hoogsteen or reversed Hoogsteen hydrogen bonding, between complementary bases. Generally, nucleic acid consists of nitrogenous bases that are either pyrimidines (cytosine (C), uracil (U), and thymine (T)) or purines (adenine (A) and guanine (G)). These nitrogenous bases form hydrogen bonds between a pyrimidine and a purine, and the bonding of the pyrimidine to the purine is referred to as "base pairing." More specifically, A will hydrogen bond to T or U, and G will bond to C. "Complementary" refers to the base pairing that occurs

between two distinct nucleic acid sequences or two distinct regions of the same nucleic acid sequence. For example, an oligonucleotide can be complementary to an mRNA, a DNA, or dsDNA encoded by one of the genes in Table 3, 4, 11, or 13.

"Specifically hybridizable" and "specifically complementary" are terms that indicate a sufficient degree of complementarity such that stable and specific binding occurs between the oligonucleotide (or it's analog) and the DNA or RNA target. The oligonucleotide or oligonucleotide analog need not be 100% complementary to its target sequence to be specifically hybridizable. An oligonucleotide or analog is specifically hybridizable when there is a sufficient degree of complementarity between the oligonucleotide or analog to the target DNA or RNA molecule (for example a DNA or RNA in Table 3, 4, 11, or 13) to avoid non-specific binding of the oligonucleotide or analog to non-target sequences under conditions where specific binding is desired. Such binding is referred to as specific hybridization.

Hybridization conditions resulting in particular degrees of stringency will vary depending upon the nature of the hybridization method of choice and the composition and length of the hybridizing nucleic acid sequences. Generally, the temperature of hybridization and the ionic strength (especially the $Na^+$ concentration) of the hybridization buffer will determine the stringency of hybridization, though waste times also influence stringency. Hybridization of an oligonucleotide sequence can be modified by incorporating un-natural bases into the sequence, such as incorporating locked nucleic acids or peptide nucleic acids.

**Isolated:** An "isolated" biological component (such as a nucleic acid molecule, protein or organelle) has been substantially separated or purified away from other biological components in the cell of the organism in which the component naturally occurs, *e.g.*, other chromosomal and extra-chromosomal DNA and RNA, proteins and/organelles. Nucleic acids and proteins that have been "isolated" include nucleic acids and proteins purified by standard purification methods. The term also embraces nucleic acids and proteins prepared by recombinant expression in a host cell as well as chemically synthesized nucleic acids, such as probes and primers, for example probes and primer for the detection and/or amplification of nucleic acids shown in Table 3, 4, 11, or 13.

**Label:** A detectable compound or composition, which can be conjugated directly or indirectly to another molecule, such as an antibody (for example an antibody that specifically binds a biomarker (*e.g.*, protein) shown in Table 3, 4, 11, or 13) or a nucleic acid probe (for example a nucleic acid probe that specifically binds or indirectly binds to a nucleic acid in Table 3, 4, 11, or 13) or a protein, to facilitate detection of that molecule. Specific, non-limiting examples of labels, and methods of labeling nucleic acids and proteins are described throughout this disclosure.

**Melanoma:** A malignant tumor of melanocytes. Melanocytes are cells that produce the dark pigment, melanin, which is responsible for the color of skin. They predominantly occur in skin, but are also found in other parts of the body, including the bowel and the eye. Thus primary melanomas can occur in areas of the body other than the skin (*e.g.*, uveal melanoma). A **primary melanoma** is neoplasia at the site of origin; even if the primary tumor has metastasized the original site remains primary and the distant site is the metastasis.

**Nevus (plural nevi):** A sharply circumscribed pigmented spot on the skin, or other part of the body, such as the bowel or eye. Nevi may be commonly referred to as birthmarks or moles. Nevi comprise melanocytes, which contribute to the nevi's pigmented appearance. Typically, nevi are considered benign. However, a dysplastic nevus (also sometimes referred to as an atypical mole) is a type of nevus with abnormal features. A dysplastic nevus may be bigger than and its color, surface, and border may be different from a non-dysplastic nevus. On the skin surface, a dysplastic nevus can appear as having a mixture of several colors (*e.g.*, from pink to dark brown), a smooth or slightly scaly or pebbly surface, and irregular edges that may fade into the surrounding skin. Dysplastic nevi are more likely than "ordinary" nevi to develop into melanoma, and about half of melanomas arise from dysplastic nevi. However, most dysplastic nevi never become malignant; thus, it is important to be able to determine which nevi (whether dysplastic or non-dysplastic) may, in fact, mistakenly be or be biologically transforming (*e.g.*, at the molecular level) to primary melanoma.

**Nuclease:** An enzyme that cleaves a phosphodiester bond. An endonuclease is an enzyme that cleaves an internal phosphodiester bond in a nucleotide chain (in contrast to exonucleases, which cleave a phosphodiester bond at the end of a nucleotide chain). Some nucleases have both endonuclease and exonuclease activities. Illustrative nucleases are described throughout this disclosure.

**Primer**: A short nucleic acid molecule, such as a DNA oligonucleotide, for example sequences of at least 15 nucleotides, which can be annealed to a complementary target nucleic acid molecule (such as one of the biomarkers in Table 3, 4, 11, or 13) by nucleic acid hybridization to form a hybrid between the primer and the target nucleic acid strand, for example under very high stringency hybridization conditions.

A primer can be extended along the target nucleic acid molecule by a polymerase enzyme. Therefore, primers can be used to amplify a target nucleic acid molecule (such as a portion of a nucleic acid molecule shown in Table 3, 4, 11, or 13), wherein the sequence of the primer is

specific for the target nucleic acid molecule, for example so that the primer will hybridize to the target nucleic acid molecule under very high stringency hybridization conditions.

The specificity of a primer typically increases with its length. Thus, for example, a primer that includes 30 consecutive nucleotides will anneal to a target sequence with a higher specificity than a corresponding primer of only 15 nucleotides. Thus, to obtain greater specificity, probes and primers can be selected that include at least 15, 20, 25, 30, 35, 40, 45, 50 or more consecutive nucleotides of the target sequence.

In particular examples, a primer is at least 10 nucleotides in length, such as at least 15 contiguous nucleotides complementary to a target nucleic acid molecule. Particular lengths of primers that can be used to practice the methods of the present disclosure (for example, to amplify a region of a nucleic acid molecule shown in Table 3, 4, 11, or 13) include primers having at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, at least 24, at least 25, at least 30, at least 35, at least 40, at least 45, at least 50, or more contiguous nucleotides complementary to the target nucleic acid molecule to be amplified, such as a primer of 10-60 nucleotides, 10-50 nucleotides, or 10-30 nucleotides.

Primer pairs can be used for amplification of a nucleic acid sequence, for example, by PCR, real-time PCR, or other nucleic-acid amplification methods known in the art and as described elsewhere in this disclosure. An "upstream" or "forward" primer is a primer 5' to a reference point on a nucleic acid sequence. A "downstream" or "reverse" primer is a primer 3' to a reference point on a nucleic acid sequence.

**Probe**: A probe comprises an isolated nucleic acid capable of hybridizing to a target nucleic acid (such as a nucleic acid sequence of a biomarker shown in Table 3, 4, 11, or 13), and a detectable label or reporter molecule can be attached to a nucleic acid molecule. For example, a label can be attached at the 5'- or 3'-end of the probe, or anywhere in between. In specific examples, the label is attached to the base at the 5'-end of the probe, the base at its 3'-end, the phosphate group at its 5'-end or a modified base, such as a T internal to the probe. Exemplary labels, methods for labeling and guidance in the choice of labels appropriate for various purposes are discussed elsewhere in this disclosure.

Probes are generally at least 15 nucleotides in length, such as at least 10, at least 15, at least 16, at least 17, at least 18, at least 19, least 20, at least 21, at least 22, at least 23, at least 24, at least 25, at least 30, at least 35, at least 40, at least 45, at least 50, at least 55, at least 60, at least 70, at least 80, at least 90, at least 100, at least 120, at least 140, at least 160, at least 180, at least 200, at

least 250, at least 300, at least 350, at least 400, at least 450, at least 500, or more contiguous nucleotides complementary to the target nucleic acid molecule (such as those in Table 3, 4, 11, or 13), such as 20-500 nucleotides, 100-250 nucleotides, 20-50 nucleotides, or 20-30 nucleotides.

**Sequence identity/similarity:** The identity/similarity between two or more nucleic acid sequences, or two or more amino acid sequences, is expressed in terms of the identity or similarity between the sequences. Sequence identity can be measured in terms of percentage identity; the higher the percentage, the more identical the sequences are. Homologs or orthologs of nucleic acid or amino acid sequences possess a relatively high degree of sequence identity/similarity when aligned using standard methods.

Methods of alignment of sequences for comparison are well known in the art; for example, Altschul *et al.*, *J. Mol. Biol.* 215:403-10, 1990, presents a detailed consideration of sequence alignment methods and homology calculations. The NCBI Basic Local Alignment Search Tool (BLAST) (Altschul *et al.*, *J. Mol. Biol.* 215:403-10, 1990) is available from several sources, including the National Center for Biological Information (NCBI, National Library of Medicine, Building 38A, Room 8N805, Bethesda, MD 20894) and on the Internet.

Homologs and variants of the sequences for those molecules shown in Table 4, 11, or 13 are encompassed by this disclosure typically characterized by possession of at least about 75%, for example at least about 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98% or at least 99% sequence identity counted over the full length alignment with the amino acid or nucleic acid sequence of interest, and can retain the activity of the native protein or nucleic acid. One of skill in the art will appreciate that these sequence identity ranges are provided for guidance only; it is entirely possible that strongly significant homologs could be obtained that fall outside of the ranges provided.

One functional indication that two nucleic acid molecules are closely related is that the two molecules hybridize to each other under stringent conditions.

**Methods and Compositions for Characterizing Melanocyte-Containing Samples**

For most cancers, including melanoma, early detection has the greatest impact on survival and can contribute to better cure rates. In some cases, it is difficult to distinguish between a benign and malignant lesion based solely on classical methods (*e.g.*, histopathology). Thus, methods that permit benign nevi to be distinguished from melanomas (*e.g.*, primary melanomas) are needed. Evolving testing methods can help identify malignancies on the molecular level, *e.g.*, before such malignancies can reliably be recognized at the microscopic or organismal level. Molecular testing

involves identifying cancer phenotypes to clinically relevant gene expression patterns, as described herein for distinguishing a benign nevus from a malignant melanoma (*e.g.*, primary melanoma). Such distinctions can avoid unnecessary therapies for those having only a benign nevus, and help to ensure those who have primary melanoma receive appropriate therapies after the initial biopsy.

***Preparing to Collect Gene Expression Data***

Gene expression is the process by which information encoded in the genome (gene) is transformed (*e.g.*, via transcription and translation processes) into corresponding gene products (*e.g.*, RNA (such as, mRNA and miRNA) and protein), which function interrelatedly to give rise to a set of characteristics (aka, phenotype). For purposes of this disclosure, gene expression may be measured by any technique known now or in the future. Commonly, gene expression is measured by detecting the products of the genes (*e.g.*, mRNA, miRNA, and/or protein) expressed in samples collected from subjects of interest.

*Subjects and Samples*

Appropriate samples for use in the methods disclosed herein include any conventional biological sample containing melanocytes for which information about gene expression (*e.g.*, mRNA, miRNA or protein expression; such as those in Table(s) 3, 4, 11, and/or 13) is desired.

Samples include those obtained from a subject, such as clinical samples obtained from a subject (including samples from a healthy or apparently healthy human subject or a human patient affected by a condition or disease to be diagnosed or investigated, such as melanoma). A subject is a living multicellular vertebrate organism, a category that includes, for example, mammals. A "mammal" includes both human and non-human mammals, such as dogs, mice or other veterinary subjects. In one example, the sample is from a subject who has no history of prior melanoma, or is from a subject who has previously had or been diagnosed with melanoma. In some examples, a subject is a patient, such as a patient presenting for skin cancer (*e.g.*, melanoma) screening, or diagnosed with melanoma or at risk (or higher risk) for developing melanoma; for example, as described below. In some examples, the sample is from a subject who has no history of prior melanoma or from a subject who previously was diagnosed with melanoma.

The highest rates of melanoma in humans are reported in Australia (followed by New Zealand, Norway, Sweden, Switzerland, Denmark, United States, Austria, Iceland, Netherlands). Risk factors for a human subject developing melanoma include (a) family or personal history of melanoma; (b) multiple nevi (*e.g.*, greater than 50 or 100 nevi), (c) multiple dysplastic nevi (*e.g.*, at least three), (d) high exposure to sunlight (*e.g.*, before age 10), (e) pale Caucasian skin, (f) red or blond hair, (g) history of at least one blistering sunburn, (h) higher socioeconomic class, (i) history

of sunbed use (especially before age 30), (j) occupation as an airline crew member, and (k) pesticide exposure (MacKie *et al.*, *Annals of Oncology*, 20(Supp. 6), vi1–7, 2009).

In some examples, a prior-used method was unable to reliably determine if the melanocyte-containing sample was malignant or benign. Thus, the disclosed methods can include using and/or determining that the sample to be analyzed cannot reliably be diagnosed as malignant or benign by another method; for example, by histopathology. Such an optional step can occur before determining levels of gene expression levels in the sample (*e.g.*, gene expression of at least two different biomarkers in Table(s) 4, 11 and/or 13 (such as, gene combinations in Tables 6, 8 or 14), and/or at least one normalization biomarker(s)).

Exemplary samples include, without limitation, cells, cell lysates, cytocentrifuge preparations, cytology smears, tissue biopsies (*e.g.,* skin biopsies, such as those that include a nevus or an ocular tissue biopsy), fine-needle aspirates, and/or tissue sections (*e.g.*, cryostat tissue sections and/or paraffin-embedded tissue sections. Tissue is a plurality of functionally related cells. In particular examples, a tissue can be in suspension or intact. In one example the melanocyte-containing sample (such as, a tissue sample) includes a nevus, dysplastic nevus, atypical nevus, or suspected melanoma. In particular examples, samples are used directly (*e.g.*, fresh or frozen), or can be manipulated prior to use, for example, by fixation (*e.g.*, using formalin) and/or embedding in wax (such as formalin-fixed paraffin-embedded (FFPE) tissue samples). Thus, in some examples, the melanocyte-containing sample to be analyzed is fixed. Other method embodiments include fixing the sample (*e.g.*, skin biopsy) in a fixative (*e.g.*, formalin), embedding the sample (*e.g.*, with paraffin), cutting or sectioning the sample, or combinations thereof.

Standard techniques for acquisition of samples useful in the present disclosure are available (see *e.g.*, Schluger *et al., J. Exp. Med.* 176:1327-33 (1992); Bigby *et al., Am. Rev. Respir. Dis.* 133:515-18 (1986); Kovacs *et al., NEJM* 318:589-93 (1988); and Ognibene *et al., Am. Rev. Respir. Dis.* 129:929-32 (1984)). In some examples, a sample is a skin sample or ocular tissue obtained by excisional biopsy, incisional biopsy, punch biopsy, saucerization biopsy or fine-needle aspiration biopsy. An excisional biopsy excises, or cuts away, the entire growth with a margin of normal surrounding skin or ocular tissue. Generally, an additional wide local excision of normal surrounding skin will be required if the biopsy is positive. The width of the margin will depend on the thickness of the cancer. An incisional biopsy, or core biopsy, removes only a sample of the growth. A punch biopsy removes a small, cylindrical shaped sample of skin or ocular tissue. It can include the epidermis, dermis, and parts of the underlying tissue. A saucerization biopsy removes the entire lesion by cutting under the lesion in a "scoop like" manner, and provides the

practitioner with a complete specimen to better analyze the tumor architecture. A fine-needle aspiration biopsy is done with a very thin needle and syringe. It removes a very small sample of tissue. This type of biopsy can be done on a suspicious mole or skin or eye growth. In addition, it can be done on other deeper tissue, such as nearby lymph nodes or an internal organ, to see if melanoma has spread. It will appreciated that any method of obtaining tissue from a subject can be utilized, and that the selection of the method used will depend upon various factors such as the type of tissue, age of the subject, or procedures available to the practitioner.

In some embodiments, a sample containing melanocytes is a cell and/or tissue lysate. Cell lysate contains many of the proteins and nucleic acids contained in a cell, and include for example, the biomarkers shown in Table 3, 4, 11, or 13. Methods for obtaining a cell lysate are well known in the art and can be found for example in Ausubel *et al.* (In Current Protocols in Molecular Biology, John Wiley & Sons, New York, 1998). In some examples, cells in the sample are lysed or permeabilized in an aqueous solution (for example using a lysis buffer). The aqueous solution or lysis buffer may include detergent (such as sodium dodecyl sulfate) and one or more chaotropic agents (such as formamide, guanidinium HCl, guanidinium isothiocyanate, or urea). The solution may also contain a buffer (for example SSC). In some examples, the lysis buffer includes about 8% to 60% formamide (v/v) about 0.01% to 0.5% SDS, and about 0.5-6X SSC (for example, about 3X SSC). The buffer may optionally include tRNA at about 0.001 to about 2.0 mg/ml or a ribonuclease. The lysis buffer may also include a pH indicator, such as Phenol Red. Cells are incubated in the aqueous solution for a sufficient period of time (such as about 1 minute to about 60 minutes, for example about 5 minutes to about 20 minutes, or about 10 minutes) and at a sufficient temperature (such as about 22°C to about 115°C, for example, about 37°C to about 105°C, or about 90°C to about 100°C) to lyse or permeabilize the cell. In some examples, lysis is performed at about 95°C, for example if the nucleic acid to be detected is RNA. In other examples, lysis is performed at about 105°C, for example if the nucleic acid to be detected is DNA. In some examples, lysis conditions can be such that genomic DNA is not accessible to the probes whereas RNA (for example, mRNA) is, or such that the RNA is destroyed and only the DNA is accessible for probe hybridization. In some examples, the crude cell lysate is used directly without further purification.

Reference Standards

A reference standard also may be referred to as a "control." A control can be a known value or range of values indicative of basal levels or amounts of expression (such as expression of a

biomarker shown in Table 4, 11, or 13) present in a tissue or a cell or populations thereof (such as a normal non-cancerous skin tissue or cell). A control can also be a cellular or tissue control.

Control samples include any suitable sample (*e.g.*, cell, tissue or organ control sample) against which to compare expression of a melanoma biomarker shown in Table 4, 11 or 13, such as the normalization markers shown in Table 3. In some embodiments, the control sample is non-tumor tissue, such as a plurality of non-tumor tissue samples. In one example, non-tumor tissue is tissue known to be benign, such as benign nevus. In some examples, non-tumor tissue includes a skin sample that appears normal, that is it has the absence of nevi, benign lesion, or melanoma. In some examples, the non-tumor tissue is obtained from the same subject, such as non-tumor tissue that is adjacent or even distant from a malignant melanoma. In other examples, the non-tumor tissue is obtained from a healthy control subject or several healthy control subjects. For example, non-tumor tissue can be obtained from a plurality of healthy control subjects (*e.g.*, those not having any cancers, including melanoma, such as samples containing benign nevi from a plurality of such subjects).

In some embodiments, the control sample is known tumor tissue, such as a plurality of known melanoma samples, such as a training set of melanoma (*e.g.*, primary melanoma) samples. Other embodiments involve controls of tissue known to be benign nevi, such as a training set of nevi samples. Training sets of samples (*e.g.*, nevi and melanoma) are useful, in some embodiments, to develop or "train" algorithms (*e.g.*, machine learning algorithms) that distinguish between such sample types.

A difference between a test sample and a control can be an increase or conversely a decrease, for example a decrease or increase in the expression of a biomarker shown in Table 4, 11 or 13. The difference can be a qualitative difference or a quantitative difference, for example a statistically significant difference. In some examples, a difference is an increase or decrease in amount, relative to a control, of at least about 1 %, such as at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, at least about 100%, at least about 150%, at least about 200%, at least about 250%, at least about 300%, at least about 350%, at least about 400%, at least about 500%, or greater than 500%. In some embodiments, the control is a reference value or ranges of values, such as expected expression levels for the biomarkers shown in Table 4, 11, or 13 for a sample(s) known to be primary melanoma(s), or benign nevus(nevi). In other embodiments, a reference value obtained from control samples may be a population central tendency ("CT") (such as a mean (*e.g.*, arithmatic or geometric mean), median, mode or average), or reference range of

values such as plus and/or minus 0.5, 1.0, 1.5 or 2.0 standard deviation(s) around a population CT. For example, one or more reference values can be derived from the average expression values obtained from a group of healthy control subjects (*e.g.*, from a plurality of known benign nevi) or from a group of cancer patients with melanoma (*e.g.*, from a plurality of known malignant nevi).

Sample Analytical Options

In particular examples, the sample to be analyzed, such as a melanocyte-containing sample (*e.g.*, skin biopsy) is or has been fixed. Fixation techniques may vary from site-to-site, country-to-country, investigator-to-investigator, *etc.* (Dissecting the Molecular Anatomy of Tissue, ed. by Emmert-Buck, Gillespie and Chuaqui, New York: Springer-Verlag, 244 pages (2010)) and may affect the integrity of and/or accessibility to the gene product(s) to be detected. Thus, in some disclosed methods involving fixed sample (*e.g.*, methods embodiments with steps for isolating the gene expression product(s), such as PCR or nucleic acid sequencing), RNA recovery (*e.g.*, using reversible cross linking agents, ethanol-based fixatives and/or RNA extraction or purification (in whole or in part)) may be advantageous. Notably, in other representative methods (*e.g.*, involving qNPA) RNA recovery is optional or RNA recovery expressly is not needed. Similarly, tissue conditioning can be used to recover protein gene products from fixed tissue in some method embodiments and, thereby, aid in the detection of such protein products.

The percentage of tumor or suspected tumor (*e.g.*, melanoma) in biological samples may vary; thus, in some disclosed embodiments, at least 5%, at least 10%, at least 25%, at least 50%, at least 75%, at least 80% or at least 90% of the sample area (or sample volume) or total cells in the sample are tumor or suspected tumor (*e.g.*, melanoma). In other examples, samples may be enriched for tumor (or suspected tumor) cells, *e.g.*, by macrodissecting areas or cells from a sample that are or appear to be abnormal (*e.g.*, dysplastic). Optionally, a pathologist or other appropriately trained professional may review the sample (*e.g.*, H&E-stained tissue section) to determine if sufficient abnormality (*e.g.*, suspected tumor) is present in the sample for testing and/or mark the area to be macrodissected. In specific examples, macrodissection of sample to be tested avoids as much as possible necrotic and/or hemorrhagic areas. Samples useful in some disclosed methods will have less than 25%, 15%, 10%, 5%, 2%, or 1% necrosis by sample volume or area or total cells.

Sample load influences the amount and/or concentration of gene product (*e.g.*, one or more of the biomarkers in Table 3, 4, 11, or 13) available for detection. In particular embodiments, at least 1 ng, 10 ng, 100 ng, 1 ug, 10 ug, 100 ug, 500 ug, 1 mg total RNA (*e.g.*, mRNA or miRNA), at least 1 ng, 10 ng, 100 ng, 1 ug, 10 ug, 100 ug, 500 ug, 1 mg total DNA, or at least 0.01 ng, 0.1 ng, 1

ng, 10 ng, 100 ng, 1 ug, 10 ug, 100 ug, 500 ug, or 1 mg total protein is isolated from and/or present in a sample (such as a sample lysate). Some embodiments use tissue samples (*e.g.*, FFPE sectioned skin biopsies) that are at least 3, 5, 8, or 10 μm (*e.g.*, about 3 to about 10 μm) thick and/or at least 0.15, 0.2, 0.5, 1, 1.5, 2, 5 or 10 $cm^2$ in area. The concentration of sample suspended in buffer in some method embodiments is at least 0.006 $cm^2$/ul (*e.g.*, 0.15 $cm^2$ FFPE tissue per 25 uL of buffer (*e.g.*, lysis buffer)).

*Genes and Gene Sets*

Among the innovations disclosed herein are genes (also referred to as biomarkers) and sets of genes, the expression of which (*e.g.*, as measured by mRNA, miRNA or protein expression) is useful in disclosed methods, arrays and kits for distinguishing between benign (*e.g.*, nevi) and malignant (*e.g.*, primary melanoma) melanocyte-containing samples. Also disclosed are genes and gene sets useful as normalizers (*e.g.*, sample-to-sample controls) for nevus and melanoma (*e.g.*, primary melanoma) samples.

In some examples, changes in expression (such as upregulation or downregulation) of at least two different biomarkers from any or all of Table(s) 4, 11 and/or 13 (including, without limitation, genes combinations in Tables 6, 8 or 14), for example normalized to at least one normalization marker (such as one or more of those in Table 3), can be used as specific markers of nevus or melanoma or as markers of the transition between a benign nevus and a primary melanoma. Such markers are useful for a variety of methods and compositions as describe in more detail in this disclosure and, for example, include methods for diagnosing a subject, such as a human subject, as having a benign nevus or as having melanoma, by measuring or detecting expression levels of two or more different biomarkers from any or all of Table(s) 4, 11 and/or 13 (including, *e.g.*, genes combinations in Tables 6, 8 or 14). In one example, the human subject is at risk for developing melanoma.

This disclosure has identified significantly differentially expressed (SDE) genes in melanocyte-containing samples (populations) of interest (*e.g.*, nevi vs. melanoma samples), and exemplary combinations of the identified SDE genes were analyzed to identify combinations of those SDE genes having predictive value to permit characterization of a melanocyte-containing sample as a benign nevus or primary melanoma (see, *e.g.*, Example 2, 3 and 4). Although particular combinations of identified SDE genes are described herein, one ordinarily skilled in the art will appreciate that this disclosure now enables the identification of other combinations of the

SDE genes shown in Table(s) 4, 11 and/or 13 that will robustly characterize a sample as a nevus or melanoma. For example, any non-repeating combination of biomarkers in any or all of Table(s) 4, 11 and/or 13 in which all predictor Xn variables (expression value for the selected biomarker) have a variance inflation factor (VIF) less than 10 are expected to have a useful predictive value for differentiating between samples from benign nevi versus those from primary melanoma and, accordingly, are contemplated by this disclosure. Additionally, nevi-melanoma classifiers of any combination of genes in Table(s) 4, 11 and/or 13 may be tested for acceptable classification performance (*e.g.*, misclassification of fewer than 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8% or 10% of samples, or classification accuracy of greater than or equal to 75%, 80%, 85%, 90%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99%) using any of the methods disclosed herein (*e.g.*, AUC) or commonly known in the art.

Particular method embodiments described throughout this disclosure include determining in a sample (*e.g.*, a skin sample) obtained from a subject, an expression level (such as a nucleic acid or protein level) of at least two different (*i.e.*, no repeated) biomarkers selected from any one or more (a) - (r) below and, in some cases, at least one normalization biomarker (such as listed in Table 3). Similarly, particular compositions embodiments described throughout this disclosure may include specific binding agents (*e.g.*, probes, primers, aptamers, antibodies, *etc.*) that can be used to specifically measure an expression level (such as a nucleic acid or protein level) of at least two different (*i.e.*, no repeated) biomarkers selected from any one or more (a) - (r) below and, in some cases, at least one normalization biomarker (such as listed in Table 3). In some examples, as applicable, an expression level (such as a nucleic acid or protein level) for at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, at least 24, at least 25, at least 26, at least 27, at least 28, at least 29, at least 30, at least 31, or all of the biomarkers listed in any one of (a) - (r) (such as 2 to 20, 2 to 10, 4 to 10, 4 to 15, or 2 to 5 of the biomarkers listed) is determined in the sample or can be specifically detected using a disclosed composition (*e.g.*, array or kit). In other examples, an expression level (such as a nucleic acid or protein level) for at least two different (*i.e.*, no repeated) biomarkers selected from any one or more (a) - (r) below are at least 50%, at least 75%, at least 80%, at least 90%, at least 95%, or at least 98% of the plurality of genes listed in the particular group (*e.g.*, Table(s) 4, 11 and/or 13) from which the biomarkers are selected.

(a) Genes described in Table 4 (*i.e.*, NR4A1, B4GALT1, SAT1, TP53, TADA3, BRAF, TFRC, RUNX1, SOCS3, PDLIM7, SP100, PIP4K2A, SOX4, PDIA4, MCM6, CTNNB1, RPL37A,

GNAS, TGFB1, PPIA, PTEN, MAGED2, 1PRAME, GALNTL1, MAGEA2, TEX13A, CREBBP, TPSAB1, CDK2, STAT2, SQSTM1, and B2M); and/or

(b) Genes described in Table 11 (*i.e.*, B4GALT1, BAX, MAGEA2, NR4A1, PDIA4, PRAME, RUNX1, SOCS3, SAT1, PDLIM7, BIRC5, HIF1A, MET, MAGEC2, ERCC1, POLR2J3, LDHA, PICALM, ZFYVE16, and BEST1), and/or

(c) Genes described in Table 13 (*i.e.*, genes expressing the products hsa.miR.122, hsa.miR.1291, hsa.miR.191, hsa.miR.19b, hsa.miR.200a, hsa.miR.200c, hsa.miR.203, hsa.miR.205, hsa.miR.21, hsa.miR.23b, hsa.miR.29c, hsa.miR.342.3p, hsa.miR.375, hsa.miR.665, hsa.miR.1304, hsa.miR.142.5p, hsa.miR.1254, hsa.let.7a, hsa.miR.140.5p, and hsa.miR.183); and/or

(d) NR4A1, B4GALT1, SOX4, SQSTM1, B2M, TFRC, TP53, GALNTL1, CREBBP, SOCS3 and CTNNB1; and/or

(e) NR4A1, B4GALT1, SOX4, SQSTM1, B2M, TFRC, TP53, CREBBP, SOCS3, RPL37A, SAT1, BRAF, and TPSAB1; and/or

(f) NR4A1, B4GALT1, SOX4, SQSTM1, B2M, TFRC, TP53, CREBBP, and SOCS3; and/or

(g) NR4A1, B4GALT1, SOX4, SQSTM1, B2M, TFRC, TP53, SOCS3, and BRAF; and/or

(h) NR4A1, B4GALT1, SOX4, SQSTM1, B2M, TFRC, TP53, CREBBP, SOCS3, and BFAF; and/or

(i) MAGEA2, PRAME, PDIA4, NR4A1, PDLIM7, B4GALT1, SAT1, RUNX1, and SOCS3; and/or

(j) Any gene set described in Table 6; and/or

(k) Any gene set described in Table 8; and/or

(l) Any gene set described in Table 14; and/or

(m) Any of the specific combinations paired in square brackets ([...]) below:
[NR4A1,B4GALT1], [NR4A1,SOX4], [NR4A1,SQSTM1], [NR4A1,B2M], [NR4A1,TFRC], [NR4A1,TP53], [NR4A1,CREBBP], [NR4A1,SOCS3], [NR4A1,BRAF], [B4GALT1,SOX4], [B4GALT1,SQSTM1], [B4GALT1,B2M], [B4GALT1,TFRC], [B4GALT1,TP53], [B4GALT1,CREBBP], [B4GALT1,SOCS3], [B4GALT1,BRAF], [SOX4,SQSTM1], [SOX4,B2M], [SOX4,TFRC], [SOX4,TP53], [SOX4,CREBBP], [SOX4,SOCS3], [SOX4,BRAF], [SQSTM1,B2M], [SQSTM1,TFRC], [SQSTM1,TP53], [SQSTM1,CREBBP], [SQSTM1,SOCS3], [SQSTM1,BRAF], [B2M,TFRC], [B2M,TP53], [B2M,CREBBP], [B2M,SOCS3], [B2M,BRAF], [TFRC,TP53], [TFRC,CREBBP],

[TFRC,SOCS3], [TFRC,BRAF], [TP53,CREBBP], [TP53,SOCS3], [TP53,BRAF], [CREBBP,SOCS3], [CREBBP,BRAF], and [SOCS3,BRAF]; and/or

(n) Combinations of three (or four) described by any of the pairs in (m) in combination with one (or two) other non-repetitive genes from the list of NR4A1, B4GALT1, SOX4, SQSTM1, B2M, TFRC, TP53, CREBBP, SOCS3, and BRAF; and/or

(o) Any of the specific combinations paired in square brackets ([...]) below: [MAGEA2,PRAME], [MAGEA2,PDIA4], [MAGEA2,NR4A1], [MAGEA2,PDLIM7], [MAGEA2,B4GALT1], [MAGEA2,SAT1], [MAGEA2,RUNX1], [MAGEA2,SOCS3], [PRAME,PDIA4], [PRAME,NR4A1], [PRAME,PDLIM7], [PRAME,B4GALT1], [PRAME,SAT1], [PRAME,RUNX1], [PRAME,SOCS3], [PDIA4,NR4A1], [PDIA4,PDLIM7], [PDIA4,B4GALT1], [PDIA4,SAT1], [PDIA4,RUNX1], [PDIA4,SOCS3], [NR4A1,PDLIM7], [NR4A1,B4GALT1], [NR4A1,SAT1], [NR4A1,RUNX1], [NR4A1,SOCS3], [PDLIM7,B4GALT1], [PDLIM7,SAT1], [PDLIM7,RUNX1], [PDLIM7,SOCS3], [B4GALT1,SAT1], [B4GALT1,RUNX1], [B4GALT1,SOCS3], [SAT1,RUNX1], [SAT1,SOCS3], or [RUNX1,SOCS3]; and/or

(p) Combinations of three (or four) described by any of the pairs in (o) in combination with one (or two) other non-repetitive gene(s) from the list of MAGEA2, PRAME, PDIA4, NR4A1, PDLIM7, B4GALT1, SAT1, RUNX1, and SOCS3; and/or

(q) Any of the specific combinations paired in square brackets ([...]) below ("hsa" has been removed in each case but is intended as part of the identifier): [miR.122, miR.1291], [miR.122, miR.191], [miR.122, miR.19b], [miR.122, miR.200a], [miR.122, miR.200c], [miR.122, miR.203], [miR.122, miR.205], [miR.122, miR.21], [miR.122, miR.23b], [miR.122, miR.29c], [miR.122, miR.342.3p], [miR.122, miR.375], [miR.122, miR.665], [miR.122, miR.1304], [miR.122, miR.142.5p], [miR.122, miR.1254], [miR.122, let.7a], [miR.122, miR.140.5p], [miR.122, miR.183], [miR.1291, miR.191], [miR.1291, miR.19b], [miR.1291, miR.200a], [miR.1291, miR.200c], [miR.1291, miR.203], [miR.1291, miR.205], [miR.1291, miR.21], [miR.1291, miR.23b], [miR.1291, miR.29c], [miR.1291, miR.342.3p], [miR.1291, miR.375], [miR.1291, miR.665], [miR.1291, miR.1304], [miR.1291, miR.142.5p], [miR.1291, miR.1254], [miR.1291, let.7a], [miR.1291, miR.140.5p], [miR.1291, miR.183], [miR.191, miR.19b], [miR.191, miR.200a], [miR.191, miR.200c], [miR.191, miR.203], [miR.191, miR.205], [miR.191, miR.21], [miR.191, miR.23b], [miR.191, miR.29c], [miR.191, miR.342.3p], [miR.191, miR.375], [miR.191, miR.665], [miR.191, miR.1304], [miR.191, miR.142.5p], [miR.191,

miR.1254], [miR.191, let.7a], [miR.191, miR.140.5p], [miR.191, miR.183], [miR.19b, miR.200a], [miR.19b, miR.200c], [miR.19b, miR.203], [miR.19b, miR.205], [miR.19b, miR.21], [miR.19b, miR.23b], [miR.19b, miR.29c], [miR.19b, miR.342.3p], [miR.19b, miR.375], [miR.19b, miR.665], [miR.19b, miR.1304], [miR.19b, miR.142.5p], [miR.19b, miR.1254], [miR.19b, let.7a], [miR.19b, miR.140.5p], [miR.19b, miR.183], [miR.200a, miR.200c], [miR.200a, miR.203], [miR.200a, miR.205], [miR.200a, miR.21], [miR.200a, miR.23b], [miR.200a, miR.29c], [miR.200a, miR.342.3p], [miR.200a, miR.375], [miR.200a, miR.665], [miR.200a, miR.1304], [miR.200a, miR.142.5p], [miR.200a, miR.1254], [miR.200a, let.7a], [miR.200a, miR.140.5p], [miR.200a, miR.183], [miR.200c, miR.203], [miR.200c, miR.205], [miR.200c, miR.21], [miR.200c, miR.23b], [miR.200c, miR.29c], [miR.200c, miR.342.3p], [miR.200c, miR.375], [miR.200c, miR.665], [miR.200c, miR.1304], [miR.200c, miR.142.5p], [miR.200c, miR.1254], [miR.200c, let.7a], [miR.200c, miR.140.5p], [miR.200c, miR.183], [miR.203, miR.205], [miR.203, miR.21], [miR.203, miR.23b], [miR.203, miR.29c], [miR.203, miR.342.3p], [miR.203, miR.375], [miR.203, miR.665], [miR.203, miR.1304], [miR.203, miR.142.5p], [miR.203, miR.1254], [miR.203, let.7a], [miR.203, miR.140.5p], [miR.203, miR.183], [miR.205, miR.21], [miR.205, miR.23b], [miR.205, miR.29c], [miR.205, miR.342.3p], [miR.205, miR.375], [miR.205, miR.665], [miR.205, miR.1304], [miR.205, miR.142.5p], [miR.205, miR.1254], [miR.205, let.7a], [miR.205, miR.140.5p], [miR.205, miR.183], [miR.21, miR.23b], [miR.21, miR.29c], [miR.21, miR.342.3p], [miR.21, miR.375], [miR.21, miR.665], [miR.21, miR.1304], [miR.21, miR.142.5p], [miR.21, miR.1254], [miR.21, let.7a], [miR.21, miR.140.5p], [miR.21, miR.183], [miR.23b, miR.29c], [miR.23b, miR.342.3p], [miR.23b, miR.375], [miR.23b, miR.665], [miR.23b, miR.1304], [miR.23b, miR.142.5p], [miR.23b, miR.1254], [miR.23b, let.7a], [miR.23b, miR.140.5p], [miR.23b, miR.183], [miR.29c, miR.342.3p], [miR.29c, miR.375], [miR.29c, miR.665], [miR.29c, miR.1304], [miR.29c, miR.142.5p], [miR.29c, miR.1254], [miR.29c, let.7a], [miR.29c, miR.140.5p], [miR.29c, miR.183], [miR.342.3p, miR.375], [miR.342.3p, miR.665], [miR.342.3p, miR.1304], [miR.342.3p, miR.142.5p], [miR.342.3p, miR.1254], [miR.342.3p, let.7a], [miR.342.3p, miR.140.5p], [miR.342.3p, miR.183], [miR.375, miR.665], [miR.375, miR.1304], [miR.375, miR.142.5p], [miR.375, miR.1254], [miR.375, let.7a], [miR.375, miR.140.5p], [miR.375, miR.183], [miR.665, miR.1304], [miR.665, miR.142.5p], [miR.665, miR.1254], [miR.665, let.7a], [miR.665, miR.140.5p], [miR.665, miR.183], [miR.1304, miR.142.5p], [miR.1304, miR.1254], [miR.1304, let.7a], [miR.1304, miR.140.5p], [miR.1304, miR.183],

[miR.142.5p, miR.1254], [miR.142.5p, let.7a], [miR.142.5p, miR.140.5p], [miR.142.5p, miR.183], [miR.1254, let.7a], [miR.1254, miR.140.5p], [miR.1254, miR.183], [let.7a, miR.140.5p], [let.7a, miR.183], or [miR.140.5p, miR.183]; and/or

(r) Combinations of three (or four) described by any of the pairs in (q) in combination with one (or two) other non-repetitive miRNA(s) from the list of hsa.miR.122, hsa.miR.1291, hsa.miR.191, hsa.miR.19b, hsa.miR.200a, hsa.miR.200c, hsa.miR.203, hsa.miR.205, hsa.miR.21, hsa.miR.23b, hsa.miR.29c, hsa.miR.342.3p, hsa.miR.375, hsa.miR.665, hsa.miR.1304, hsa.miR.142.5p, hsa.miR.1254, hsa.let.7a, hsa.miR.140.5p, or hsa.miR.183.

Particular method embodiments include normalizing expression of the disease-specific biomarker(s) (*e.g.*, see (a) - (r) above) to at least one normalization biomarker. As discussed in further detail elsewhere in this disclosure, normalization is a step included in some method embodiments that is useful to control for certain types of confounding variability in gene expression values. Adjusting the values of all disease-specific variables to the expression of specified normalization biomarkers (*e.g.*, by division or subtraction) is one, non-limiting way to normalize such disease-specific variables. As a general rule, a specified normalization biomarker has no statistically significant difference in expression between the sample types of interest (such as between nevi and primary melanoma sample types). Exemplary normalization biomarkers for nevi and melanoma samples are listed in Table 3. Some disclosed methods contemplate normalizing disease-specific biomarker (see, *e.g.*, Table(s) 4, 11, and/or 13) expression to an expression level for at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, or all of the normalization biomarker(s) listed in Table 3, or as selected from the group consisting of (i) MFI2, RAP2B, BMP1, NCOR2, RPS6KB2 and SDHA (ii) BMP-1, MFI2, NCOR2, and RAP2b; or (iii) RPS6KB2 and SDHA. While illustrative normalization biomarkers are specified here, other methods of normalization useful in the disclosed methods are discussed below.

***Detecting Gene Expression***

Disclosed methods further involve detecting the expression of the genes discovered herein (see Table(s) 4, 11 and/or 13) that distinguish benign (*e.g.*, nevi) from malignant (*e.g.*, primary melanoma) melanocyte-containing samples, or are suitable for normalizating expression levels in such sample types (see Table 3). A variety of techniques are (or may become) available for measuring gene expression in a sample of interest. However, the disclosure is not limited to particular methods of obtaining, measuring, or detecting gene expression. Many such techniques involve detecting the products of the genes (*e.g.*, nucleic acids (such as mRNA or miRNA) and/or protein) expressed in such samples. It may also be (or become) possible to directly detect the

activity of a gene or of chromosomal DNA (*e.g.*, transcription rate) independent of measuring its resultant gene products and such techniques also are useful in methods disclosed herein.

Gene expression levels can be determined in the disclosed methods using a solution-based (*i.e., ex situ*) assay, such as PCR or a nuclease protection assay or nucleic acid sequencing. In other examples, expression levels are determined or detected using an *in situ* assay, for example using immunohistochemistry or in situ hybridization.

*Detecting Nucleic-acid Gene Products*

Nucleic-acid gene products are, as the name suggests, products of gene expression that are nucleic acids. Exemplary nucleic acids whose expression can be detected include DNA or RNA, such as cDNA, protein-coding RNA (*e.g.*, mRNA) or non-coding RNA (e.g., miRNA or lncRNA). In a particular exmaples, the method includes detecting mRNA expression, miRNA expression, or both. Base pairing between complementary strands of RNA or DNA (*i.e.*, nucleic acid hybridization) forms all or part of the basis for a large representative class of techniques for detecting nucleic-acid gene products. Other representative detection techniques involve nucleic acid sequencing, which may or may not involve hybridization steps and/or bioinformatics steps (*e.g.*, to associate nucleic acid sequence information to its corresponding gene). These and other methods of detecting nucleic acids are known in the art and, while representative techniques are described herein, this disclosure is not intended to be limited to particular methods of nucleic acid detection.

In some embodiments of the disclosed methods, determining the level of gene expression in a melanocyte-containing sample includes detecting two or more nucleic acids shown in Table(s) 4, 11, and/or 13 (and in some examples also one or more nucleic acids shown in Table 3), for example by determining the relative or actual amounts of such nucleic acids in the sample. Exemplary nucleic acids include DNA or RNA, such as cDNA, miRNA, or mRNA.

The level of expression of nucleic acid molecules can be detected or measured using, for instance, *in vitro* nucleic acid amplification and/or nucleic acid hybridization. The results of such detection methods can be quantified, for instance by determining the amount of hybridization or the amount of amplification. Thus, in some examples, determining the expression level of a biomarker (such as those in Table(s) 3, 4, 11, and/or 13, individually or in any combination, including the combinations in Tables 6, 8 or 14) in the methods provided herein can include contacting the sample with a plurality of nucleic acid probes (such as a nuclease protection probe, NPP) or paired amplification primers, wherein each probe or paired primers is/are specific and complementary to one of the least two, non-repeated biomarkers in Table(s) 4, 11, and/or 13, under conditions that

permit the plurality of nucleic acid probes or paired primers to hybridize to its/their complementary at least two biomarkers in Table(s) 4, 11, and/or 13. In one example, the method can also include after contacting the sample with the plurality of nucleic acid probes (such as NPPs), contacting the sample with a nuclease that digests single-stranded nucleic acid molecules.

Optional Nucleic Acid Isolation

In some examples, nucleic acids are isolated or extracted from the melanocyte-containing sample prior to contacting such nucleic acids in the sample with a complementary nucleic acid probe or primer and/or otherwise detecting such nucleic acids in the sample. Nucleic acids (such as RNA (*e.g.*, mRNA or miRNA) or DNA) can be isolated from the sample according to any of a number of methods. Representative methods of isolation and purification of nucleic acids are described in detail in Chapter 3 of *Laboratory Techniques in Biochemistry and Molecular Biology: Hybridization With Nucleic Acid Probes, Part I. Theory and Nucleic Acid Preparation*, P. Tijssen, ed. Elsevier, N.Y. (1993). Representative methods for RNA (*e.g.*, mRNA or miRNA) extraction similarly are well known in the art and are disclosed in standard textbooks of molecular biology, including Ausubel *et al.*, *Current Protocols of Molecular Biology*, John Wiley and Sons (1997).

Specific methods can include isolating total nucleic acid from a sample using, for example, an acid guanidinium-phenol-chloroform extraction method and/or isolating polyA+ mRNA by oligo dT column chromatography or by (dT)n magnetic beads (see, for example, Sambrook et al, *Molecular Cloning: A Laboratory Manual* (2nd ed.), Vols. 1-3, Cold Spring Harbor Laboratory, (1989), or *Current Protocols in Molecular Biology*, F. Ausubel et al., ed. Greene Publishing and Wiley-Interscience, N.Y. (1987)). In other examples, nucleic acid isolation can be performed using purification kit, buffer set and protease from commercial manufacturers, such as QIAGEN® (Valencia, CA), according to the manufacturer's instructions. For example, total RNA from cells (such as those obtained from a subject) can be isolated using QIAGEN® RNeasy mini-columns. Other commercially available nucleic acid isolation kits include MASTERPURE® Complete DNA and RNA Purification Kit (EPICENTRE® Madison, Wis.), and Paraffin Block RNA Isolation Kit (Ambion, Inc.). Total RNA from tissue samples can be isolated using RNA Stat-60 (Tel-Test). RNA prepared from tumor or other biological sample can be isolated, for example, by cesium chloride density gradient centrifugation. Methods for RNA extraction from paraffin embedded tissues are disclosed, for example, in Rupp and Locker, *Biotechniques* 6:56-60 (1988), and De Andres *et al.*, *Biotechniques* 18:42-44 (1995).

After isolation or extraction of nucleic acids (*e.g.*, RNA (such as mRNA or miRNA) or DNA) from a sample, any of a number of optional other steps may be performed to prepare such

nucleic acids for detection, including measuring the concentration of the isolated nucleic acid, repair (or recovery) of degraded or damaged RNA, RNA reverse transcription, and/or amplification of RNA or DNA.

In other examples, a sample (*e.g.*, FFPE melanocyte-containing tissue sample) is suspended in a buffer (*e.g.*, lysis buffer) and nucleic acids (such as RNA or DNA) present in the suspended sample are not isolated or extracted (*e.g.*, purified in whole or in part) from such suspended sample and are contacted in such suspension with one or more complementary nucleic acid probe(s) (*e.g.*, nuclease protection probes); thereby, eliminating a need for isolation or extraction of nucleic acids (*e.g.*, RNA) from the sample. This embodiment is particularly advantageous where the nucleic acids (such as RNA or DNA) present in the suspended sample are crosslinked or fixed to cellular structures and are not readily isolatable or extractable. Relatively short (*e.g.*, less than 100 base pairs, such as 75-25 base pairs or 50-25 base pairs) probes for which no extension of such probe is required for detection are useful in some non-extraction method embodiments. An ordinarily skilled artisan will appreciate that methods requiring probe extension (*e.g.*, PCR or primer extension) are not reliable where the nucleic acid template (*e.g.*, RNA) for such extension is degraded or otherwise inaccessible. Specific methods (*e.g.*, qNPA) for detecting nucleic acids (*e.g.*, RNA) in a sample without prior extraction of such nucleic acids are described in detail elsewhere herein.

Nucleic Acid Hybridization

In some examples, determining the expression level of a disclosed biomarker (such as those in Table(s) 4, 11, and/or 13) or normalization biomarker (*e.g.*, Table 3) in the methods provided herein can include contacting the sample with a plurality of nucleic acid probes (such as a nuclease protection probe, NPP, or adjoining ligatable probes) or paired amplification primers, wherein each probe (or set of ligatable probes) or paired primers in the plurality is/are specific and complementary to one of at least two biomarkers in Table(s) 4, 11, and/or 13 or a or normalization biomarker in Table 3, under conditions that permit the plurality of nucleic acid probes or paired primers to hybridize to its/their complementary biomarker in Table(s) 4, 11, and/or 13. In one example, the method can also include after contacting the sample with the plurality of nucleic acid probes (such as NPPs), contacting the sample with a nuclease that digests single-stranded nucleic acid molecules. In other examples, each of the at least two biomarkers in Table(s) 4, 11, and/or 13, or a or normalization biomarker in Table 3, is contacted with a "probe set" that consists of multiple (*e.g.*, 2, 3, 4, 5, or 6) probes specific for each such biomarker, which design can be useful, for

example, to increase the signal obtained from such gene product or to detect multiple variants of the same gene product.

In some examples, variable (*e.g.*, Table(s) 4, 11, and/or 13) or normalization (*e.g.*, Table 3) nucleic acids are detected by nucleic acid hybridization. Nucleic acid hybridization involves providing a denatured probe and target nucleic acid (*e.g.*, those in Table(s) 4, 11, and/or 13) under conditions where the probe and its complementary target can form stable hybrid duplexes through complementary base pairing. In some examples, the nucleic acids that do not form hybrid duplexes are then removed (*e.g.*, washed away, digested by nuclease or physically removed) leaving the hybridized nucleic acids to be detected, typically through detection of an (directly or indirectly) attached detectable label. In specific examples, nucleic acids that do not form hybrid duplexes, such as any excess probe that does not hybridize to its respective target, and the regions of the target sequence that are not complementary to the probes, can be digested away by addition of nuclease, leaving just the hybrid duplexes of target sequence of complementary probe.

It is generally recognized that nucleic acids are denatured by increasing the temperature and/or decreasing the salt concentration of the buffer containing the nucleic acids. Under low stringency conditions (*e.g.*, low temperature and/or high salt) hybrid duplexes (*e.g.*, DNA:DNA, RNA:RNA, or RNA:DNA) will form even where the annealed sequences are not perfectly complementary. Thus, specificity of hybridization is reduced at lower stringency. Conversely, at higher stringency (*e.g.*, higher temperature or lower salt) successful hybridization requires fewer mismatches. One of skill in the art will appreciate that hybridization conditions can be designed to provide different degrees of stringency. The strength of hybridization can be increased without lowering the stringency of hybridization, and thus the specificity of hybridization can be maintained in a high stringency buffer, by including unnatural bases in the probes, such as by including locked nucleic acids or peptide nucleic acids.

In general, there is a tradeoff between hybridization specificity (stringency) and signal intensity. Thus, in one embodiment, the wash is performed at the highest stringency that produces consistent results and that provides a signal intensity greater than approximately 10% of the background intensity. Thus, the hybridization complexes (*e.g.*, as captured on an array surface) may be washed at successively higher stringency solutions and read between each wash. Analysis of the data sets thus produced will reveal a wash stringency above which the hybridization pattern is not appreciably altered and which provides adequate signal for the particular oligonucleotide probes of interest.

Changes in expression of a nucleic acid and/or the presence of nucleic acid detected by these methods for instance can include increases or decreases in the level (amount) or functional activity of such nucleic acids, their expression or translation into protein, or in their localization or stability. An increase or a decrease, for example relative to a normalization biomarker (see, *e.g.*, Table 3), can be, for example, at least a 1-fold, at least a 2-fold, or at least a 5-fold, such as about a 1-fold, 2-fold, 3-fold, 4-fold, 5-fold, change (increase or decrease) in the expression of and/or the presence of a particular nucleic acid, such as a nucleic acid corresponding to the biomarker shown in any of Table(s) 4, 11, and/or 13. In multiplexed method embodiments, the relative expression of non-normalizer genes (*e.g.*, variable genes; for example, Table(s) 4, 11, and/or 13) also can be compared; particularly, when each such gene has been similarly normalized (*e.g.*, to the expression of one or more co-detected normalizer genes; for example see Table 3). Hence, the normalized expression of one variable gene may be at least at least a 1-fold, at least a 2-fold, or at least a 5-fold, such as about a 1-fold, 1.5-fold, 2-fold, 3-fold, 4-fold, 5-fold higher or lower than the normalized expression of another variable gene.

Gene expression is measured using a multiplexed methodology and/or high throughput methodology in some embodiments. In multiplexed methods, a plurality of measurements (*e.g.*, gene expression measurements) is made in a single sample. Various technologies have evolved that permit the monitoring of large numbers of genes in a single sample (*e.g.*, traditional microarrays, multiplexed PCR, serial analysis of gene expression (SAGE; *e.g.*, U.S. Pat. No. 5,866,330), multiplex ligation-dependent probe amplification (MLPA), high-throughput sequencing, labeled bead-based technology (*e.g.*, U.S. Pat. Nos. 5,736,330 and 6,449,562), digital molecular barcoding technology (*e.g*, U.S. Pat. No. 7,473,767). In high-throughput methods, gene expression in multiple samples is measured contemporaneously. High-throughput methods can also be multiplexed (*i.e.*, contemporaneously detecting multiple genes in each of multiple samples).

In some embodiments, expression levels of one or more biomarkers (such as two or more of those in Table(s) 4, 11, and/or 13 (*e.g.*, any genes combination in Tables 6, 8 or 14) and/or at least one in Table 3) are determined contemporaneously in a single melanocyte-containing sample or in a plurality of melanocyte-containing samples (such as samples from different subjects). In one example, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, or, as applicable, at least 20, at least 21, at least 22, at least 23, at least 24, at least 25, at least 26, at least 27, at least 28, at least 29, at least 30, at least 31, or all of the biomarkers listed in Table(s) 4, 11, and/or 13 (such as 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, or,

as applicable, 21, 22, 23, 24 , 25, 26, 27, 28, 29, 30, 31, or all of the biomarkers in Table(s) 4, 11, and/or 13), or, *e.g.*, any of the gene combinations in Table 6, 8 or 14, can be detected contemporaneously in the same sample or in a plurality of samples, and in some examples, at least 2, at least 3, at least 4, at least 5, or all 6 of the normalization biomarkers listed in Table 3 (or other normalization biomarker(s) identified with the methods provided herein) are detected contemporaneously, for example contemporaneously with the at least two biomarkers in Table(s) 4, 11, and/or 13. The plurality of samples can be from multiple different subjects and/or be multiple samples from the same subject, such as at least 2 different samples (*e.g.*, from at least 2 different subjects and/or from different areas of the same subject's tumor or body). In some examples, at least at least 2, at least 5, at least 10, at least 20, at least 50, at least 100, at least 500, at least 1000, at least 2000, at least 5000, or even at least 10,000 melanocyte-containing samples are analyzed contemporaneously (such as 10 to 100, 10 to 1000, 100 to 1000, 100 to 5000, or 1000 to 10,000 melanocyte-containing samples are analyzed contemporaneously).

This disclosure also includes methods utilizing integrated systems for high-throughput screening. The systems typically include a robotic armature that transfers fluid from a source to a destination, a controller that controls the robotic armature, a detector, a data storage unit that records detection, and an assay component such as a microtiter plate, for example including one or more programming linkers or one that includes one or more oligonucleotides that can directly hybridize to a target (such as two or more of the biomarkers in Table(s) 4, 11, and/or 13, and one or more of the normalization markers in Table 3).

Arrays are one useful (non-limiting) set of tools for multiplex detection of gene expression. An array is a systematic arrangement of elements (*e.g.*, analyte capture reagents (such as, target-specific oligonucleotide probes, aptamers, or antibodies)) where a set of values (*e.g.*, gene expression values) can be associated with an identification key. The arrayed elements may be systematically identified on a single surface (*e.g.*, by spatial mapping or by differential tagging), using separately identifiable surfaces (*e.g.*, flow channels or beads), or by a combination thereof.

Other examples of methods and assay systems that can be used to detect the disclosed biomarkers are high throughput assay techniques disclosed in International Patent Publication Nos. WO 2003/002750 and WO 2008/121927, WO 1999/032663, WO 2000/079008, WO/2000/037684, and WO 2000/037683 and U.S. Patent Nos. 6,232,066, 6,458,533,6,238,869, and 7,659,063.

In some array embodiments, nucleic acid probes (such as oligonucleotides), which are designed to capture (directly or indirectly) one or more products of the genes shown in Table(s) 3,

4, 11, and/or 13), are plated, or arrayed, on a microchip substrate. For example, the array can include oligonucleotides complementary to at least two of the genes shown in Table(s) 3, 4, 11, and/or 13 (such as at least 3, at least 5, at least 10, at least 20, or all of such genes, or any of the genes combinations in Tables 6, 8 or 14 or as otherwise disclosed herein) and, optionally, at least one of the genes shown in Table 3. In other examples, the array can include oligonucleotides complementary to a portion of a nuclease protection probe that is complementary to a product of at least two of the genes shown in Table(s) 3, 4, 11, and/or 13 (such as at least 3, at least 5, at least 10, at least 20, or all of such genes, or any of the genes combinations in Tables 6, 8 or 14 or as otherwise disclosed herein) and, optionally, at least one of the genes shown in Table 3.

The arrayed sequences are then hybridized with isolated nucleic acids (such as cDNA, miRNA or mRNA) from the test sample (*e.g.*, melanocyte-containing sample obtained from a subject, whose characterization as benign nevus or malignant melanoma (*e.g.*, primary melanoma) is desired). In one example, the isolated nucleic acids from the test sample are labeled, such that their hybridization with the specific complementary oligonucleotide on the array can be determined. Alternatively, the test sample nucleic acids are not labeled, and hybridization between the oligonucleotides on the array and the target nucleic acid is detected using a sandwich assay, for example using additional oligonucleotides complementary to the target that are labeled.

In one embodiment, the hybridized nucleic acids are detected by detecting one or more labels attached to the sample nucleic acids or attached to a nucleic acid probe that hybridizes directly or indirectly to the target nucleic acids. The labels can be incorporated by any of a number of methods. In one example, the label is simultaneously incorporated during the amplification step in the preparation of the sample nucleic acids. Thus, for example, polymerase chain reaction (PCR) with labeled primers or labeled nucleotides will provide a labeled amplification product. In one embodiment, transcription amplification using a labeled nucleotide (such as fluorescein-labeled UTP and/or CTP) incorporates a label into the transcribed nucleic acids.

Detectable labels suitable for use in embodiments throughout this disclosure include any composition detectable by spectroscopic, photochemical, biochemical, immunochemical, electrical, optical or chemical means. Useful labels include biotin for staining with labeled streptavidin conjugate, magnetic beads (for example DYNABEADS$^{TM}$), fluorescent dyes (for example, fluorescein, Texas red, rhodamine, green fluorescent protein, and the like), chemiluminescent markers, radiolabels (for example, $^{3}H$, $^{125}I$, $^{35}S$, $^{14}C$, or $^{32}P$), enzymes (for example, horseradish peroxidase, alkaline phosphatase and others commonly used in an ELISA), and colorimetric labels such as colloidal gold or colored glass or plastic (for example, polystyrene, polypropylene, latex,

*etc.*) beads. Patents teaching the use of such labels include U.S. Patent No. 3,817,837; U.S. Patent No. 3,850,752; U.S. Patent No. 3,939,350; U.S. Patent No. 3,996,345; U.S. Patent No. 4,277,437; U.S. Patent No. 4,275,149; and U.S. Patent No. 4,366,241. In some embodiments, labels are attached by spacer arms of various lengths to reduce potential steric hindrance.

Means of detecting such labels are also well known. Thus, for example, radiolabels may be detected using photographic film or scintillation counters, fluorescent markers may be detected using a photodetector to detect emitted light. Enzymatic labels are typically detected by providing the enzyme with a substrate and detecting the reaction product produced by the action of the enzyme on the substrate, and colorimetric labels are detected by simply visualizing the colored label.

The label may be added to the target (sample) nucleic acid(s) prior to, or after, the hybridization. So-called "direct labels" are detectable labels that are directly attached to or incorporated into the target (sample) nucleic acid prior to hybridization. In contrast, so-called "indirect labels" are joined to the hybrid duplex after hybridization. Often, the indirect label is attached to a binding moiety that has been attached to the target nucleic acid prior to the hybridization. Thus, for example, the target nucleic acid may be biotinylated before the hybridization. After hybridization, an avidin-conjugated fluorophore will bind the biotin bearing hybrid duplexes providing a label that is easily detected (see *Laboratory Techniques in Biochemistry and Molecular Biology,* Vol. 24: *Hybridization With Nucleic Acid Probes*, P. Tijssen, ed. Elsevier, N.Y., 1993).

*In situ* hybridization (ISH), such as chromogenic *in situ* hybridization (CISH) or silver *in situ* hybridization (SISH), is an exemplary method for detecting and comparing expression of genes of interest (such as those in Table(s) 3, 4, 11, and/or 13). ISH is a type of hybridization that uses a complementary nucleic acid to localize one or more specific nucleic acid sequences in a portion or section of tissue (*in situ*), or, if the tissue is small enough, in the entire tissue (whole mount ISH). RNA ISH can be used to assay expression patterns in a tissue, such as the expression of the biomarkers in Table(s) 4, 11, and/or 13. Sample cells or tissues may be treated to increase their permeability to allow a probe, such as a probe specific for one or more of the biomarkers in Table(s) 4, 11, and/or 13, to enter the cells. The probe is added to the treated cells, allowed to hybridize at pertinent temperature, and excess probe is washed away. A complementary probe is labeled with a detectable label, such as a radioactive, fluorescent or antigenic tag, so that the probe's location and quantity in the tissue can be determined, for example using autoradiography, fluorescence microscopy or immunoassay.

*In situ* PCR is the PCR-based amplification of the target nucleic acid sequences followed by *in situ* detection of target and amplicons. Prior to *in situ* PCR, cells or tissue samples generally are fixed and permeabilized to preserve morphology and permit access of the PCR reagents to the intracellular sequences to be amplified; optionally, an intracellular reverse transcription step is introduced to generate cDNA from RNA templates, which enables detection of low copy RNA sequences. PCR amplification of target sequences is next performed; then, intracellular PCR products are visualized by ISH or immunohistochemistry.

Quantitative Nuclease Protection Assay (qNPA)

In particular embodiments of the disclosed methods, nucleic acids are detected in the sample utilizing a quantitative nuclease protection assay and array (such as an array described below). The quantitative nuclease protection assay is described in International Patent Publications WO 99/032663; WO 00/037683; WO 00/037684; WO 00/079008; WO 03/002750; and WO 08/121927; and U.S. Pat. Nos. 6,238,869; 6,458,533; and 7,659,063. See also, Martel et al, *Assay and Drug Development Technologies.* 2002, 1 (1-1):61-71; Martel et al, *Progress in Biomedical Optics and Imaging*, 2002, 3:35-43; Martel et al, *Gene Cloning and Expression Technologies*, Q. Lu and M. Weiner, Eds., Eaton Publishing, Natick (2002); Seligmann, B. *PharmacoGenomics*, 2003, 3:36-43; Martel et al, "Array Formats" in "Microarray Technologies and Applications," U.R. Muller and D. Nicolau, Eds, Springer-Verlag, Heidelberg; Sawada et al, *Toxicology in Vitro*, 20:1506-1513; Bakir *et al., Biorg. & Med. Chem Lett,* 17: 3473-3479; Kris, et al, *Plant Physiol.* 144: 1256-1266; Roberts *et al., Laboratory Investigation*, 87: 979-997; Rimsza *et al., Blood*, 2008 Oct 15, 112 (8): 3425-3433; Pechhold *et al., Nature Biotechnology*, 27, 1038-1042.

Using qNPA methods, a nuclease protection probe (NPP) is allowed to hybridize to the target sequence, which is followed by incubation of the sample with a nuclease that digests single stranded nucleic acid molecules. Thus, if the probe is detected, (*e.g.* it is not digested by the nuclease) then the target of the probe, for example a target nucleic acid shown in Table(s) 3, 4, 11 and/or 13, is present in the sample, and this presence can be detected (*e.g.*, quantified). NPPs can be designed for individual targets and added to an assay as a cocktail for identification on an array; thus, multiple genes targets can be measured within the same assay and/or array.

In some examples, cells in the melanocyte-containing sample are used directly, or are first lysed or permeabilized in an aqueous solution (for example using a lysis buffer). The aqueous solution or lysis buffer may include detergent (such as sodium dodecyl sulfate) and/or one or more

chaotropic agents (such as formamide, guanidinium HCl, guanidinium isothiocyanate, or urea). The solution may also contain a buffer (for example SSC). In some examples, the lysis buffer includes about 15% to 25% formamide (v/v), about 0.01% to 0.1% SDS, and about 0.5-6X SSC. The buffer may optionally include tRNA (for example, about 0.001 to about 2.0 mg/ml) or a ribonuclease. The lysis buffer may also include a pH indicator, such as Phenol Red. In a particular example, the lysis buffer includes 20% formamide, 3X SSC (79.5%), 0.05% DSD, 1 μg/ml tRNA, and 1 mg/ml Phenol Red. Cells are incubated in the aqueous solution for a sufficient period of time (such as about 1 minute to about 60 minutes, for example about 5 minutes to about 20 minutes, or about 10 minutes) and at a sufficient temperature (such as about 22°C to about 115°C, for example, about 37°C to about 105°C, or about 90°C to about 110°C) to lyse or permeabilize the cell. In some examples, lysis is performed at about 95°C, if the nucleic acid to be detected is RNA. In other examples, lysis is performed at about 105°C, if the nucleic acid to be detected is DNA.

In some examples, a nucleic acid protection probe (NPP) (such as those shown in SEQ ID NOS: 1-36 and 123-164) complementary to the target can be added to a sample at a concentration ranging from about 10 pM to about 10 nM (such as about 30 pM to 5 nM, about 100 pM to about 1 nM), in a buffer such as, for example, 6X SSPE-T (0.9 M NaCl, 60 mM $NaH_2PO_4$, 6 mM EDTA, and 0.05% Triton X-100) or lysis buffer (described above). In one example, the probe is added to the sample at a final concentration of about 30 pM. In another example, the probe is added to the sample at a final concentration of about 167 pM. In a further example, the probe is added to the sample at a final concentration of about 1 nM. In such examples, NPPs not digested by a nuclease, such as S1, if the NPP is hybridized to (forms a duplex with) a complementary sequence, such as a target sequence.

One of skill in the art can identify conditions sufficient for an NPP to specifically hybridize to its target present in the test sample. For example, one of skill in the art can determine experimentally the features (such as length, base composition, and degree of complementarity) that will enable a nucleic acid (*e.g.,* NPP) to hybridize to another nucleic acid (*e.g.,* a target nucleic acid in Table(s) 3, 4, 11 and/or 13) under conditions of selected stringency, while minimizing non-specific hybridization to other substances or molecules. Typically, the nucleic acid sequence of an NPP will have sufficient complementarity to the corresponding target sequence to enable it to hybridize under selected stringent hybridization conditions, for example hybridization at about 37°C or higher (such as about 37°C, 42°C, 50°C, 55°C, 60°C, 65°C, 70°C, 75°C, or higher). Among the hybridization reaction parameters which can be varied are salt concentration, buffer, pH, temperature, time of incubation, amount and type of denaturant such as formamide.

The nucleic acids in the sample are denatured (for example at about 95°C to about 105°C for about 5-15 minutes) and hybridized to a NPP for between about 10 minutes and about 24 hours (for example, at least about 1 hour to 20 hours, or about 6 hours to 16 hours) at a temperature ranging from about 4°C to about 70°C (for example, about 37°C to about 65°C, about 45°C to about 60°C, or about 50°C to about 60°C). In some examples, the probes are incubated with the sample at a temperature of at least about 40°C, at least about 45°C, at least about 50°C, at least about 55°C, at least about 60°C, at least about 65°C, or at least about 70°C. In one example, the probes are incubated with the sample at about 60°C. In another example, the NPPs are incubated with the sample at about 50°C. These hybridization temperatures are exemplary, and one of skill in the art can select appropriate hybridization temperature depending on factors such as the length and nucleotide composition of the NPPs.

In some embodiments, the methods do not include nucleic acid purification (for example, nucleic acid purification is not performed prior to contacting the sample with the probes and/or nucleic acid purification is not performed following contacting the sample with the probes). In some examples, no pre-processing of the sample is required except for cell lysis. In some examples, cell lysis and contacting the sample with the NPPs occur sequentially, in some non-limiting examples without any intervening steps. In other examples, cell lysis and contacting the sample with the NPPs occur concurrently.

Following hybridization of the one or more NPPs and nucleic acids in the sample, the sample is subjected to a nuclease protection procedure. NPPs which have hybridized to a full-length nucleic acid are not hydrolyzed by the nuclease and can be subsequently detected.

Treatment with one or more nucleases will destroy nucleic acid molecules other than the probes which have hybridized to nucleic acid molecules present in the sample. For example, if the sample includes a cellular extract or lysate, unwanted nucleic acids, such as genomic DNA, cDNA, tRNA, rRNA and mRNAs other than the gene of interest, can be substantially destroyed in this step. One of skill in the art can select an appropriate nuclease, for example based on whether DNA or RNA is to be detected. Any of a variety of nucleases can be used, including, pancreatic RNAse, mung bean nuclease, S1 nuclease, RNAse A, Ribonuclease T1 , Exonuclease III, Exonuclease VII, RNAse CLB, RNAse PhyM, RNAse U2, or the like, depending on the nature of the hybridized complexes and of the undesirable nucleic acids present in the sample. In a particular example, the nuclease is specific for single-stranded nucleic acids, for example S1 nuclease. An advantage of using a nuclease specific for single-stranded nucleic acids in some method embodiments disclosed here is to remove such single-stranded ("sticky") molecules from subsequent reaction steps where

they may lead to unnecessary background or cross-reactivity. S1 nuclease is commercially available from, for example, Promega, Madison, WI (cat. no. M5761); Life Technologies/Invitrogen, Carlsbad, CA (cat. no. 18001-016); Fermentas, Glen Burnie, MD (cat. no. EN0321), and others. Reaction conditions for these enzymes are well-known in the art and can be optimized empirically.

In some examples, S1 nuclease diluted in an appropriate buffer (such as a buffer including sodium acetate, sodium chloride, zinc sulfate, and detergent, for example, 0.25 M sodium acetate, pH 4.5, 1.4 M NaCl, 0.0225 M $ZnSO_4$, 0.05% KATHON) is added to the hybridized probe mixture and incubated at about 50°C for about 30-120 minutes (for example, about 60-90 minutes) to digest non-hybridized nucleic acid and unbound NPP.

The samples optionally are treated to otherwise remove non-hybridized material and/or to inactivate or remove residual enzymes (*e.g.,* by phenol extraction, precipitation, column filtration, *etc.*). In some examples, the samples are optionally treated to dissociate the target nucleic acid from the probe (*e.g.,* using base hydrolysis and heat). After hybridization, the hybridized target can be degraded, *e.g.*, by nucleases or by chemical treatments, leaving the NPPs in direct proportion to how much NPP had been hybridized to target. Alternatively, the sample can be treated so as to leave the (single strand) hybridized portion of the target, or the duplex formed by the hybridized target and the probe, to be further analyzed.

The presence of the NPPs (or the remaining target or target:NPP complex) is then detected. Any suitable method can be used to detect the probes (or the remaining target or target:NPP complex). In some examples, the NPPs include a detectable label and detecting the presence of the NPP(s) includes detecting the detectable label. In some examples, the NPPs are labeled with the same detectable label. In other examples, the NPPs are labeled with different detectable labels (such as a different label for each target). In other examples, the NPPs are detected indirectly, for example by hybridization with a labeled nucleic acid. In some examples, the NPPs are detected using a microarray, for example, a microarray including detectably labeled nucleic acids (for example labeled with biotin or horseradish peroxidase) that are complementary to the NPPs. In other examples, the NPPs are detected using a microarray including capture probes and programming linkers, wherein a portion of the programming linker is complementary to a portion of the NPPs and subsequently incubating with detection linkers, a portion of which is complementary to a separate portion of the NPPs. The detection linkers can be detectably labeled, or a separate portion of the detection linkers are complementary to additional nucleic acids including a detectable label (such as biotin or horseradish peroxidase). In some examples, the NPPs

are detected on a microarray, for example, as described in International Patent Publications WO 99/032663; WO 00/037683; WO 00/037684; WO 00/079008; WO 03/002750; and WO 08/121927; and U.S. Pat. Nos. 6,238,869; 6,458,533; and 7,659,063.

Briefly, in one non-limiting example, following hybridization and nuclease treatment, the solution is neutralized and transferred onto a programmed ARRAYPLATE (HTG Molecular Diagnostics, Tucson, AZ; each element of the ARRAYPLATE is programmed to capture a specific probe, for example utilizing an anchor attached to the plate and a programming linker associated with the anchor), and the NPPs are captured during an incubation (for example, overnight at about 50°C). The probes can instead be captured on X-MAP beads (Luminex, Austin, TX), an assay referred to as the QBEAD assay, or processed further, including as desired PCR amplification or ligation reactions, and for instance then measured by sequencing). The media is removed and a cocktail of probe-specific detection linkers are added, in the case of the ARRAYPLATE and QBEAD assays, which hybridize to their respective (captured) probes during an incubation (for example, 1 hour at about 50°C). Specific for the ARRAYPLATE and QBEAD assays, the array or beads are washed and then a triple biotin linker (an oligonucleotide that hybridizes to a common sequence on every detection linker, with three biotins incorporated into it) is added and incubated (for example, 1 hour at about 50°C). For the ARRAYPLATE (mRNA assay), HRP-labeled avidin (avidin-HRP) or streptavidin poly-HRP is added and incubated (for example at about 37°C for 1 hour), then washed to remove unbound avidin-HRP or streptavidin poly-HRP. Substrate is added and the plate is imaged to measure the intensity of every element within the plate. In the case of QBEAD Avidin-PE is added, the beads are washed, and then measured by flow cytometry using the Luminex 200, FLEXMAP 3D, or other appropriate instrument. One of skill in the art can design suitable capture probes, programming linkers, detection linkers, and other reagents for use in a quantitative nuclease protection assay based upon the NPPs utilized in the methods disclosed herein.

In some examples, instead of using a detection linker, NPPs are directly biotinylated.

Nucleic Acid Amplification

In some method examples, nucleic acid molecules (such as nucleic acid gene products (*e.g.*, mRNA, miRNA or lncRNA) or nuclease protection probes) are amplified prior to or as a means to their detection. In some examples, nucleic acid expression levels are determined during amplification, for example by using real time RT-PCR.

In one example, a nucleic acid sample can be amplified prior to hybridization, for example hybridization to complementary oligonucleotides present on an array. If a quantitative result is desired, a method is utilized that maintains or controls for the relative frequencies of the amplified nucleic acids. Methods of "quantitative" amplification are well known. For example, quantitative PCR involves simultaneously co-amplifying a known quantity of a control sequence using the same primers. This provides an internal standard that can be used to calibrate the PCR reaction. The array can then include probes specific to the internal standard for quantification of the amplified nucleic acid.

In some examples, the primers used for the amplification are selected so as to amplify a unique segment of the gene product of interest (such as RNA of a gene shown in any of Table(s) 3, 4, 11, and/or 13). In other embodiments, the primers used for the amplification are selected so as to amplify a NPP specific for a gene product of interest (such as RNA of a gene shown in any of Table(s) 3, 4, 11, and/or 13). Primers that can be used to amplify variable gene products (*e.g.*, shown in any of Table(s) 4, 11, and/or 13), as well as normalization gene products (*e.g.*, see Table 3), are commercially available or can be designed and synthesized according to well-known methods.

In one example, RT-PCR can be used to detect RNA (*e.g.*, mRNA, miRNA or lncRNA) levels in melanocyte-containing tissue samples (*e.g.*, skin biopsy). Generally, the first step in gene expression profiling by RT-PCR is the reverse transcription of the RNA template into cDNA, followed by its exponential amplification in a PCR reaction. Two commonly used reverse transcriptases are avian myeloblastosis virus reverse transcriptase (AMV-RT) and Moloney murine leukemia virus reverse transcriptase (MMLV-RT). The reverse transcription step is typically primed using specific primers, random hexamers, or oligo-dT primers, depending on the circumstances and the goal of expression profiling.

Although PCR can use a variety of thermostable DNA-dependent DNA polymerases, it typically employs the Taq DNA polymerase. TaqMan® PCR typically utilizes the 5'-nuclease activity of Taq or Tth polymerase to hydrolyze a hybridization probe bound to its target amplicon, but any enzyme with equivalent 5' nuclease activity can be used. Two oligonucleotide primers are used to generate an amplicon typical of a PCR reaction. A third oligonucleotide, or probe, is designed to detect nucleotide sequence located between the two PCR primers. The probe is non-extendable by Taq DNA polymerase enzyme, and is labeled with a reporter fluorescent dye and a quencher fluorescent dye. Any laser-induced emission from the reporter dye is quenched by the quenching dye when the two dyes are located close together as they are on the probe. During the

amplification reaction, the Taq DNA polymerase enzyme cleaves the probe in a template-dependent manner. The resultant probe fragments dissociate in solution, and signal from the released reporter dye is free from the quenching effect of the second fluorophore. One molecule of reporter dye is liberated for each new molecule synthesized, and detection of the unquenched reporter dye provides the basis for quantitative interpretation of the data.

A variation of RT-PCR is real time quantitative RT-PCR, which measures PCR product accumulation through a dual-labeled fluorogenic probe (*e.g.*, Taqman® probe). Real time PCR is compatible both with quantitative competitive PCR, where internal competitor for each target sequence is used for normalization, and with quantitative comparative PCR using a normalization gene contained within the sample, or a normalization gene for RT-PCR (see Heid *et al.*, *Genome Research* 6:986-994, 1996). Quantitative PCR is also described in U.S. Pat. No. 5,538,848. Related probes and quantitative amplification procedures are described in U.S. Pat. No. 5,716,784 and U.S. Pat. No. 5,723,591. Instruments for carrying out quantitative PCR in microtiter plates are available, e.g., from PE Applied Biosystems (Foster City, CA).

An alternative quantitative nucleic acid amplification procedure is described in U.S. Pat. No. 5,219,727. In this method, the amount of a target sequence (*e.g.*, the expression product of a gene listed in any of Table(s) 4, 11 and/or 13) in a sample is determined by simultaneously amplifying the target sequence and an internal standard nucleic acid segment. The amount of amplified nucleic acid from each segment is determined and compared to a standard curve to determine the amount of the target nucleic acid segment that was present in the sample prior to amplification.

RNA Sequencing

RNA sequencing provides another way to obtain multiplexed and, in some embodiments, high-throughput gene expression information. Numerous specific methods of RNA sequencing are known and/or being developed in the art (for one review, see Chu and Corey, *Nuc. Acid Therapeutics*, 22:271 (2012)). Whole-transcriptome sequencing and targeted RNA sequencing techniques each are available and are useful in the disclosed methods. Representative methods for sequencing-based gene expression analysis include serial analysis of gene expression (SAGE), gene expression analysis by massively parallel signature sequencing (MPSS), whole transcriptome shotgun sequencing (aka, WTSS or RNA-Seq), or nuclease-protection sequencing (aka, qNPS or NPSeq; see PCT Pub. No. WO2012/151111).

*Proteins for Detecting Gene Expression*

In some embodiments of the disclosed methods, determining the level of gene expression in a melanocyte-containing sample (*e.g.*, skin biopsy) includes detecting one or more proteins (for example by determining the relative or actual amounts of such proteins) in the sample. Routine methods of detecting proteins are known in the art, and the disclosure is not limited to particular methods of protein detection.

Protein gene products (*e.g.*, those in any of Table(s) 4 and/or 11) or normalization proteins (*e.g.*, those in Table 3) can be detected and the level of protein expression in the sample can be determined through novel epitopes recognized by protein-specific binding agents (such as antibodies or aptamers) specific for the target protein (such as those in any of Table(s) 3, 4, and/or 11) used in immunoassays, such as ELISA assays, immunoblot assays, flow cytometric assays, immunohistochemical assays, an enzyme immunoassay, radioimmuno assays, Western blot assays, immunofluorescent assays, chemiluminescent assays and other peptide detection strategies (Wong *et al.*, *Cancer Res.*, 46: 6029-6033, 1986; Luwor *et al.*, *Cancer Res.*, 61: 5355-5361, 2001; Mishima *et al.*, *Cancer Res.*, 61: 5349-5354, 2001; Ijaz *et al.*, *J. Med. Virol.*, 63: 210-216, 2001). Generally these methods utilize monoclonal or polyclonal antibodies.

Thus, in some embodiments, the level of target protein expression (such as those in any of Table(s) 3, 4, and/or 11) present in the biological sample and thus the amount of protein expressed is detected using a target protein specific binding agent, such as an antibody of fragment thereof, or an aptamer, which can be detectably labeled. In some embodiments, the specific binding agent is an antibody, such as a polyclonal or monoclonal antibody, that specifically binds to the target protein (such as those in any of Table(s) 3, 4, and/or 11). Thus in certain embodiments, determining the level or amount of protein in a biological sample includes contacting a sample from the subject with a protein specific binding agent (such as an antibody that specifically binds a protein shown in any of Table(s) 3, 4, and/or 11), detecting whether the binding agent is bound by the sample, and thereby measuring the amount of protein present in the sample. In one embodiment, the specific binding agent is a monoclonal or polyclonal antibody that specifically binds to the target protein (such as those in any of Table(s) 3, 4, and/or 11). One skilled in the art will appreciate that there are commercial sources for antibodies to target proteins, such as those in any of Table(s) 3, 4, and/or 11.

The presence of a target protein (such as those in any of Table(s) 3, 4, and/or 11) can be detected with multiple specific binding agents, such as one, two, three, or more specific binding agents. Thus, the methods can utilize more than one antibody. In some embodiments, one of the

antibodies is attached to a solid support, such as a multiwell plate (such as, a microtiter plate), bead, membrane or the like. In practice, microtiter plates may conveniently be utilized as the solid phase. However, antibody reactions also can be conducted in a liquid phase.

In some examples, the method can include contacting the sample with a second antibody that specifically binds to the first antibody that specifically binds to the target protein (such as those in any of Table(s) 3, 4, and/or 11). In some examples, the second antibody is detectably labeled, for example with a fluorophore (such as FITC, PE, a fluorescent protein, and the like), an enzyme (such as HRP), a radiolabel, or a nanoparticle (such as a gold particle or a semiconductor nanocrystal, such as a quantum dot (QDOT®)). In this method, an enzyme which is bound to the antibody will react with an appropriate substrate, such as a chromogenic substrate, in such a manner as to produce a chemical moiety which can be detected, for example, by spectrophotometric, fluorimetric or by visual means. Enzymes which can be used to detectably label the antibody include, but are not limited to, malate dehydrogenase, staphylococcal nuclease, delta-5-steroid isomerase, yeast alcohol dehydrogenase, alpha-glycerophosphate, dehydrogenase, triose phosphate isomerase, horseradish peroxidase, alkaline phosphatase, asparaginase, glucose oxidase, beta-galactosidase, ribonuclease, urease, catalase, glucose-6-phosphate dehydrogenase, glucoamylase and acetylcholinesterase. The detection can be accomplished by colorimetric methods which employ a chromogenic substrate for the enzyme.

Detection can also be accomplished by visual comparison of the extent of enzymatic reaction of a substrate in comparison with similarly prepared standards. It is also possible to label the antibody with a fluorescent compound. Exemplary fluorescent labeling compounds include fluorescein isothiocyanate, rhodamine, phycoerythrin, phycocyanin, allophycocyanin, o-phthaldehyde, Cy3, Cy5, Cy7, tetramethylrhodamine isothiocyanate, phycoerythrin, allophycocyanins, Texas Red and fluorescamine. The antibody can also be detectably labeled using fluorescence emitting metals such as $^{152}$Eu, or others of the lanthanide series. Other metal compounds that can be conjugated to the antibodies include, but are not limited to, ferritin, colloidal gold, such as colloidal superparamagnetic beads. These metals can be attached to the antibody using such metal chelating groups as diethylenetriaminepentacetic acid (DTPA) or ethylenediaminetetraacetic acid (EDTA). The antibody also can be detectably labeled by coupling it to a chemiluminescent compound. Examples of chemiluminescent labeling compounds are luminol, isoluminol, theromatic acridinium ester, imidazole, acridinium salt and oxalate ester. Likewise, a bioluminescent compound can be used to label the antibody. In one example, the antibody is labeled with a bioluminescence compound, such as luciferin, luciferase or aequorin.

Haptens that can be conjugated to the antibodies include, but are not limited to, biotin, digoxigenin, oxazalone, and nitrophenol. Radioactive compounds that can be conjugated or incorporated into the antibodies include but are not limited to technetium 99m ($^{99}$Tc), $^{125}$I and amino acids including any radionucleotides, including but not limited to, $^{14}$C, $^{3}$H and $^{35}$S.

Generally, immunoassays for proteins (such as those in any of Table(s) 3, 4, and/or 11) typically include incubating a biological sample in the presence of antibody, and detecting the bound antibody by any of a number of techniques well known in the art. In one example, the biological sample (such as one containing melanocytes) can be brought in contact with, and immobilized onto, a solid phase support or carrier such as nitrocellulose or a multiwell plate, or other solid support which is capable of immobilizing cells, cell particles or soluble proteins. The support may then be washed with suitable buffers followed by treatment with the antibody that specifically binds to the target protein (such as those in any of Table(s) 3, 4, and/or 11). The solid phase support can then be washed with the buffer a second time to remove unbound antibody. If the antibody is directly labeled, the amount of bound label on solid support can then be detected by conventional means. If the antibody is unlabeled, a labeled second antibody, which detects that antibody that specifically binds to the target protein (such as those in any of Table(s) 3, 4, and/or 11) can be used.

Alternatively, antibodies are immobilized to a solid support, and then contacted with proteins isolated from a biological sample, such as a tissue biopsy from the skin or eye, under conditions that allow the antibody and the protein to bind specifically to one another. The resulting antibody: protein complex can then be detected, for example by adding another antibody specific for the protein (thus forming an antibody:protein:antibody sandwich). If the second antibody added is labeled, the complex can be detected, or alternatively, a labeled secondary antigay can be used that is specific for the second antibody added.

A solid phase support or carrier includes materials capable of binding a sample, antigen or an antibody. Exemplary supports include glass, polystyrene, polypropylene, polyethylene, dextran, nylon, amylases, natural and modified celluloses, polyacrylamides, gabbros and magnetite. The nature of the carrier can be either soluble to some extent or insoluble. The support material may have virtually any possible structural configuration so long as the coupled molecule is capable of binding to its target (such as an antibody or protein). Thus, the support configuration may be spherical, as in a bead, or cylindrical, as in the inside surface of a test tube, or the external surface of a rod. Alternatively, the surface may be flat such as a sheet or test strip.

In one embodiment, an enzyme linked immunosorbent assay (ELISA) is utilized to detect the target protein(s) (*e.g.*, see Voller, "The Enzyme Linked Immunosorbent Assay (ELISA)," *Diagnostic Horizons* 2:1-7, 1978). ELISA can be used to detect the presence of a protein in a sample, for example by use of an antibody that specifically binds to a target protein (such as those in any of Table(s) 3, 4, and/or 11). In some examples, the antibody can be linked to an enzyme, for example directly conjugated or through a secondary antibody, and a substance is added that the enzyme can convert to a detectable signal.

Detection can also be accomplished using any of a variety of other immunoassays; for example, by radioactively labeling the antibodies or antibody fragments. In another example, a sensitive and specific tandem immunoradiometric assay may be used (see Shen and Tai, *J. Biol. Chem.*, 261:25, 11585-11591, 1986). The radioactive isotope can be detected by such means as the use of a gamma counter or a scintillation counter or by autoradiography.

In one example, a spectrometric method is utilized to detect or quantify an expression level of a target protein (such as those in any of Table(s) 3, 4, and/or 11). Exemplary spectrometric methods include mass spectrometry, nuclear magnetic resonance spectrometry, and combinations thereof. In one example, mass spectrometry is used to detect the presence of a target protein (such as those in any of Table(s) 3, 4, and/or 11) in a melanocyte-containing sample, such as a skin biopsy (see for example, Stemmann *et al.*, *Cell* 107(6):715-26, 2001).

A target protein (such as those in any of Table(s) 3, 4, and/or 11) also can be detected by mass spectrometry assays coupled to immunaffinity assays, the use of matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass mapping and liquid chromatography/quadrupole time-of-flight electrospray ionization tandem mass spectrometry (LC/Q-TOF-ESI-MS/MS) sequence tag of proteins separated by two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) (Kiernan *et al.*, *Anal. Biochem.*, 301: 49-56, 2002).

Quantitative mass spectroscopic methods, such as SELDI, can be used to analyze protein expression in a melanocyte-containing sample, such as a skin biopsy. In one example, surface-enhanced laser desorption-ionization time-of-flight (SELDI-TOF) mass spectrometry is used to detect protein expression, for example by using the ProteinChip (Ciphergen Biosystems, Palo Alto, CA). Such methods are well known in the art (*e.g.*, see U.S. Pat. Nos. 5,719,060; 6,897,072; and 6,881,586). Briefly, one version of SELDI uses a chromatographic surface with a chemistry that selectively captures analytes of interest, such as those in any of Table(s) 3, 4, and/or 11.

*Optional Assay Control Measures*

Optionally, assays used to detect gene expression products (*e.g.*, nucleic acids (such as mRNA, miRNA, lncRNA) or protein) will have both positive and negative process control elements used to assess assay performance.

A positive control can be any known element, preferably of a similar nature to the target (*e.g.*, RNA target, then RNA (or cDNA) positive control), that can be included in an assay (or sample) and detected in parallel with the target(s) and that does not interfere (*e.g.*, crossreact) with such target(s) detection. In one example, the positive control is an *in vitro* transcript (IVT) that is run in parallel as a separate sample or is "spiked" into each sample at a known amount. IVT-specific binding agents (*e.g.*, oligonucleotide probes, such as a nuclease protection probe)) and, if applicable, IVT-specific detection agents also are included in each assay to ensure a positive result for such *in vitro* transcript. In another example, an IVT transcript can be designed from non-crossreacting regions of the Methanobacterium sp. AL-21 chromosome (NC_015216).

Negative process control elements can include analyte-specific binding agents (*e.g.*, oligonucleotides or antibodies) designed or selected to detect a gene product that is not expected to be expressed in the applicable test sample. For example, an analyte-specific binding agent that does recognize any gene expression product in the human transcriptome or proteome may be included in a multiplexed assay (such as an oligonucleotide probe or antibody specific for a plant or insect or nematode RNA or protein, respectively, where human gene expression products are the desired targets). This negative control element should not generate signal in the applicable assay. Any above-background signal for such negative process control element is an indicator of assay failure. In one example, the negative control is ANT.

Gene expression can vary across sample types or subjects due to the biology and/or due to variability related to specimen stability, integrity or input level as well as the assay process and system. In order to minimize non-biological related sources of variability (especially in multiplexed assays), gene expression products that do not or are found by bioinformatic methods not to significantly vary (*e.g.*, "housekeepers" or normalizers) among samples of interest are measured in particular embodiments. In some such embodiments, expression levels for candidate normalization gene products will demonstrate adequate (*e.g.*, above-background) and/or non-saturated intensity values. Further discussion of normalizer gene expression products is found elsewhere in this disclosure.

In some situations, anomalous signals may result from unexpected process-related issues that are not otherwise controlled, *e.g.*, by analysis of normalizers; thus, in some embodiments, it is

useful to include a sample-independent process control element(s) to indicate a successful or failed assay on any specimen, irrespective of the specimen stability, integrity, or input level. Method embodiments in which nucleic acid gene expression products are detected may include a known concentration of a RNA sample (*e.g.*, *in vitro* transcript RNA or IVT) in every assay. Such a control element (*e.g.*, IVT) will be measured in each assay and act as an assay process quality control.

The MAQC (Microarray Array Quality Control) project proposed that a "Universal Human Reference RNA" could be a useful external-control standard for microarray gene expression assays. Accordingly, some disclosed method embodiments involving RNA gene expression products may, but need not, include a parallel-processed sample containing Universal Human Reference RNA. If such universal RNA sample includes all or some of the RNAs targeted for detection by the applicable assay, a positive signal can be expected for such included RNAs, which may serve as an (or another) assay process quality control.

***Gene Expression Data***

It is well accepted that gene expression data "contain the keys to address fundamental problems relating to the prevention and cure of diseases, biological evolution mechanisms and drug discovery" (Lu and Han, *Information Systems*, 28:243-268 (2003)). In some examples, distilling the information from such data is as simple as making a qualitative determination from the presence, absence or qualitative amount (*e.g.*, high, medium, low) of one or more gene products detected. In other examples, raw gene expression data may be pre-processed (*e.g.*, background subtracted, log transformed, and/or corrected), normalized, and/or applied in classification algorithms. These aspects are described in more detail below.

*Data Pre-processing*

Background Subtraction

In some method embodiments, raw gene expression data is background subtracted. This correction may be used, for example, where data has been collected using multiplexed methods, such as microarrays. One aim of such transformation is to correct for local effects, *e.g.*, where one portion of a microarray surface may look "brighter" than another portion of the surface without any biological reason. Methods of background subtraction are well known in the art and include, *e.g.*, (i) local background subtraction (*e.g.*, consider all pixels that are outside the spot mask but within the bounding box centered at the spot center), (ii) morphological opening background estimation (relies on non-linear morphological filters, such as opening, erosion, dilation and rank filters (see, Soille, Morphological Image Analysis: Principles and Applications, Berlin: Springer-Verlag

(1999), to create a background image for subtraction from the original image), (iii) constant background (subtracts a constant background for all spots), Normexp background correction (a convolution of normal and exponential, distributions is fitted to the foreground intensities, using the background intensities as a covariate, and the expected signal given the observed foreground becomes the corrected intensity).

Data Transformation

Many biological variables (*e.g.*, gene expression data) do not meet the assumptions of parametric statistical tests, *e.g.*, such variables are not normally distributed, the variances are not homogeneous, or both (Durbin *et al.*, *Bioinformatics*, 18:S105 (2002)). In some cases, transforming the data will make it fit the statistical assumptions better. In some method embodiments, useful data transformation can include (i) log transformation, which consists of taking the log of each observation, *e.g.*, base-10 logs, base-2 logs, base-e logs (also known as natural logs); the log selection makes no difference because such logs differ by a constant factor; or variance-stabilizing transformation, *e.g.*, as described by Durbin (supra). In specific examples, raw expression values for each biomarker detected in such method (*e.g.*, at least two Table 4, 11 and/or 13 biomarkers and/or at least one normalization biomarker) are log (*e.g.*, log 2 or log 10) transformed. In other embodiments, the normalizing step can include dividing each of the at least two Table 4, 11 and/or 13 biomarkers log (*e.g.*, log 2 or log 10) transformed raw expression values by the log (*e.g.*, log 2 or log 10) transformed raw expression value(s) of the at least one normalization biomarker.

Data Filters

Gene expression data may be filtered in some method embodiments to remove data that may be considered unreliable. It is understood that there are many methods known in the art for assessing the reliability of gene expression data and the following non-limiting examples are merely representative.

Gene expression data may be excluded from a disclosed method, in some cases, if it is not expressed or is expressed at an undetectable level (not above background). Oppositely, gene expression data may be excluded from analysis, in some cases, if the expression of a negative control (*e.g.*, ANT) gene is greater than an standard cut off (*e.g.*, more than 100, 200, 250, or 300 relative light units, or more than 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, or 10% above background).

For embodiments involving probe-sets or genes, there are a number of specific data filters that may be useful, including:

(i) Data arising from unreliable probe sets may be selected for exclusion from analysis by ranking probe-set reliability against a series of reference datasets. For example, RefSeq and Ensembl (EMBL) are considered very high quality reference datasets. Data from probe sets matching RefSeq or Ensembl sequences may in some cases be specifically included in microarray analysis experiments due to their expected high reliability. Similarly data from probe-sets matching less reliable reference datasets may be excluded from further analysis, or considered on a case by case basis for inclusion; or

(ii) Probe-sets that exhibit no, or low variance may be excluded from further analysis. Low-variance probe-sets may be excluded from the analysis via a Chi-Square test. A probe-set is considered to be low-variance if its transformed variance is to the left of the 99 percent confidence interval of the Chi-Squared distribution with (N-1) degrees of freedom; or

(iii) Probe-sets for a given gene or transcript cluster may be excluded from further analysis if they contain less than a minimum number of probes, *e.g.*, following other data pre-processing steps. For example in some embodiments, probe-sets for a given gene or transcript cluster may be excluded from further analysis if they contain less than 1, 2, 3, 4, or 5 probes.

Optionally, a statistical outlier program can be used that determines whether one of several replicates is statistically an outlier compared to the others, such as judged by being "x" standard deviations (SD) (*e.g.* at least 2-SD or at least 3-SD) away from the average, or CV% of replicates greater than a specified amount (*e.g.*, at least 8% in log-transformed space). In an array-based assay, an outlier could result from there being a problem with one of the array spots, or due to an imaging artifact. Outlier removal is typically performed on a gene-by-gene basis, and if most of the genes in one replicate are outliers, one can apply a pre-established rule that eliminates the entire replicate. For instance, a pipetting error resulting in the improper addition of a critical reagent could cause the entire replicate to be an outlier.

In some examples where gene expression is measured in sample replicates (*e.g.*, triplicates), reproducibility can be measured by pairwise correlation and by pairwise sample linear regression, and a correlation r >= 0.95 used as acceptance of replicate (*e.g.*, triplicate) reproducibility. In more specific examples, replicates with pairwise correlation r => 0.90 can be further reviewed by a simple regression model; in which case, if the intercept of the linear regression is statistically signicantly different from zero, the replicate removed from further consideration. Any sample with more than 25% (*e.g.*, 1 out of 4) or more, 33% (*e.g.*, 1 out of 3) or more, 50% (*e.g.*, 2 out of 4) or

more, or 67% (*e.g.*, 2 out of 3) or more failed replicates may be considered a "failed sample" and removed from further analysis.

*Normalization*

The objective of normalization is to remove variability due to experimental error (for example due to be due to pipetting, plate position, image artifacts, different amounts of total RNA, *etc.*) so that variation due to biological effects can be observed and quantified. This process helps ensure the differences observed between different sample types is due truly to difference in sample biology and not due to some technical artifact. There are several points during experimentation at which errors can be introduced and which can be eliminated by normalization. Methods for normalization of gene expression data are well established in the art (*e.g.*, *Methods in Microarray Normalization*, ed. by Phillip Stafford, Baton Rouge, FL: CRC Press an imprint of Taylor & Francis Group, 2008).

Normalization typically involves comparing an experimental value, such as the expression value of one or more Table 4, 11 and/or 13 biomarkers, to one or more normalizing value(s) or factor(s) (*e.g.*, by dividing (or subtracting, typically, after log transformation). A normalizing value can be the raw (or log transformed) expression value of a single normalizer biomarker or can be calculated, *e.g.*, from the expression values of a plurality of normalizers or using methods and calculations known in the art. In some examples, normalizing uses a mean value of the expression of a plurality of normalization biomarkers to generate normalized expression values for each Table 4, 11 and/or 13 biomarker tested. In some examples, normalizing uses raw expression values for each of the Table 4, 11 and/or 13 biomarkers tested, and raw expression values for at least one normalization marker in Table 3, to generate normalized expression values for each Table 4, 11 and/or 13 biomarker tested.

In some embodiments, the expression of one or more "normalization biomarkers" can be determined or measured, such as one or more those in Table 3. For example, expression of 1, 2, 3, 4, 5, 6, 7, 8 or all of BMP-1, MFI2, NCOR2, RAP2b, RPS6KB2, SDHA, RPL19, RPLP0, and ALDOA can be detected in the test sample.

Alternatively, one or more normalization biomarkers useful in a disclosed method can be identified using the methods provided herein. For example, a normalization biomarker is any constitutively expressed gene (or protein) against whose expression another expressed gene (or protein) can be compared (*e.g.*, by dividing (or subtracting, typically, after log transformation) the expression of one by the other). In other examples, a normalization biomarker can be any gene expression product (*e.g.*, mRNA, miRNA, or protein) the expression of which does not

significantly differ across a representative plurality of samples, such as nevi and melanoma samples. Accordingly, in some methods, a normalization biomarker can be any gene expression product not listed in Table(s) 4, 11, and/or 13, the expression of which does not significantly differ between melanocyte-containing samples (*e.g.*, a representative population of nevi and melanoma samples). In other examples, the at least one normalization biomarker(s) can include a plurality of normalization biomarkers, none of whose expression is statistically significant difference between nevi and primary melanoma samples.

Another way to identify normalization biomarkers useful in disclosed methods is to determine if, when comparing raw data, the expression of putative normalizers track with one another (*i.e.*, if one normalization biomarker goes up, the other normalization biomarkers should as well). Useful normalizers will track one another across multiple samples of interest. The ratio between putative normalization biomarkers also can be determined and normalizers identified if the ratio between them remains constant across a plurality of samples of interest (*e.g.*, melanocyte-containing samples).

Having identified normalization biomarkers, *e.g.*, as described in this disclosure, some method embodiments include normalizing raw (or log transformed) expression values for each of the at least two biomarkers in Table(s) 4, 11 and/or 13 to raw (or log transformed expression values for at least one normalization biomarker(s).

Alternatively, a normalization value can be determined and such value used to normalize the experimental values (*e.g.*, the gene expression values of at least two different biomarkers from Table(s) 4, 11 and/or 13). For example, a population CT (*e.g.*, mean (such as, arithmetic or geometric mean), median, mode, or average) of a plurality of biomarkers whose range and distribution of expression values is representative of the range and distribution of expression of the gene population in the transcriptome of the sample(s) of interest (*e.g.*, melanocyte-containing samples, such as nevi and/or melanoma samples) may serve as a normalization value in some disclosed methods. In other examples, the expression values of outliers (*e.g.*, +/- one or two standard deviations from the population CT) in the plurality of biomarkers are removed from the original calculation of biomarker plurality's population CT and an outlier-free population CT is determined for the plurality of biomarkers and serves as the normalization value for experimental variables (*e.g.*, gene expression values for at least two genes in Table(s) 4, 11, and/or 13).

In other specific examples, the robust multi-array average (RMA) method may be used to normalize the raw data. The RMA method begins by computing background-corrected intensities for each matched cell on a number of microarrays. The background corrected values are restricted

to positive values as described by Irizarry *et al.* (*Biostatistics*, 4:249 (2003)). After background correction, the base-2 logarithm of each background-corrected matched-cell intensity is then obtained. The background-corrected, log-transformed, matched intensity on each microarray is then normalized using the quantile normalization method in which for each input array and each probe expression value, the array percentile probe value is replaced with the average of all array percentile points, this method is more completely described by Bolstad *et al.* (*Bioinformatics*, 19(2):185 (2003)). Following quantile normalization, the normalized data may then be fit to a linear model to obtain an expression measure for each probe on each microarray.

In some examples, a first normalization can be across the replicates within a treatment or within technical replicates. This is a normalization to all the tested biomarkers (such as two or more of those in Table(s) 4, 11 and/or 13) weighted to a constant level of the total signal for that set of replicates. In this step, the total signal intensity for each assay (such as a well or bead or lane) in a set of replicates is adjusted so that all are equal. The average total signal is calculated for all the replicates, and then a normalization factor is calculated for each sample which adjusts the total signal form that replicate to the total average signal for all replicates. This normalization factor is use then to normalize the signal for each biomarker in that replicate.

*Feature Selection (FS)*

Classification algorithms typically perform suboptimally with thousands of features (genes/proteins). Thus, feature selection methods are used to identify features that are most predictive of a phenotype. The selected genes/proteins are presented to a classifier or a prediction model. The following benefits result from reducing the dimensionality of the feature space: (i) improve classification accuracy, (ii) provide a better understanding of the underlying concepts that generated the data, and (iii) overcome the risk of data overfitting, which arises when the number of features is large and the number of training patterns is comparatively small. Feature selection was used to determine the disclosed gene sets; therefore the corresponding classifiers have the foregoing advantages built in.

Feature selection techniques including filter techniques (which assess the relevance of features by looking at the intrinsic properties of the data), wrapper methods (which embed the model hypothesis within a feature subset search), and embedded techniques (in which the search for an optimal set of features is built into a classifier algorithm). Filter FS techniques useful in disclosed methods include: (i) parametric methods such as the use of two sample t-tests or moderated t-tests (*e.g.*, LIMMA), ANOVA analyses, Bayesian frameworks, and Gamma distribution models, (ii) model free methods such as the use of Wilcoxon rank sum tests, between-

within class sum of squares tests, rank products methods, random permutation methods, or total number of misclassifications (TNoM) which involves setting a threshold point for fold-change differences in expression between two datasets and then detecting the threshold point in each gene that minimizes the number of missclassifications, and (iii) multivariate methods such as bivariate methods, correlation based feature selection methods (CFS), minimum redundancy maximum relavance methods (MRMR), Markov blanket filter methods, tree-based methods, and uncorrelated shrunken centroid methods. Wrapper methods useful in disclosed methods include sequential search methods, genetic algorithms, and estimation of distribution algorithms. Embedded methods useful in the methods of the present disclosure include random forest (RF) algorithms, weight vector of support vector machine algorithms, and weights of logistic regression algorithms. Saeys *et al.* describe the relative merits of the filter techniques provided above for feature selection in gene expression analysis. In some embodiments, feature selection is provided by use of the LIMMA software package (Smyth, LIMMA: Linear Models for Microarray Data, In: Bioinformatics and Computational Biology Solutions, ed. by Gentleman et al., New York:Springer, pages 397-420 (2005)).

*Classifier Algorithms*

In some methods, gene expression information (*e.g.*, for the biomarkers described in Table(s) 3, 4, 11 and/or 13) is applied to an algorithm in order to classify the expression profile (*e.g.*, whether a melanocyte-containing sample (such as a skin biopsy) is a benign nevus or a primary melanoma or neither (such as, indeterminant)). The methods disclosed herein can include gene expression-based classifiers for characterizing melanocyte-containing samples as nevi or melanoma. Specific classifier embodiments are described and, based on the provided gene sets and classification methods, others now are enabled.

A classifier is a predictive model (*e.g.*, algorithm or set of rules) that can be used to classify test samples (*e.g.*, melanocyte-containing samples) into classes (or groups) (*e.g.*, nevus or melanoma) based on the expression of genes in such samples (such as the genes in Table(s) 4, 11 and/or 13). Unlike cluster analysis for which the number of clusters is unknown in advance, a classifier is trained on one or more sets of samples for which the desired class value(s) (*e.g.*, nevus or melanoma) is (are) known. Once trained, the classifier is used to assign class value(s) to future observations.

Illustrative algorithms useful in disclosed methods include, but are not limited to, methods that reduce the number of variables such as principal component analysis algorithms, partial least squares methods, and independent component analysis algorithms. Illustrative algorithms further

include, but are not limited to, methods that handle large numbers of variables directly such as statistical methods and methods based on machine learning techniques. Statistical methods include penalized logistic regression, prediction analysis of microarrays (PAM), methods based on shrunken centroids, support vector machine analysis, and regularized linear discriminant analysis. Machine learning techniques include bagging procedures, boosting procedures, random forest algorithms, and combinations thereof. Boulesteix *et al.* (*Cancer Inform.*, 6:77 (2008)) provide an overview of the classification techniques provided above for the analysis of multiplexed gene expression data.

Machine learning is where a computer uses adaptive technology to recognize patterns and anticipate actions; thereby sorting through vast amounts of data and analyzing and identifying patterns. Machine learning algorithms (*e.g.*, Logistic Regression (LR), Random Forest (RF), Support Vector Machine (SVM), K-nearest neighbor (KNN)) can be useful for developing software in applications too complex for people to manually design the algorithm.

In some embodiments, test samples are classified using a trained algorithm. Trained algorithms of the present disclosure include algorithms that have been developed using a reference set of known nevi and melanoma samples. Algorithms suitable for categorization of samples include, but are not limited to, k-nearest neighbor algorithms, concept vector algorithms, naive bayesian algorithms, neural network algorithms, hidden markov model algorithms, genetic algorithms, and mutual information feature selection algorithms or any combination thereof. In some cases, trained algorithms of the present disclosure may incorporate data other than gene expression data such as but not limited to scoring or diagnosis by cytologists or pathologists of the present disclosure, information provided by a disclosed pre-classifier algorithm or gene set, or information about the medical history of a subject from whom a tested sample is taken.

In some specific embodiments, a support vector machine (SVM) algorithm, a random forest algorithm, or a combination thereof provides classification of samples (*e.g.*, melanocyte-containing samples) into nevus or melanoma (*e.g.*, primary melanoma) and, optionally, indeterminant classes. In some embodiments, identified markers that distinguish samples (*e.g.*, nevi vs. melanoma) are selected based on statistical significance. In some cases, the statistical significance selection is performed after applying a Benjamini Hochberg correction for false discovery rate (FDR) (see, *J. Royal Statistical Society*, Series B (Methodological) 57:289 (1995)).

In some cases, a disclosed classifier algorithm may be supplemented with a meta-analysis approach such as that described by Fishel *et al.* (*Bioinformatics*, 23:1599 (2007)). In some cases, the classifier algorithm may be supplemented with a meta-analysis approach such as a repeatability

analysis. In some cases, the repeatability analysis selects markers that appear in at least one predictive expression product marker set.

*Exemplary Decision Tree Models*

A decision tree algorithm is a flow-chart-like tree structure where each internal node denotes a test on an attribute, and a branch represents an outcome of the test. Leaf nodes represent class labels or class distribution. To generate a decision tree, all the training examples are used at the root, the logical test at the root of the tree is applied and training data then is partitioned into sub-groups based on the values of the logical test. This process is recursively applied (*i.e.*, select attribute and split) and terminated when all the data elements in one branch are of the same class. To classify an unknown sample, its attribute values are tested against the decision tree.

As one example of machine learning, Random Forests are ensemble learning methods for classification (and regression) that operate by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes output by individual trees. In one particular Random Forest algorithm (Breiman, *Machine Learning*, 45:5-32 (2001)), each tree is constructed as follows:

1. Let the number of training cases be "N," and the number of variables in the classifier be "M."
2. "m" is the number of input variables to be used to determine the decision at a node of the tree; m should be less than M.
3. Choose a training set for this tree by choosing n times with replacement from all N available training cases (*i.e.*, take a bootstrap sample). Use the rest of the cases to estimate the error of the tree, by predicting their classes.
4. For each node of the tree, randomly choose m variables on which to base the decision at that node. Calculate the best split based on these m variables in the training set.
5. Each tree is fully grown and not pruned (as may be done in constructing a normal tree classifier).

For prediction, a new sample is pushed down the tree. It is assigned the label of the training sample in the terminal node it ends up in. This procedure is iterated over all trees in the ensemble, and the mode vote of all trees is reported as the random forest class prediction.

*Exemplary Logistic Regression Models*

One representative method for developing statistical predictive models using the genes in Table(s) 4, 11 and/or 13 is logistic regression with a binary distribution and a logit link function. Estimation for such models can be performed using Fischer Scoring. However, models estimated

with exact logistic regression, Empirical Sandwich Estimators or other bias corrected, variance stabilized or otherwise corrective estimation techniques will also, under many circumstances, provide similar models which while yielding slightly different parameter estimates will yield qualitatively consistent patterns of results. Similarly, other link functions, including but not limited to a cumulative logit, complementary log-log, probit or cumulative probit may be expected to yield predictive models that give the same qualitative pattern of results.

One representative form of a predictive model (algorithm) is:

$$\text{Logit}(Yi) = \beta 0 + \beta 1X1 + \beta 2X2 + \beta 3X3...\beta nXn$$

where βo is an intercept term, βn is a coefficient estimate and Xn is the log expression value for a given gene (*e.g.*, any log, such as log base 2 or log base 10). Typically, the value for all β will be greater than -1,000 and less than 1,000. Often, the β0 intercept term will be greater than -200 and less than 200 with cases in which it is greater than -100 and less than 100. The additional βn, where n>0, can be greater than -100 and less than 100.

In particular method embodiments, the Logit(Yi) output is referred to as a consolidated expression value (CEV) for the at least two Table(s) 4, 11 and/or 13 biomarkers. The CEV is determined by (a) weighting the expression level of the at least two Table(s) 4, 11 and/or 13 biomarkers with a constant predetermined for each of the at least two Table(s) 4, 11 and/or 13 biomarkers, and (b) combining the weighted expression levels of the at least two Table(s) 4, 11 and/or 13 biomarkers to produce the CEV. Such a method can also include comparing the CEV to a reference value that distinguishes known melanoma (*e.g.*, primary melanoma) samples from known benign nevus samples. In one example, the method further includes characterizing the sample as malignant (*e.g.*, primary melanoma) if the CEV falls on the same side of the reference value as do the known melanoma samples. In another example, the method further includes characterizing the sample as benign (*e.g.*, nevus) if the CEV falls on the same side of the reference value as do known benign nevi samples.

Performance of any predictive model contemplated herein may be validated with a number of tests known in the art, including, but not limited to, Wald Chi-Square test (overall model fit), and Hosmer and Lemeshow lack fit test (no statistically detectable lack of fit for the model). Predictors for each gene in the model should be stastically significant (*e.g.*, $p<0.05$).

A number of cross validation methods are available to ensure reproducibility of the results. An exemplary method is a one-step maximum likelihood estimate approximation implemented as part of the SAS Proc Logistic classification table procedure. In some examples, ten (10)-fold cross validation and 66-33% split validation in the open source package Weka can be used for

confirmation of results. In other examples, n-fold, including leave-one-out (LOO), cross validation and split sample training/testing provides useful confirmation of results.

In some method embodiments, algorithms (aka, fitted model) provide a predicted event probability, which, for example, is the probability of a melanocyte-containing sample (*e.g.*, skin biopsy) sample being a melanoma (*e.g.*, primary melanoma), being malignant, being a nevus, or being benign. In some instances, a SAS computation method known to those of ordinary skill in the art can be used to compute a reduced-bias estimate of the predicted probability (see, support.sas.com/documentation/cdl/en/statug/63347/HTML/default/viewer.htm#statug_logistic_sect044.htm (as of March 15, 2013)). In other examples, a series of threshold values, z, where z is between 0 and 1 are set, as typically determined by the ordinarily skilled artisan based on the desired clinical utility of a model or application requirement. If the predicted probability calculated for a particular sample exceeds or equals the pre-set threshold value, z, the sample is assigned to the nevus group; otherwise, it was assigned to the melanoma group or *vice versa*. In other examples, two threshold values can be set where sample values falling between the two thresholds are assigned an "indeterminant" or "not otherwise assigned" or the like label.

Based on the algorithm output, a determination is made as to whether a tested sample (*e.g.*, a skin sample) is malignant or benign, for example, by comparing the output to a reference standard (*e.g.*, a cutoff determined from known malignant and benign melanocyte-containing samples). In some examples, the steps of calculating the output from the algorithm and/or determining from the algorithm output that the sample is or is not malignant by comparing the output to a reference standard, are performed by a suitably programmed computer. In some examples, the method can also include providing to a user a report comprising the algorithm output or the determination that the sample is or is not malignant or is "consistent with melanoma," or "consistent with nevus" or "indeterminant" or the like. In some examples the report includes a CEV for the at least two biomarkers from Table(s) 4, 11 and/or 13 analyzed.

The resulting output value is compared to a cut-off value. The cut-off value can be determined by a machine learning or logistic regression analysis of normalized expression values for the at least two biomarkers from Table(s) 4, 11 and/or 13 in a plurality of melanocyte-containing samples known in advance to be benign or malignant. Cut-off values may be determined by individual users on a case-by-case basis, for example, by selecting particular sensitivity and specificity values and/or AUC value for the nevi-melanoma classifer being used. Other methods for determine cut-off values are provided in WO 02/103320 and U.S. Patent Nos.

7,171,311; 7,514,209; 7,863,001; and 8,019,552.

In some examples, a tested sample (*e.g.*, a skin biopsy) is characterized as benign if the algorithm output value is on the same side of the cut-off value as the plurality of known benign samples, or characterized as malignant if the output value is on the same side of the cut-off value as the plurality of known malignant samples. In one example, the sample is characterized as benign if the output value is below the cut-off value or as malignant if the output value is above the cut-off value. In another example, the sample is characterized as benign if the output value is above the cut-off value or as malignant if the output value is below the cut-off value.

*Molecular Profiling and Classifier Outputs*

There typically are four possible outcomes when classifying a biological sample, such as a melanocyte-containing sample, with a disclosed method that includes a binary classifier. If the outcome from a prediction is p and the actual value is also p, then it is called a true positive (TP); however if the actual value is n then it is said to be a false positive (FP). Conversely, a true negative has occurred when both the prediction outcome and the actual value are n, and false negative is when the prediction outcome is n while the actual value is p. Consider an embodiment that seeks to determine whether a sample is a melanoma (*e.g.*, a primary melanoma). A false positive in this case occurs when a sample tests positive, but is not actually a melanoma (*e.g.*, a primary melanoma). A false negative, on the other hand, occurs when the sample tests negative (*i.e.*, not melanoma), when it actually is a melanoma (*e.g.*, a primary melanoma). In some embodiments, ROC curve assuming real-world prevalence of subtypes can be generated by re-sampling errors achieved on available samples in relevant proportions.

The positive predictive value (PPV), or precision rate, or post-test probability of melanoma (*e.g.*, a primary melanoma), is the proportion of samples with positive test results that correctly are melanoma (*e.g.*, a primary melanoma). PPV reflects the probability that a positive test reflects the underlying hypothesis being tested (*e.g.*, a sample is a melanoma (such as, a primary melanoma)). In one example:

False positive rate ($\alpha$) = FP/(FP+TN)-specificity

False negative rate ($\beta$) = FN/(TP+FN)-sensitivity

Power = sensitivity = 1-$\beta$

Likelihood-ratio positive = sensitivity/(1-specificity)

Likelihood-ratio negative = (1-sensitivity)/specificity

where TN is true negative, FN is false negative and TP and FP are as defined above.

Negative predictive value (NPV) is the proportion of subjects or samples with a negative test result (*e.g.*, nevus or indeterminant) who are correctly diagnosed or subtyped. A high NPV for a given test means that when the test yields a negative result, it is most likely correct in its assessment.

In some embodiments, the results of the gene expression analysis of the disclosed methods provide a statistical confidence level that a given diagnosis (*e.g.*, nevus or melanoma or indeterminant) is correct. In some embodiments, such statistical confidence level is above 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 99.5%.

In one aspect of the present disclosure, samples that have been processed by another method (*e.g.*, histopathology and/or immunocytochemistry) and diagnosed are, then, subjected to disclosed molecular profiling as a second diagnostic screen. This second diagnostic screen enables, at least: 1) a significant reduction of false positives and false negatives, 2) a determination of the underlying genetic, metabolic, or signaling pathways responsible for the resulting pathology, 3) the ability to assign a statistical probability to the accuracy of the diagnosis, 4) the ability to resolve ambiguous results, and 5) the ability to properly characterize a previously ambiguous sample.

In some embodiments, the biological sample is classified as nevus or melanoma (*e.g.*, primary melanoma) with an accuracy of greater than 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.5%. The term accuracy as used in the foregoing sentence includes specificity, sensitivity, positive predictive value, negative predictive value, and/or false discovery rate.

In other cases, receiver operator characteristic (ROC) analysis may be used to determine the optimal assay parameters to achieve a specific level of accuracy, specificity, positive predictive value, negative predictive value, and/or false discovery rate. A ROC curve is a graphical plot that illustrates the performance of a binary classifier system as its discrimination threshold is varied. It is created by plotting the fraction of true positives out of the positives (TPR = true positive rate) vs. the fraction of false positives out of the negatives (FPR = false positive rate) at various threshold settings.

***Method Implementation***

The methods, such as those involving classifiers, described herein can be implemented in numerous ways. Several representative non-limiting embodiments are described below.

In some method embodiments, gene expression data is input (*e.g.*, manually or automatically) into a computer or other device, machine or apparatus for application of the various algorithms described herein, which is particularly advantageous where a large number of gene

expression data points are collected and processed. Other embodiments involve use of a communications infrastructure, for example the internet. Various forms of hardware, software, firmware, processors, or a combination thereof are useful to implement specific classifier and method embodiments. Software can be implemented as an application program tangibly embodied on a program storage device, or different portions of the software implemented in the user's computing environment (*e.g.*, as an applet) and on the reviewer's computing environment, where the reviewer may be located at a remote site associated (*e.g.,* at a service provider's facility).

For example, during or after data input by the user, portions of the data processing can be performed in the user-side computing environment. For example, the user-side computing environment can be programmed to provide for defined test codes to denote a likelihood "score," where the score is transmitted as processed or partially processed responses to the reviewer's computing environment in the form of test code for subsequent execution of one or more algorithms to provide a results and/or generate a report in the reviewer's computing environment. The score can be a numerical score (representative of a numerical value) or a non-numerical score representative of a numerical value or range of numerical values (*e.g.*, "A" representative of a 90-95% likelihood of an outcome).

The application program for executing the algorithms described herein may be uploaded to, and executed by, a machine comprising any suitable architecture. In general, the machine involves a computer platform having hardware such as one or more central processing units (CPU), a random access memory (RAM), and input/output (I/O) interface(s). The computer platform also includes an operating system and microinstruction code. The various processes and functions described herein may either be part of the microinstruction code or part of the application program (or a combination thereof) which is executed via the operating system. In addition, various other peripheral devices may be connected to the computer platform such as an additional data storage device and a printing device.

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

Part or all of the input and output data can also be sent electronically; certain output data (*e.g.*, reports) can be sent electronically or telephonically (*e.g.*, by facsimile, using devices such as fax back). Exemplary output receiving devices can include a display element, a printer, a facsimile

device and the like. Electronic forms of transmission and/or display can include email, interactive television, and the like. In one embodiment, all or a portion of the input data and/or all or a portion of the output data (*e.g.*, usually at least the final report) are maintained on a web server for access, preferably confidential access, with typical browsers. The data may be accessed or sent to health professionals as desired. The input and output data, including all or a portion of the final report, can be used to populate a patient's medical record which may exist in a confidential database at the healthcare facility. In some examples, the method includes generating a report. In some examples the report includes an icon indicating the classification of a sample, such as a "+" or "M" for melanoma or a "-" or "N" for nevi.

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

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

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

The input client components can be complete, stand-alone personal computers offering a full range of power and features to run applications. The client component usually operates under

any desired operating system and includes a communication element (*e.g.*, a modem or other hardware for connecting to a network), one or more input devices (*e.g.*, a keyboard, mouse, keypad, or other device used to transfer information or commands), a storage element (*e.g.*, a hard drive or other computer-readable, computer-writable storage medium), and a display element (*e.g.*, a monitor, television, LCD, LED, or other display device that conveys information to the user). The user enters input commands into the computer processor through an input device. Generally, the user interface is a graphical user interface (GUI) written for web browser applications.

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

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

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

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

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

*Computer-Readable Storage Media*

The present disclosure also contemplates a computer-readable storage medium (*e.g.* CD-ROM, memory key, flash memory card, diskette, *etc.*) having stored thereon a program which, when executed in a computing environment, provides for implementation of algorithms to carry out

all or a portion of the results of a response likelihood assessment as described herein. Where the computer-readable medium contains a complete program for carrying out the methods described herein, the program includes program instructions for collecting, analyzing and generating output, and generally includes computer readable code devices for interacting with a user as described herein, processing that data in conjunction with analytical information, and generating unique printed or electronic media for that user.

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

*Output*

In some embodiments, once a score for a particular sample (patient) is determined, an indication of that score can be displayed and/or conveyed to a clinician or other caregiver. For example, the results of the test are provided to a user (such as a clinician or other health care worker, laboratory personnel, or patient) in a perceivable output that provides information about the results of the test. In some examples, the output is a paper output (for example, a written or printed output), a display on a screen, a graphical output (for example, a graph, chart, or other diagram), or an audible output. Thus, the output can include a report that is generated.

For example, the output can be textual (optionally, with a corresponding) score. For example, textual outputs may be "consistent with nevus" or the like, or "consistent with melanoma" or the like (such as, "consistent with primary melanoma"), or "indeterminant" (*e.g.*, not consistent with either nevus or melanoma) or the like. Such textual output can be used, for example, to provide a diagnosis of benign sample (*e.g.*, nevus) or malignant sample (*e.g.*, primary melanoma),

or can simply be used to assist a clinician in distinguishing a nevus from a melanoma (*e.g.*, a primary melanoma).

In other examples, the output is a numerical value (*e.g.*, quantitative output), such as an amount of gene or protein expression (such as those in any of Table(s) 3, 4, 11 and/or 13) in the sample or a relative amount of gene or protein expression (such as those in any of 4, 11 and/or 13) in the sample as compared to a control. In additional examples, the output is a graphical representation, for example, a graph that indicates the value (such as amount or relative amount) of gene or protein expression (such as those in any of Table(s) 3, 4, 11 and/or 13) in the sample from the subject on a standard curve. In a particular example, the output (such as a graphical output) shows or provides a cut-off value or level that characterizes the sample tested as nevus or melanoma (*e.g.*, primary melanoma). In other examples, the output is an icon, such as a "N" or "-" if the sample is classified as a nevus, "M" or "+" if the sample is classified as a melanoma", or "I" or "?" if the sample is classified as a indeterminant (*e.g.*, not consistent with either nevus or melanoma). In some examples, the output is communicated to the user, for example by providing an output via physical, audible, or electronic means (for example by mail, telephone, facsimile transmission, email, or communication to an electronic medical record).

In additional examples, the output can provide qualitative information regarding the relative amount of gene or protein expression (such as those in any of Table(s) 3, 4, 11 and/or 13) in the sample, such as identifying presence of an increase in gene or protein expression (such as those in any of any of Table(s) 4, 11 and/or 13) relative to a control, a decrease in gene or protein expression (such as those in any of Table(s) 4, 11 and/or 13) relative to a control, or no change in gene or protein expression (such as those in any of Table(s) 4, 11 and/or 13) relative to a control.

In some examples, the output is accompanied by guidelines for interpreting the data, for example, numerical or other limits that indicate the presence or absence of primary melanoma. The guidelines need not specify whether a nevus or melanoma (*e.g.*, primary melanoma), is present or absent, although it may include such a diagnosis. The indicia in the output can, for example, include normal or abnormal ranges or a cutoff, which the recipient of the output may then use to interpret the results, for example, to arrive at a diagnosis or treatment plan. In other examples, the output can provide a recommended therapeutic regimen. In some examples, the test may include determination of other clinical information (such as determining the amount of one or more additional melanoma biomarkers in the sample).

***Clinical Use Steps***

Disclosed methods may result in a melanocyte-containing sample (*e.g.*, skin biopsy) being characterized as benign (*e.g.*, nevus) or malignant (*e.g.*, melanoma, such as primary melanoma) or indeterminate or suspicious (*e.g.*, suggestive of a cancer, disease, or condition), or non-diagnostic (*e.g.*, providing inadequate information concerning the presence or absence of a cancer, disease, or condition). Each of these (and other possible) results is useful to the trained clinical professional. Some method embodiments include clinically relevant steps as described in more detail below.

*Diagnosis Indications*

A diagnosis informs a subject (*e.g.*, patient) what disease or condition s/he has or may have. As more particularly described throughout this disclosure, any result of any disclosed method that characterizes a melanocyte-containing sample can be provided, *e.g.*, to a subject or health professional, as a diagnosis. Accordingly, some method embodiments contemplated providing a diagnosis (such as, benign (*e.g.*, nevus) or malignant (*e.g.*, melanoma, such as primary melanoma) or indeterminate or suspicious (*e.g.*, suggestive of a cancer, disease, or condition), or non-diagnostic (*e.g.*, providing inadequate information concerning the presence or absence of a cancer, disease, or condition) to a subject or health professional.

*Prognostic Indications*

Prognosis is the likely health outcome for a subject whose sample received a particular test result (*e.g.*, nevus versus melanoma). A poor prognosis means the long-term outlook for the subject is not good, *e.g.*, the 1-, 2-, 3- or 5-year survival is 50% or less (e.g., 40%, 30%, 25%, 20%, 15%, 10%, 5%, 2% or 1% or less). On the other hand, a good prognosis means the long-term outlook for the subject is fair to good, *e.g.*, the 1-, 2-, 3- or 5-year survival is greater than 30%, 40%, 50%, 60%, 70%, 75%, 80% or 90%.

A subject whose melanocyte-containing sample is characterized as malignant (*e.g.*, melanoma) is likely to have a poorer prognosis (with respect to that disease or condition) than a subject whose melanocyte-containing sample is characterized as benign (*e.g.*, nevus). Accordingly, particular method embodiments include prognosing a comparatively poor outcome (see above) for a subject from whom a test sample characterized as malignant (*e.g.*, melanoma, such as primary melanoma, or the like) is taken. Conversely, other exemplary methods include prognosing a comparatively good outcome (see above) for a subject from whom a test sample characterized as benign (*e.g.*, nevus or the like) is taken.

*Therapeutic (Predictive) Indications*

The disclosed methods can further include selecting subjects for treatment for melanoma (*e.g.*, primary melanoma), if the sample is diagnosed as a melanoma (*e.g.*, primary melanoma). Alternatively, the disclosed methods can further include selecting subjects for no treatment, if the sample is diagnosed as a benign nevus.

In some embodiments, the disclosed methods of diagnosis include one or more of the following depending on the patient's diagnosis: a) prescribing a treatment regimen for the subject if the subject's determined diagnosis is positive for a primary melanoma (such as treatment with one or more chemotherapeutic agents, additional surgery to remove more tissue, or combinations thereof); b) not prescribing a treatment regimen for the subject if the subject's determined diagnosis is negative for primary melanoma or is positive for a benign nevus; c) administering a treatment (such as treatment with one or more chemotherapeutic agents, additional surgery to remove more tissue, or combinations thereof) to the subject if the subject's determined diagnosis is positive for primary melanoma; and d) not administering a treatment regimen to the subject if the subject's determined diagnosis is primary melanoma or is positive for a benign nevus. In an alternative embodiment, the method can include recommending one or more of (a)-(d). Thus, the disclosed methods can further include treating a subject for primary melanoma, if the sample from the subject is characterized as being a primary melanoma.

In some examples, chemotherapy is used to treat a subject diagnosed with melanoma using a disclosed method. In cancer treatment, chemotherapy refers to the administration of one or more agents (chemotherapeutic agents) to kill or slow the reproduction of rapidly multiplying cells, such as tumor or cancer cells. In a particular example, chemotherapy refers to the administration of one or more agents to significantly reduce the number of tumor cells in the subject, such as by at least about 50%. "Chemotherapeutic agents" include any chemical agent with therapeutic usefulness in the treatment of cancer. Examples of chemotherapeutic agents can be found for example in Fischer et al. (eds), *The Cancer Chemotherapy Handbook*, 6th ed., Philadelphia: Mosby 2003, and/or Skeel and Khleif (eds), *Handbook of Cancer Chemotherapy*, 8th ed., Philadelphia: Lippincott, Williams & Wilkins (2011)).

Chemotherapies, typically used to treat melanoma include interleukin 2 (IL2), dacarbazine, interferon, ipilimumab, carboplatin with taxol, granulocyte macrophage colony stimulating factor (GMCSF), and/or vemurafenib. Use of chemotherapeutic agent in a subject can decrease a sign or a symptom of a cancer, such as melanoma, or can reduce, stop or reverse the progression, metastasis and/or growth of a cancer, such as inhibiting metastasis.

***Arrays***

Disclosed herein are arrays that can be used to detect expression (such as expression of two or more of the sample-type-specific biomarkers in Table(s) 4, 11 and/or 13), for example, for use in characterizing a melanocyte-containing sample as a benign nevus or a primary melanoma as discussed above. In some embodiments, the disclosed arrays can also be used to detect expression of one or more normalization biomarkers (*e.g.*, those in Table 3). In other embodiments, the disclosed arrays can also be used to detect expression of sets of genes described througout this disclosure, such as in Table 6, 8 or 14. In particular examples, the array surface comprises a plate, a bead (or plurality of beads), or flow cell (*e.g.*, with multiple channels).

In some embodiments an array can include a solid surface including specifically discrete regions or addressable locations, each region having at least one immobilized oligonucleotide capable of directly hybridizing to biomarkers in Table(s) 4, 11 and/or 13, and in some examples to a normalization gene shown in Table 3. In some examples, the array includes immobilized capture probes capable of directly or indirectly specifically hybridizing with all 32 biomarkers listed in Table 4, and all normalization biomarkers in Table 3, or all of the biomarkers listed in Table 11, and all normalization biomarkers in Table 3. The oligonucleotide probes are identifiable by position on the array. In another example, an array can include specifically discrete regions, each region having at least one or at least two immobilized capture probes. The immobilized capture probes are capable of directly or indirectly specifically hybridizing with at least two different biomarkers in Table(s) 4, 11 and/or 13, and in some examples to a normalization gene shown in Table 3. The capture probes are identifiable by position on the array. The probes on the array can be attached to the surface in an addressable manner. For example, each addressable location can be a separately identifiable bead or a channel in a flow cell.

For example, the array can include at least three addressable locations, each location having immobilized capture probes with the same specificity, and each location having capture probes having a specificity that differs from capture probes at each other location. The capture probes at two of the at least three locations are capable of directly or indirectly specifically hybridizing a biomarker listed in Table(s) 4, 11 and/or 13, and the capture probes at one of the at least three locations is capable of directly or indirectly specifically hybridizing a normalization biomarker listed in Table 3. In addition, the specificity of each capture probe is identifiable by the addressable location the array. In some examples the array further includes at least two discrete regions (such wells on a multi-well surface, or channels in a flow cell), each region having the at least three addressable locations. In some example, such an array includes immobilized capture probes

capable of directly or indirectly specifically hybridizing with all biomarkers listed in Table 4, 6, 8, 11, 13, or 14 and at least two normalizers (*e.g.*, RPS6KB2 and SDHA) in Table 3. In some examples, the capture probe(s) indirectly hybridize with the at least two biomarkers listed in Table(s) 4, 11 and/or 13 and the at least one normalization biomarker in Table 3 through a nucleic acid programming linker, wherein the programming linker is a hetro-bifunctional linker which has a first portion complementary to the capture probe(s) and a second portion complementary to a nuclease protection probe (NPP), wherein the NPP is complementary to one of the at least two biomarkers listed in Table(s) 4, 11 and/or 13 or the at least one normalization biomarker in Table 3. Thus, in some examples the array also includes the nucleic acid programming linkers.

In some embodiments the array includes oligonucleotides that include or consist essentially of oligonucleotides that are complementary to at least 2 at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, or, as applicable, at least 21, at least 22, at least 23, at least 24, at least 25, at least 26, at least 27, at least 28, at least 29, at least 30, at least 31, or all of the biomarkers in Table(s) 4, 11 and/or 13 (such as 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or, as applicable, 20, 21, 22, 23,24 ,25, 26, 27, 28, 29, 30, 31, or all of the biomarkers in Table(s) 4, 11 and/or 13). In some examples, the array further includes oligonucleotides that are complementary to normalization biomarkers, such as at least 1, at least 2 at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, or all of the biomarkers in Table 3 (such as 1, 2, 3, 4, 5, 6, 7, 8 or 9 of the normalization biomarkers in Table 3, or RPS6KB2 and SDHA). In some examples, the array further includes one or more control oligonucleotides (such as 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more control oligonucleotides), for example, one or more positive and/or negative controls. In some examples, the control oligonucleotides are complementary to one or more of DEAD box polypeptide 5 (DDX5), glyceraldehyde-3-phosphate dehydrogenase (GAPDH), fibrillin 1 (FBN1), or *Arabidopsis thaliana* AP2-like ethylene-responsive transcription factor (ANT).

In some embodiments, the array can include a surface having spatially discrete regions (such as wells on a multi-well surface, or channels in a flow cell), each region including an anchor stably (*e.g.,* covalently) attached to the surface and nucleic acid programming linker, wherein the programming linker is a hetro-bifunctional linker which has a first portion complementary to the capture probe(s) and a second portion complementary to a nuclease protection probe (NPP), wherein the NPP is complementary to a target nucleic acid (such as those in Table(s) 4, 11, and/or 13). In some embodiments the array includes or consists essentially of bifunctional linkers in

which the first portion is complementary to an anchor and the second portion is complementary to an NPP, wherein the NPP is complementary to one of the at least 2 at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, or, as applicable, at least 20, at least 21, at least 22, at least 23, at least 24, at least 25, at least 26, at least 27, at least 28, at least 29, at least 30, at least 31, or all 32 of the biomarkers in Table(s) 4, 11, and/or 13 (such as 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or, as applicable, 20, 21, 22, 23,24 ,25, 26, 27, 28, 29, 30, 31, or 32 of the biomarkers in Table(s) 4, 11, and/or 13). In some examples, the array further includes bifunctional linkers in which the first portion is complementary to an anchor and the second portion is complementary to an NPP complementary to a normalization biomarker, such as at least 1, at least 2 at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, or all of the biomarkers in Tables 3 (such as 1, 2, 3, 4, 5, 6, 7, 8 or 9 of the biomarkers in Table 3). Such arrays have attached thereto the anchor hybridized to at least a segment of the bifunctional linker that is not complementary to the NPP. In another example, the array further includes bifunctional linkers in which the second portion of the bifunctional linker is complementary to an NPP complementary to a control gene (such as DDX5, GAPDH, FBN1, or ANT). Such arrays can further include (1) the anchor probe hybridized to the first portion of the programming linker, (2) NPPs hybridized to the second portion of the programming linker (which in some examples are biotinylated), (3) bifunctional detection linkers having a first portion hybridized to the NPPs and a second portion hybridized to a detection probe, (4) a detection probe; (5) a label (such as avidin HRP), or combinations thereof.

In some examples, a collection of up to 47 different capture (*i.e.*, anchor) oligonucleotides can be spotted onto the surface at spatially distinct locations and stably associated with (*e.g.*, covalently attached to) the derivatized surface. For any particular assay, a given set of capture probes can be used to program the surface of each well to be specific for as many as 47 different targets or assay types of interest, and different test samples can be applied to each of the 96 wells in each plate. The same set of capture probes can be used multiple times to re-program the surface of the wells for other targets and assays of interest.

*Array substrates*

The solid support of the array can be formed from an organic polymer. Suitable materials for the solid support include, but are not limited to: polypropylene, polyethylene, polybutylene, polyisobutylene, polybutadiene, polyisoprene, polyvinylpyrrolidine, polytetrafluroethylene, polyvinylidene difluroide, polyfluoroethylene-propylene, polyethylenevinyl alcohol,

polymethylpentene, polycholorotrifluoroethylene, polysulfornes, hydroxylated biaxially oriented polypropylene, aminated biaxially oriented polypropylene, thiolated biaxially oriented polypropylene, ethyleneacrylic acid, thylene methacrylic acid, and blends of copolymers thereof (see U.S. Patent No. 5,985,567). Other examples of suitable substrates for the arrays disclosed herein include glass (such as functionalized glass), Si, Ge, GaAs, GaP, $SiO_2$, $SiN_4$, modified silicon nitrocellulose, polystyrene, polycarbonate, nylon, fiber, or combinations thereof. Array substrates can be stiff and relatively inflexible (for example glass or a supported membrane) or flexible (such as a polymer membrane).

In general, suitable characteristics of the material that can be used to form the solid support surface include: being amenable to surface activation such that upon activation, the surface of the support is capable of stably (*e.g.,* covalently, electrostatically, reversibly, irreversibly, or permanently) attaching a biomolecule such as an oligonucleotide thereto; amenability to "*in situ*" synthesis of biomolecules; being chemically inert such that at the areas on the support not occupied by the oligonucleotides or proteins (such as antibodies) are not amenable to non-specific binding, or when non-specific binding occurs, such materials can be readily removed from the surface without removing the oligonucleotides or proteins (such as antibodies).

In another example, a surface activated organic polymer is used as the solid support surface. One example of a surface activated organic polymer is a polypropylene material aminated via radio frequency plasma discharge. Other reactive groups can also be used, such as carboxylated, hydroxylated, thiolated, or active ester groups.

*Array Formats*

Within an array, each arrayed sample is addressable, in that its location can be reliably and consistently determined within dimensions (e.g., at least two dimensions) of the array. The feature application location on an array can assume different shapes. For example, the array can be regular (such as arranged in uniform rows and columns, or be set forth in a plurality of individually identifiable beads) or irregular. Thus, in ordered arrays the location of each sample is assigned to the sample at the time when it is applied to the array, and a key may be provided in order to correlate each location with the appropriate target or feature position. Often, ordered arrays are arranged in a symmetrical grid pattern, but samples could be arranged in other patterns (such as in radially distributed lines, spiral lines, or ordered clusters). Addressable arrays usually are computer readable, in that a computer can be programmed to correlate a particular address on the array with information about the sample at that position (such as hybridization or binding data, including for instance signal intensity). In some examples of computer readable formats, the individual features

in the array are arranged regularly, for instance in a Cartesian grid pattern, which can be correlated to address information by a computer.

One example includes a linear array of oligonucleotide bands, generally referred to in the art as a dipstick. Another suitable format includes a two-dimensional pattern of discrete cells (such as 4096 squares in a 64 by 64 array). In one example, the array includes up to 47 (*e.g.*, 5, between 5 and 16, between 5 and 47, 16, between 16 and 47) addressable locations per reaction chamber; thus, in a 96-well array, there may be 96 x 5, 96 x 16, 96 x 47addressable locations with the addressable locations within each reaction chamber (*e.g.*, well) being the same or different (*e.g.*, using programmable array technologies); provided, however, it is understood in that art that universally programmable arrays may be flexibly programmed to capture any number of analytes up to the number of addressable locations that can physically be printed on the array surface of interest. Other embodiments include arrays comprising physically separate surfaces combined together into a set of surfaces that when combined create an addressable array; for example, a set of individually identifiable (*e.g.*, addressable) beads, each programmed or printed to capture a specific analyte. As is appreciated by those skilled in the art, other array formats including, but not limited to slot (rectangular) and circular arrays are equally suitable for use (see U.S. Patent No. 5,981,185). In some examples, the array is a multi-well plate (such as a 96-well plate). In one example, the array is formed on a polymer medium, which is a thread, membrane or film. An example of an organic polymer medium is a polypropylene sheet having a thickness on the order of about 1 mil. (0.001 inch) to about 20 mil., although the thickness of the film is not critical and can be varied over a fairly broad range. The array can include biaxially oriented polypropylene (BOPP) films, which in addition to their durability, exhibit low background fluorescence.

The array formats of the present disclosure can be included in a variety of different types of formats. A "format" includes any format to which the solid support can be affixed, such as microtiter plates (*e.g.*, multi-well plates), test tubes, inorganic sheets, dipsticks, beads, and the like. For example, when the solid support is a polypropylene thread, one or more polypropylene threads can be affixed to a plastic dipstick-type device; polypropylene membranes can be affixed to glass slides. The particular format is, in and of itself, unimportant. All that is necessary is that the solid support can be affixed thereto without affecting the functional behavior of the solid support or any biopolymer absorbed thereon, and that the format (such as the dipstick or slide) is stable to any materials into which the device is introduced (such as clinical samples and hybridization solutions).

The arrays of the present disclosure can be prepared by a variety of approaches. In one example, oligonucleotide sequences are synthesized separately and then attached to a solid support

(see U.S. Patent No. 6,013,789). In another example, sequences are synthesized directly onto the support to provide the desired array (see U.S. Patent No. 5,554,501). Suitable methods for coupling oligonucleotides to a solid support and for directly synthesizing the oligonucleotides onto the support are known to those working in the field; a summary of suitable methods can be found in Matson *et al.*, *Anal. Biochem.* 217:306-10, 1994. In one example, the oligonucleotides are synthesized onto the support using conventional chemical techniques for preparing oligonucleotides on solid supports (such as PCT applications WO 85/01051 and WO 89/10977, or U.S. Patent No. 5,554,501).

A suitable array can be produced using automated means to synthesize oligonucleotides in the cells of the array by laying down the precursors for the four bases in a predetermined pattern. Briefly, a multiple-channel automated chemical delivery system is employed to create oligonucleotide probe populations in parallel rows (corresponding in number to the number of channels in the delivery system) across the substrate. Following completion of oligonucleotide synthesis in a first direction, the substrate can then be rotated by 90° to permit synthesis to proceed within a second set of rows that are now perpendicular to the first set. This process creates a multiple-channel array whose intersection generates a plurality of discrete cells.

The oligonucleotides can be bound to the support by either the 3'-end of the oligonucleotide or by the 5' end of the oligonucleotide. In one example, the oligonucleotides are bound to the solid support by the 3'-end. However, one of skill in the art can determine whether the use of the 3'-end or the 5'-end of the oligonucleotide is suitable for bonding to the solid support. In general, the internal complementarity of an oligonucleotide probe in the region of the 3'-end and the 5'-end determines binding to the support.

***Kits***

Also disclosed herein are kits that can be used to detect expression (such as expression of two or more of the biomarkers in Table(s) 4, 11 and/or 13), for example for use in characterizing a sample as a benign nevus or a primary melanoma as discussed above. In some embodiments, the disclosed kits can also be used to detect expression of one or more normalization biomarkers (*e.g.*, those in Table 3). In particular examples, the kit includes one or more of the arrays provided herein.

In some examples the kits include probes and/or primers for the detection of nucleic acid or protein expression, such as two or more of the biomarkers in Table(s) 4, 11 and/or 13, and in some examples, one or more normalization biomarkers in Table 3. In some examples, the kits include antibodies that specifically bind to biomarkers listed in Table(s) 4, 11 and/or 13. For example, the

kits can include one or more nucleic acid probes needed to construct an array for detecting the biomarkers disclosed herein.

In some examples, the kit includes nucleic acid programming linkers. The programming linkers are hetro-bifunctional having a first portion complementary to the capture probe(s) on the array and a second portion complementary to a nuclease protection probe (NPP), wherein the NPP is complementary to one of the at least two biomarkers listed in Table(s) 4, 11 and/or 13 or to at least one normalization biomarker in Table 3. In one example, the programming linkers are pre-hybridized to the capture probes, such that they are not covalently attached so that the surface includes the addressable immobilized capture probes and the nucleic acid programming linkers. In such an example, the kit does not have a separate container with programming linkers

In some examples, the kit includes NPPs. The NPPs are complementary to the second portion of the programming linker. Exemplary NPPs are shown in SEQ ID NOS: 1-36, and 123-164.

In some examples, the kit includes bifunctional detection linkers. Such linkers can be labeled with a detection probe and are capable of specifically hybridizing to the NPPs or to the target (such as those in Table(s) 4, 11 and/or 13).

In some examples, the kit includes an array disclosed herein, and one or more of a container containing a buffer (such as a lysis buffer); a container containing a nuclease specific for single-stranded nucleic acids; a container containing nucleic acid programing linkers; a container containing NPPs; a container containing a plurality of bifunctional detection linkers; a container containing a detection probe (such as one that is triple biotinylated); and a container containing a detection reagent (such as avidin HRP).

In one example, the kit includes a graph or table showing expected values or ranges of values of the biomarkers in Table(s) 4, 11 and/or 13 expected in a normal skin cell (*e.g.*, benign nevus) or a primary melanoma, or clinically useful cutoffs. In some examples, kits further include control samples, such as particular quantities of nucleic acids or proteins for those biomarkers in Table(s) 4, 11 and/or 13.

The kits may further include additional components such as instructional materials and additional reagents, for example detection reagents, such as an enzyme-based detection system (for example, detection reagents including horseradish peroxidase or alkaline phosphatase and appropriate substrate), secondary antibodies (for example antibodies that specifically bind the primary antibodies that specifically bind the targets (*e.g.*, proteins) in Table(s) 3, 4, 11, and/or 13), or a means for labeling antibodies. The kits may also include additional components to facilitate

the particular application for which the kit is designed (for example microtiter plates). In one example, the kit of further includes control nucleic acids. Such kits and appropriate contents are well known to those of ordinary skill in the art. The instructional materials may be written, in an electronic form (such as a computer diskette or compact disk) or may be visual (such as video files).

The following examples are provided to illustrate certain particular features and/or embodiments. These examples should not be construed to limit the invention to the particular features or embodiments described.

## EXAMPLES

### Example 1

### Gene Selection Using a Discovery Set of Clinically Characterized Skin Samples

Nevi and melanoma cells, like all cells, express a vast number of genes, most of which are not relevant to distinguishing between such groups. Thus, in order to extract useful gene information and reduce dimensionality, this Example describes the initial screening of the expression of greater than 2600 mRNA targets to identify significantly differentially expressed mRNAs in formalin-fixed, paraffin-embedded ("FFPE") skin samples biopsied from human subjects. Further described are methodological details used throughout the Examples.

A discovery set of 39 FFPE tissue sections, each approximately 5 um thick and mounted on a microscope slide, was provided by the John Wayne Cancer Institute (JWCI) tissue bank. The set included 14 normal skin samples, 10 nevi samples, 5 primary melanoma samples, and 10 samples of melanoma metastases.

*Sample Preparation and Lysis*

Briefly, each FFPE tissue section was measured to determine its approximate area (in $cm^2$). The tissue section then was scraped into a labeled eppendorf tube using a razor blade and avoiding any excess paraffin on the slide. The sample was suspended in 25 ul pre-warmed (50°C) SSC buffer including formamide and SDS per each 0.3 $cm^2$ of the applicable tissue section. Five-hundred (500) ul of mineral oil containing a surfactant (*e.g.*, Brij-97) ("Non-aqueous Layer") then was overlaid on the tissue suspension, and this lysis reaction was incubated at 95°C for 10-15 minutes. After briefly cooling the reaction mixture, proteinase K was added to a final concentration of 1 mg/ml and the incubation continued at 50 C for 30-60 minutes. A portion of the lysis reaction was used immediately in a nuclease protection assay (see below), or the lysis reaction

(or remaining portion thereof) was frozen and stored at -80°C. Frozen lysis reactions were thawed at 50°C for 10-15 minutes before a subsequent use.

*Nuclease Protection Assay ("NPA")*

Twenty-five (25) ul of each lysed reaction mixture was placed in a well of a 96-well plate and overlaid with 70 ul Non-aqueous Layer. To each well was added 5 ul of nuclease protection probe (NPP) mix. One (1) nM (an excess of) NPP complementary to each of the plurality of mRNA targets to be detected was present in the NPP mix. NPPs for ArrayPlate detection were (i) 50-base pairs in length with each half of the NPP having a Tm in the range of 40°C-75°C (and full length Tms in the range of 60°C-85°C) and (ii) tested in silico (using NCBI BLAST) and with in vitro transcripts for specificity to the respective mRNA target (and substantially no cross-reactivity with other NPPs, other targets, or other analytes in the NPA reaction). NPPs for ArraySlide detection differ only in that they contain an internal biotinylated base (T) biased toward the 3' end of the NPP. NPPs are further described in connection with genes specifically identified in other Examples.

The 96-well NPA plate was heated at 95°C for 10-15 minutes to denature nucleic acids and, then, allowed to incubate at 60°C for 6-16 hours to permit hybridization of the NPPs to their respective mRNA targets.

Following the hybridization step, 20 ul of excess S1 nuclease (2.5 U/ul) in sodium acetate buffer was added to the aqueous phase of each well. The S1 reaction proceeded at 50°C for 90-120 minutes to digest unbound mRNA and unbound NPPs. In some sets of reactions, BSA in molecular-biology-grade water was added to a final concentration of 40 mg/ml.

During the S1 digestion step, a 96-well "Stop" plate was prepared by adding 10 ul of solution contain 0.1 M EDTA and 1.6 N NaOH to each well corresponding to the reactions in the 96-well NPA plate. The entire volume (approx. 120 ul) of each reaction in the 96-well NPA plate was transferred to a corresponding well in the second 96-well Stop plate. The Stop plate was incubated at 95°C for 15-20 minutes and, then, cooled for 5-10 minutes at room temperature prior to the addition of 10 ul 1.6 N HCl to neutralize the NaOH previously added to each reaction.

The nuclease protection assay reactions in this Example were interrogated directly (*e.g.*, without purification or reverse transcription of target mRNA analytes) using (i) a first, 96-well-plate-based array (ArrayPlate No. 1) custom designed to detect in each well the expression of 34 human putative melanoma-related mRNAs (or controls), (ii) a second ArrayPlate (*i.e.*, No. 2) custom designed to detect 33 human putative melanoma-related mRNAs (or controls), (iii) a first, glass-slide-based, 21-well (ArraySlide No. 1) custom "cancer transcriptome" array capable of

detecting 1829 human putative cancer-related mRNAs (or controls) in each well, and (iv) a second ArraySlide (*i.e.*, No. 2) "whole transcriptome" custom array capable of detecting in each 2600 mRNAs putatively representative of the human transcriptome. The targets to be interrogated by each of the foregoing arrays was determined, *e.g.*, on the basis of literature searches and public knowledge.

*ArrayPlate Capture and Detection*

ArrayPlate Nos. 1 and 2 were programmed with 40 ul 50-base pair programming linkers ("PL") at 5nM in SSC buffer containing SDS ("SSC-S"). The PLs were artificial, 25-base pair, bi-functional synthetic oligonucleotide constructs (adaptors) complementary in part to a universal anchor sequence affixed to the array surface and complementary in the other part to the particular NPP addressed to the particular array location. Following the programming step, the entire aqueous phase (60-65 ul) of each reaction from the Stop plate was added to a corresponding well of the programmed ArrayPlate and incubated at 50°C for 16-24 hour to capture undigested NPPs (which were bound to target during the nuclease step and, therefore, are quantifiable surrogates for targets present in the sample). Thereafter, 5 nM bi-functional detection linker ("DL") in SSC-S including 1% nonfat dry milk was added to each reaction followed by 1 hour incubation at 60°C. The DLs were artificial 25-base pair, bi-functional synthetic oligonucleotide constructs complementary in part to its respective NPP and complementary in the other part to one or more (*e.g.*, two or three) copies of a biotin-labeled detection probe ("DP"), which DP was capable of specifically binding the detection-region designed into all DLs. To complete the detection "sandwich," 40 ul of 3 nM DP was added to the reactions followed by 50°C incubation for 45-60 min. Next, 40 ul avidin peroxidase (1:600) in SSC-S including 1% nonfat dry milk was added followed by incubation at 37°C for 30-45 minute. Finally, a chemiluminescent substrate mix was added that, in the presence of peroxidase enzyme, generated light that was captured using a HTG OMIX™ imager. Gene expression is directly related to the intensity of light emitted at each addressable position of the ArrayPlate.

*ArraySlide Capture and Detection*

The entire aqueous phase of each nuclease protection assay reaction (60-65 ul) was then hybridized to ArraySlide No. 1 or No. 2 for 16-24 hour at 50°C for capture of the NPPs. After capture of the biotinylated NPPs, the respective ArraySlide was washed rigorously with 1x SSC containing 1% Tween ("Wash Buffer"). Fifty (50) ul of avidin-peroxidase (1:600) in detection enzyme buffer (1x SSC-S, 0.05% Tween and non-fat, dry milk) was added for 45 minutes at 37°C. ArraySlides were washed followed by addition of TSA-Plus Cy3 reagent in amplification diluent

(Perkin Elmer) for detection. After a 3-minute room temperature incubation, TSA-Plus Cy3 reactions were stopped by washing the ArraySlides in Wash Buffer. Finally, the ArraySlides were spun dry and scanned at 5 um resolution using a GenePix 4200AL microarray slide scanner (Molecular Devices, Sunnyvale, CA). Probe intensities were extracted from TIFF images using NimbleScan 2.5 software (Roche NimbleGen, Madison, WI) for analysis as described below.

*Data Analysis*

Raw data from each of the arrays in this Example were processed using BRB array tools (freely available for research use, as of June 4, 2012, on the internet at linus.nci.nih.gov/~brb/download_full_v4_2_1_stable.html). Briefly, data was subjected to minimum intensity thresholding, quantile normalization and certain data filters were applied to remove non-differential data points from further analysis. Data was log2 transformed and analyzed to find statistically significant differential genes among the group arrays based on p-values and log fold change values.

Seventy-eight (78) genes were selected for further study based on (a) significant ($p \leq 0.05$) differential expression in nevi versus primary melanoma samples and, in some cases, (b) mRNA expression that exceeded 3000 raw signal intensity in each sample population in which such expression was measured. An additional four (4) genes (SDHA, RPS6KB2, RPL37A, and TFRC) originally included as putative controls also were carried forward for further study.

## Example 2

## Genes Significantly Differentially Expressed in a Second Set of Clinically Characterized Skin Samples - Normalization to Four

This Example describes the identification of a set of 32 genes, the mRNA expression of which is significantly different between human skin biopsies characterized by the JWCI tissue bank as either nevi or primary melanomas.

Two custom ArrayPlates (referred to as ArrayPlates No. 3 and 4) were constructed to measure the expression of the 82 mRNA targets identified in or carried forward from Example 1 plus 6 additional targets identified by pathway analysis or used as negative controls. The gene lists for ArrayPlates No. 3 and 4 are shown in Table 1 below:

Table 1. ArrayPlate Gene Lists

| ArrayPlate No. 3 | |
|---|---|
| **Symbol** | **GenBank Accession No.** |
| SDHA | NM_004168 |
| RPS6KB2 | NM_003952 |
| RPL37A | NM_000998 |
| TFRC | NM_003234 |
| ANT | NM_119937 |
| MAGEA2 | NM_005361 |
| PAX3 | NM_181457 |
| CDK2 | NM_001798 |
| PRAME | NM_206953 |
| MFI2 | NM_005929 |
| MCM6 | NM_005915 |
| S100B | NM_006272 |
| PDIA4 | NM_004911 |
| SOX4 | NM_003107 |
| BRAF | NM_004333 |
| PPIA | NM_021130 |
| MAGED2 | NM_014599 |
| GALNTL1 | NM_001168368 |
| PTEN | NM_000314 |
| HRAS | NM_005343 |
| TP53 | NM_000546 |
| CTNNB1 | NM_001904 |
| TYR | NM_000372 |
| TEX13A | NM_031274 |
| BMP1 | NM_001199 |
| TGFB1 | NM_000660 |
| NR4A1 | NM_002135 |
| PIP4K2A | NM_005028 |
| PDLIM7 | NM_213636 |
| TADA3L | NM_006354 |
| B4GALT1 | NM_001497 |
| RAP2B | NM_002886 |
| B2M | NM_004048 |
| NCOR2 | NM_001077261 |
| SP100 | NM_003113 |
| SAT1 | NM_002970 |
| STAT2 | NM_005419 |
| RUNX1 | NM_001001890 |
| GNAS | NM_016592 |
| SOCS3 | NM_003955 |

| ArrayPlate No. 4 | |
|---|---|
| **Symbol** | **GenBank Accession No.** |
| SDHA | NM_004168 |
| RPS6KB2 | NM_003952 |
| RPL37A | NM_000998 |
| TFRC | NM_003234 |
| ANT | NM_119937 |
| BIRC7 | NM_139317 |
| BIRC5 | NM_001168 |
| MET | NM_001127500 |
| HIF1A | NM_001530 |
| ALK | NM_004304 |
| DAZAP2 | NM_014764 |
| EVI2B | NM_006495 |
| LDHA | NM_005566 |
| ERCC1 | NM_001983 |
| ESR1 | NM_000125 |
| ALDOA | NM_000034 |
| CTNNB1 | NM_001904 |
| ARID1A | NM_139135 |
| NPHP1 | NM_001128179 |
| AF090940 | AF090940 |
| DUX4 | NM_033178 |
| POLR2J3 | NM_001097615 |
| HADHA | NM_000182 |
| AK027225 | AK027225 |
| IGFBP5 | NM_000599 |
| BC017937 | BC017937 |
| OAZ1 | NM_004152 |
| TACSTD2 | NM_002353 |
| ATXN2L | NM_148416 |
| PLIN2 | NM_001122 |
| PFDN6 | NM_014260 |
| HMGA1 | NM_002131 |
| ZFYVE16 | NM_014733 |
| AF168811 | AF168811 |
| BAX | NM_004324 |
| AU159040 | AU159040 |
| BRD7P3 | NR_002730 |
| RNF126 | NM_194460 |
| ETV2 | NM_014209 |
| TPSAB1 | NM_003294 |

Table 1. ArrayPlate Gene Lists

| ArrayPlate No. 3 | |
|---|---|
| **Symbol** | **GenBank Accession No.** |
| BAX | NM_004324 |
| CREBBP | NM_001079846 |
| HIST1H2BN | NM_003520 |
| HP1BP3 | NM_016287 |
| LZTS1 | NM_021020 |
| SQSTM1 | NM_003900 |
| TPSAB1 | NM_003294 |

| ArrayPlate No. 4 | |
|---|---|
| **Symbol** | **GenBank Accession No.** |
| ZFPL1 | NM_006782 |
| COX16 | NM_016468 |
| AK023563 | AK023563 |
| BEST1 | NM_004183 |
| PICALM | NM_001008660 |
| NOP56 | NM_006392 |
| PTMS | NM_002824 |

mRNA expression was measured in 100 FFPE tissue sections, consisting of 39 nevus samples (from melanoma-naive patients) and 61 primary melanoma samples.

Sample preparation and lysis, nuclease protection assay, and array capture and detection were performed substantially as described for ArrayPlates Nos. 1 and 2 in Example 1.

Table 2 shows NPP sequences for (i) targets found in this Example to be significantly differentially expressed between nevi and melanoma samples and (ii) targets whose expression was used for normalization. Other NPP sequences useful in a disclosed invention are describe elsewhere or can be determined by one of ordinary skill in the art using guidance provided in this disclosure and publicly available sequences of the disclosed targets (*e.g.*, SEQ ID NOs. shown in Tables 11 and 13).

Table 2. Exemplary Nuclease Protection Probe Sequences

| **Gene Name** | **Accession No.** | **NPP Sequence (5' - 3'; wrapped at line break)** | **SEQ ID NO.** |
|---|---|---|---|
| B2M | NM_004048 | CTGCTGGATGACGTGAGTAAACCTGAA TCTTTGGAGTACGCTGGATAGCC | 1 |
| B4GALT1 | NM_001497 | GTCTTGGAACCTGAGCCCAGGCTGGAC CTGGCAAAGGCGCTCAGTGGTAG | 2 |
| BMP1 | NM_001199 | CCGCAAGGTCGATAGGTGAACACAATA TAGCTGTCCTCGTCAGTGCGCTC | 3 |
| BRAF | NM_004333 | GTAAGTGGAACATTCTCCAACACTTCC ACATGCAATTCTTCTCCAGTAAG | 4 |
| CDK2 | NM_001798 | CAAGTTCAGAGGGCCCACCTGAGTCCA AATAGCCCAAGGCCAAGCCTGGT | 5 |
| CREBBP | NM_001079846 | CCTGGGTTGATACTAGAGCCGCTGCCT CCTCGTAGAAGCTCCGACAGTTG | 6 |
| CTNNB1 | NM_001904 | CAGCATCTGTGATGGTTCAGCCAAACG CTGGACATTAGTGGGATGAGCAG | 7 |

Table 2. Exemplary Nuclease Protection Probe Sequences

| Gene Name | Accession No. | NPP Sequence (5' - 3'; wrapped at line break) | SEQ ID NO. |
|---|---|---|---|
| GALNTL1 | NM_001168368 | GGGCTCAGCTTGTCACTCTCCAGCTGGT TGAAGGCGTGCTGTCTGTAGGG | 8 |
| GNAS | NM_016592 | CTCGCTGAGTCTTAGATTCCGCAGCCTA AGACTCGAGAGAGGTGCCTCCG | 9 |
| MAGEA2 | NM_005361 | CTCAGGCTCTCCACCTGGATGCTTGGCA GATCCTAGAACCACTGCATCTG | 10 |
| MAGED2 | NM_014599 | CTTCACCTTTCGGGCTTTCTTGGCTTTG ACCTTGGGCCGAGTATCCTGAT | 11 |
| MCM6 | NM_005915 | TCCTGGTGTGCTAAGCTTGGAGACGTC AGGCACAACAATCAGTGTCCCTG | 12 |
| MFI2 | NM_005929 | GCTGGCATTGAAGAACTCGCTCACTGC TGTGAGGACGTCACAGTCCTTGG | 13 |
| NCOR2 | NM_001077261 | CCCGGTACAGCAGCGGGTACACAGCAC TCCGGGAGTGCCCTGGCTCCGTC | 14 |
| NR4A1 | NM_002135 | CGCCACAGCTGCCACGTGCTCCTTCAG GCAGCTGGCGATGCGGTTCTGCA | 15 |
| PDIA4 | NM_004911 | CACATCAAACCTGCTGGCCAGCACAGA CGCTGAGGTTGCATCGATCTTGG | 16 |
| PDLIM7 | NM_213636 | CTTCGATGTGTGTGAGGCTACCCGCATT CTCGCCATCGATGCTCAGCACC | 17 |
| PIP4K2A | NM_005028 | ATTCACTCACTCACTCACTCACTCATTC ATTCGGCCATAGCTGGAATCAA | 18 |
| PPIA | NM_021130 | TGGTATCACCCAGGGAATACGTAACCA GACAACACACAAGACTGAGATGC | 19 |
| PRAME | NM_206953 | GTCTGGCTGTGTCTCCCGTCAAAGGCTG CCATGAAGAGTGGCGGGAAGAG | 20 |
| PTEN | NM_000314 | CTTCACCTTTAGCTGGCAGACCACAAA CTGAGGATTGCAAGTTCCGCCAC | 21 |
| RAP2B | NM_002886 | CCTCTCCTCCTGCTCCTTCATATGGTTC TCCCGGACTTCCTTCCATGTAT | 22 |
| RPL37A | NM_000998 | CTGATGGCGGACTTTACCGTGACAGCG GAAGTGGTATTGTACGTCCAGGC | 23 |
| RPS6KB2 | NM_003952 | GCTTCACATACGTGGCGCCGTCTGTCCT GGACAGCATCAAGGAGGGCTTC | 124 |
| RUNX1 | NM_001001890 | GCAGAGTCACACACATGCAAACACGCA CTCTTCGGAAGGCAGCCACTGTC | 24 |
| SAT1 | NM_002970 | ATTTCAAACATGCAACAACGCCACTGG TAATAAAGCTTTGGAATGGGTGC | 25 |
| SDHA | NM_004168 | GAAGAAGCCCTTTGAGGAGCACTGGAG GAAGCACACCCTGTCCTATGTGG | 123 |
| SOCS3 | NM_003955 | GTCTTCTCTACCAGGAGCCTGAGGTGA AAGATGTCCCGTCTCCTCCATCC | 26 |
| SOX4 | NM_003107 | CTCCGCCTCTCGAATGAAAGGGATCTT GTCGCTGTCTTTGAGCAGCTTCC | 27 |

Table 2. Exemplary Nuclease Protection Probe Sequences

| Gene Name | Accession No. | NPP Sequence (5' - 3'; wrapped at line break) | SEQ ID NO. |
|---|---|---|---|
| SP100 | NM_003113 | CCATGGTTGTGTAGCTCTGCCTCTGGGC TTTCTTCATCACAGGGCAACGG | 28 |
| SQSTM1 | NM_003900 | CCCAGGAAACATCAGCACACACACACA CAGGGACCCTCCCTTCATGTCAC | 29 |
| STAT2 | NM_005419 | CGGGATTCAATCTCATGTTGCTGGCTCT CCACAGGTGTTTCGAGAACTGG | 30 |
| TADA3 | NM_006354 | CTACCCATCCAGCAGCTTCAGGATGCT CTCACGCTCCTTCAGAGTCTTCC | 31 |
| TEX13A | NM_031274 | AGTATGAGTATGAGGCAGGGAGCTGGA CAGGAAGAGGTTCTGATGAGGCT | 32 |
| TFRC | NM_003234 | GACGTGCTGCAGGGAAGTCCTCTCCTG GCTCCTCCCTCACTGGAGACTCG | 33 |
| TGFB1 | NM_000660 | GGTAGTGAACCCGTTGATGTCCACTTG CAGTGTGTTATCCCTGCTGTCAC | 34 |
| TP53 | NM_000546 | CCCGGGACAAAGCAAATGGAAGTCCTG GGTGCTTCTGACGCACACCTATT | 35 |
| TPSAB1 | NM_003294 | CGCCAGCAGCAGCAGATTCAGCATCCT GGCCGCTCCCTGTTCCTTCTACC | 36 |

*Data Analysis*

All analysis in Examples 2 and 3was performed in SAS version 9.3 unless otherwise specified.

A. Transformation and Quality Control

The data was processed using a HTG OMIX™ imaging device and a 16 bit image was extracted. As is standard practice in genomic research, the raw intensity values were log base 2 transformed in order to make the scale of the data more linear. Each gene had three independent observations and all three observations were averaged with a geometric mean (although an arithmetic mean would serve equally well) to create a composite average log base 2 expression value for each gene. The plant gene ANT (AP2-like ethylene-responsive transcription factor; GenBank mRNA RefSeq No. NM_119937; SEQ ID NO. 122) was used as a negative control on each array. Samples for which ANT was detectible above background was used to screen and remove assay failures. Descriptive statistical analyses were also conducted to screen for errors in the data file.

B. Selection of Genes for Normalization

The scientific dogma that any gene remains constant in its expression across all sample types or subjects (*i.e.*, universal "housekeeper" gene) is losing favor (*e.g.*, Avison, Measuring Gene

Expression, Psychology Press, 2007, p. 128). Thus, other alternatives for selecting genes suitable for normalization, especially, of microarray data have been developed. Some suitable methods are described herein and others are known to those of ordinary skill in the art.

Expression of "normalization" genes were used to normalize the data to uncontrollable process variables such as cellular content in sample loads. The first step in screening candidate normalization genes for this Example was to run a Satterthwaite T-Test to determine that there was no statistically significant difference in expression of such candidate normalization genes between the samples in the populations of interest, *i.e.*, nevi and primary melanoma samples. Initially, this analysis was performed using an average of triplicate raw expression values and later confirmed with normalized expression values. A p-value exceeding 0.05 was set as a lower bound for determining a lack of significance.

Expression levels for candidate normalization genes were then inspected to ensure adequate and non-saturated intensity values. Adequate and non-saturated intensity values were defined as 1.5 expression units above background and below saturation.

Candidate normalization genes were also selected on the basis of minimal standard deviations. An upper bound of 2.0 expression units was set as a cutoff. Candidate normalization genes with standard deviations larger than this cutoff were removed from consideration. The goal was to select among remaining candidate normalization genes those which had the lowest standard errors between the sample populations of interest (*i.e.*, nevi and primary melanoma samples).

It is noted that a coefficient of variation (CV) can also be used in place of a standard deviation in this and other applicable analyses. A CV is a statistical method for describing the dispersion of data or a variable irrespective of the unit of measurement. Since a CV is calculated by dividing the standard deviation (or in some SAS procedures the root mean square error) by the mean and the unit of expression measurement for genes across an array is very similar, using a CV or a standard deviation rarely, if ever, results in qualitatively different patterns of results in which one would be led to draw different conclusions as to the validity of a housekeeper.

An exemplary normalization genes (also referred to as "normalizers") selected throughout these Examples as representative for human nevi and primary melanoma skin biopsies are shown in Table 3. The box plots, means plots and SAS diffograms for therepresentative normalizers BMP-1, MF12, NCOR2 and RAP2b are shown in FIGS. 2A and 2B. In the SAS diffograms, for example, the dashed diagonal line (from bottom left to top right; colored blue) represents p=0.5; the x and y axes plot the normalized average log2 intensity value; and lines on each axis denote the mean normalized average log2 intensity value for each group as indicated. The solid diagonal line (from

top left to bottom right; colored red), crosses the dashed p=0.5 reference line, which illustrates no statistically significant difference between nevi and primary samples with p>0.05.

Table 3. Representative Normalization Genes Identified for Nevi and Primary Melanoma Samples

| Symbol | Name | GenBank Ref. No. | SEQ ID NO(s). |
|---|---|---|---|
| BMP-1 | Homo sapiens bone morphogenetic protein 1, variant 1 | NM_001199 (var 1) NR_033404 (var 5/nc); NR_033403 (var 4/nc); NM_006129 (var 3) | 40-43 |
| MFI2 | Homo sapiens antigen p97 (melanoma associated) identified by monoclonal antibodies 133.2 and 96.5 (MFI2) | NM_005929 (var 1) NM_033316 (var 2) | 37, 38 |
| NCOR2 | Homo sapiens nuclear receptor corepressor 2 | NM_001077261 (var 2) NM_001206654 (var 3); NM_006312 (var 1) | 44-46 |
| RAP2b | Homo sapiens RAP2B, member of RAS oncogene family | NM_002886 | 39 |
| RPS6KB2 | Homo sapiens ribosomal protein S6 kinase, 70kDa, polypeptide 2 | NM_003952 | 120 |
| SDHA | Homo sapiens succinate dehydrogenase complex, subunit A, flavoprotein (Fp) (SDHA), nuclear gene encoding mitochondrial protein | NM_004168 | 121 |
| RPL19 | Ribosomal Protein L19 | NM_000981 | |
| RPLP0 | Large Ribosomal Phosphoprotein P0 | NM_001002 (var 1); NM_053275 (var 2) | |
| ALDOA | Fructose-bisphosphate Aldolase A (aka, Fructose-1,6-Bisphosphate Aldolase A; ALDA; Aldolase 1; Fructoaldolase A | NM_000034 (var 1); NM_184041 (var 2); NM_184043 (var 3); NM_001127617 (var 4) NM_001243177 (var 6) | |

C. Univariate Screening of Genes

To normalize the data with the foregoing normalizers, the average log2 expression value for all replicates for each gene was divided by the geometric mean of the BMP-1, MF12, NCOR2 and RAP2b normalizers (this is also known to some in the art as "normalization to some" and may be referred to as "normalization to four" herein). As previously mentioned, an arithmetic mean also would suffice for the foregoing purposes. The resulting value was multiplied by a constant of 10.

Following normalization, each other (non-normalizer) gene was screened to determine if there was a statistically significant difference in expression of that gene between nevi and primary

melanoma samples. A statistically significant difference indicates that the gene has some ability to differentiate between the two groups. A Bonferroni correction was used to select a nominal level of alpha (p-value cutoff for significance) in order to protect against alpha inflation and multiple testing. A Satterthwaite T-Test was used to screen each gene in a univariate fashion. A Satterthwaite corrected T-Test was used to ensure accurate estimates in the case of unequal variances between groups.

Table 4 shows the list of genes that were found to have statistically significant differences in mRNA expression between nevi and primary melanoma samples.

Table 4. Genes Differentiating Between Nevi and Primary Melanoma Samples

| Symbol | Name | GenBank Ref. No(s). | P-value | SEQ ID NO(s). |
|---|---|---|---|---|
| B2M | Homo sapiens beta-2-microglobulin | NM_004048 | <0.01 | 119 |
| B4GALT1 | Homo sapiens UDP-Gal: betaGlcNAc beta 1,4-galactosyltransferase, polypeptide 1 | NM_001497 | <0.01 | 50 |
| BRAF | Homo sapiens v-raf murine sarcoma viral oncogene homolog B1 | NM_004333 | <0.01 | 63 |
| CDK2 | Homo sapiens cyclin-dependent kinase 2 | NM_001798 (var 1); NM_052827 (var 2) | <0.01 | 112, 113 |
| CREBBP | Homo sapiens CREB binding protein | NM_004380 (var 1); NM_001079846 (var 2) | <0.01 | 109, 110 |
| CTNNB1 | Homo sapiens catenin (cadherin-associated protein), beta 1 | NM_001904 | <0.01 | 83 |
| GALNTL1 | Homo sapiens UDP-N-acetyl-alpha-D-galactosamine: polypeptide N-acetyl-galactosaminyltransferase-like 1 | NM_001168368 (var 1); NM_020692 (var 2) | <0.01 | 103, 104 |
| GNAS | Homo sapiens GNAS complex locus | NM_000516 (var 1); NM_080425 (var 2); NM_080426 (var 3); NM_016592 (var 4); NM_001077488 (var 6) NM_001077489 (var 7); NR_003259 (var 8/nc) | <0.01 | 85-91 |
| MAGEA2 | Homo sapiens melanoma antigen family A, 2 | NM_005361 (var 1); NM_175742 (var 2); NM_175743 (var 3) | <0.01 | 105-107 |

Table 4. Genes Differentiating Between Nevi and Primary Melanoma Samples

| Symbol | Name | GenBank Ref. No(s). | P-value | SEQ ID NO(s). |
|---|---|---|---|---|
| MAGED2 | Homo sapiens melanoma antigen family D, 2 | NM_014599 (var 1); NM_177433 (var 2); NM_201222 (var 3) | <0.01 | 95-97 |
| MCM6 | Homo sapiens minichromosome maintenance complex component 6 | NM_005915 | <0.01 | 82 |
| NR4A1 | Homo sapiens nuclear receptor subfamily 4, group A, member 1 | NM_002135 (var 1); NM_173157 (var 2); NM_001202233 (var 3) | <0.01 | 47-49 |
| PDIA4 | Homo sapiens protein disulfide isomerase family A, member 4 | NM_004911 | <0.01 | 81 |
| PDLIM7 | Homo sapiens PDZ and LIM domain 7 (enigma) | NM_005451 (var 1); NM_203352 (var 2); NM_213636 (var 4) | <0.01 | 70-72 |
| PIP4K2A | Homo sapiens phosphatidylinositol-5-phosphate 4-kinase, type II, alpha | NM_005028 | <0.01 | 79 |
| PPIA | Homo sapiens peptidylprolyl isomerase A (cyclophilin A) | NM_021130 | <0.01 | 93 |
| PRAME | Homo sapiens preferentially expressed antigen in melanoma | NM_006115 (var 1); NM_206953 (var 2); NM_206954 (var 3); NM_206955 (var 4); NM_206956 (var 5) | <0.01 | 98-102 |
| PTEN | Homo sapiens phosphatase and tensin homolog | NM_000314 | <0.01 | 94 |
| RPL37A | Homo sapiens ribosomal protein L37a (RPL37A) | NM_000998 | <0.01 | 84 |
| RUNX1 | Homo sapiens runt-related transcription factor 1 | NM_001754 (var 1); NM_001001890 (var 2); NM_001122607 (var 3) | <0.01 | 66-68 |
| SAT1 | Homo sapiens spermidine/spermine N1-acetyltransferase 1 | NM_002970 (var 1) NR_027783 (var 2/nc) | <0.01 | 51, 52 |
| SOCS3 | Homo sapiens suppressor of cytokine signaling 3 | NM_003955 | <0.01 | 69 |
| SOX4 | Homo sapiens SRY (sex determining region Y)-box 4 | NM_003107 | <0.01 | 80 |

Table 4. Genes Differentiating Between Nevi and Primary Melanoma Samples

| Symbol | Name | GenBank Ref. No(s). | P-value | SEQ ID NO(s). |
|---|---|---|---|---|
| SP100 | Homo sapiens SP100 nuclear antigen | NM_001080391 (var 1); NM_003113 (var 2); NM_001206701 (var 3); NM_001206702 (var 4); NM_001206703 (var 5); NM_001206704 (var 6) | <0.01 | 73-78 |
| SQSTM1 | Homo sapiens sequestosome 1 | NM_003900 (var 1); NM_001142298 (var 2); NM_001142299 (var 3) | <0.01 | 116-118 |
| STAT2 | Homo sapiens signal transducer and activator of transcription 2 | NM_005419 (var 1); NM_198332 (var 2) | <0.01 | 114, 115 |
| TADA3 | Homo sapiens transcriptional adaptor 3 | NM_006354 (var 1); NM_133480 (var 2) | <0.01 | 61, 62 |
| TEX13A | Homo sapiens testis expressed 13A | NM_031274 | <0.01 | 108 |
| TFRC | Homo sapiens transferrin receptor (p90, CD71) | NM_003234 (var 1); NM_001128148 (var 2) | <0.01 | 64, 65 |
| TGFB1 | Homo sapiens transforming growth factor, beta 1 | NM_000660 | <0.01 | 92 |
| TP53 | Homo sapiens tumor protein p53 | NM_000546 (var 1); NM_001126112 (var 2); NM_001126114 (var 3); NM_001126113 (var 4); NM_001126115 (var 5); NM_001126116 (var 6); NM_001126117 (var 7); NM_001126118 (var 8) | <0.01 | 53-60 |
| TPSAB1 | Homo sapiens tryptase alpha/beta 1 | NM_003294 | <0.01 | 111 |

A covariance matrix for the normalized data with the disease variable being a binary-coded dummy variable, where "0" represented nevi and "1" represented primary melanoma, was created. Table 5 shows how the expression of each indicated gene covaries with the disease variable:

5

Table 5. Disease Covariance

| Symbol | Covariance v. Disease Variable | Symbol | Covariance v. Disease Variable | Symbol | Covariance v. Disease Variable |
|---|---|---|---|---|---|
| B2M | 0.1253 | NR4A1 | 0.766 | SOX4 | 0.1853 |
| B4GALT1 | 0.2552 | PDIA4 | 0.165 | SP100 | 0.203 |
| BRAF | 0.3014 | PDLIM7 | 0.1693 | SQSTM1 | -0.092 |
| CDK2 | 0.146 | PIP4K2A | 0.4079 | STAT2 | 0.0731 |
| CREBBP | 0.086 | PPIA | 0.1539 | TADA3 | 0.184 |
| CTNNB1 | 0.205 | PRAME | 0.3603 | TEX13A | 0.2098 |
| GALNTL1 | 0.1324 | PTEN | 0.164 | TFRC | 0.2609 |
| GNAS | 0.205 | RPL37A | -0.409 | TGFB1 | 0.1114 |
| MAGEA2 | 0.3195 | RUNX1 | 0.2568 | TP53 | 0.265 |
| MAGED2 | 0.129 | SAT1 | 0.4122 | TPSAB1 | 0.188 |
| MCM6 | 0.2381 | SOCS3 | 0.4256 | | |

As shown in Table 5, the mean expression value for each gene in Table 4 is higher (positive value) in primary melanoma than in nevi except for RPL37A and SQSTM1 (negative value) where the means are higher in nevi as compared to primary melanoma. In other words, except as noted, the genes in Table 4 tend to be upregulated in primary melanoma as compared to their expression in nevi.

Using these genes individually or in combinations will yield predictive models (*e.g.*, regression models or, in more specific examples, linear regression models) capable of characterizing (*e.g.*, diagnosing) test samples as benign nevi or primary melanoma. Illustrative, non-limiting gene combinations for use in disclosed methods, arrays or kits are at least 2, 3, 4, 5, 6, 7, 8, or all 9 of MAGEA2, PRAME, PDIA4, NR4A1, PDLIM7, B4GALT1, SAT1, RUNX1, and/or SOCS3.

In addition to overall significance, when selecting model combinations among the set of 32 genes, a number of measures were used to help determine which genes paired or combined well together to form a predictive model. One specific method was to minimize multicollinearity between predictors (*i.e.*, the Xn variables; see below) in the model as measured by the variance inflation factor (VIF) of each Xn variable gene in a model. Any combinations of the genes (e.g., mRNA or miRNA) in Table(s) 4, 11 and/or 13 in which all predictor Xn variables have a variance inflation factor (VIF) less than 10 is likely to have useful predictive value for differentiating between samples from benign nevi versus those from primary melanoma and, accordingly, are contemplated by this disclosure.

D. Logistic Regression Models

The basis used for developing statistical predictive models using the genes in Table 4 was logistic regression with a binary distribution and a logit link function. Estimation for the models was performed using Fischer Scoring. However, models estimated with exact logistic regression, Empirical Sandwich Estimators or other bias corrected, variance stabilized or otherwise corrective estimation techniques will also, under many circumstances, provide similar models which while yielding slightly different parameter estimates will yield qualitatively consistent patterns of results. Similarly, other link functions, including but not limited to a cumulative logit, complementary log-log, probit or cumulative probit may be expected to yield predictive models that give the same qualitative pattern of results.

The primary form of the model (algorithm) in this Example is:

$$\text{Logit}(Yi) = \beta 0 + \beta 1X1 + \beta 2X2 + \beta 3X3...\beta nXn$$

where $\beta o$ is an intercept term, $\beta n$ is a coefficient estimate and Xn is the log base 2 expression value for a given gene. Typically, the value for all $\beta$ will be greater than -1,000 and less than 1,000. Often, the $\beta 0$ intercept term will be greater than -200 and less than 200 with cases in which it is greater than -100 and less than 100. The additional $\beta n$, where $n>0$, will likely be greater than -100 and less than 100.

To validate model performance a number of tests were conducted. A Wald Chi-Square test was used and the test needed to show a statistically significant result for overall model fit. A Hosmer and Lemeshow lack fit test needed to indicate not statistically detectable lack of fit for the model. Predictors for each gene in the model needed to be significant $p<0.05$.

A number of cross validation methods were used to ensure reproducibility of the results. The primary method was a one-step maximum likelihood estimate approximation implemented as part of the SAS Proc Logistic classification table procedure. Ten (10)-fold cross validation and 66-33% split validation was also performed in the open source package Weka for additional confirmation of results. While logistic regression is the mathematical underpinning in this Example, other statistical, mathematical and data mining procedures (such as probit regression, support vector machines or clustering algorithms) can produce models which give the same qualitative pattern of results.

Applying logistic regression modeling to the data in the present Example, the following Table 6 shows non-limiting combinations of genes that accurately differentiate between nevi and primary melanoma samples and the values for the corresponding predictive algorithm.

Table 6. Exemplary Predictive Combinations (Normalized to Four) with Algorithm Values

| Embodiment | Gene Combination (From left to right, each gene represents X1, X2...Xn, as applicable, in the algorithm: Output = β0 + β1X1 + β2X2...βnXn) | Regression Coefficients and Intercept (β0) |
|---|---|---|
| B1 | N4RA1, B4GALT1<br>As an example:<br>Output = -59.0958+1.5998(NR4A1)+4.2115(B4GALT1) | β0 = -59.0958<br>β1 = 1.5998<br>β2 = 4.2115 |
| B2 | NR4A1, SOX4 | β0 = -39.1063<br>β1 = 2.0554<br>β2 = 1.8234 |
| B3 | NR4A1, SOX4, B4GALT1 | β0 = -75.3582<br>β1 = 1.9674<br>β2 = 1.5141<br>β3 = 4.0622 |
| B4 | NR4A1, SOX4, SQSTM1, B2M | β0 = -34.8327<br>β1 = 2.2925<br>β2 = 2.2998<br>β3 = -3.2193<br>β4 = 2.1559 |
| B5 | MAGED2, SAT1, SOX4 | β0 = -49.3358<br>β1 = -0.2919<br>β2 = 3.0513<br>β3 = 2.3171 |
| B6 | N4RA1, BRAF | β0 = -43.4593<br>β1 = 2.1785<br>β2 = 2.3159 |
| B7 | NR4A1, RPL37A | β0 = -9.6524<br>β1 = 3.2965<br>β2 = -2.1656 |
| B8 | NR4A1, SQSTM1, TPSAB1 | β0 = -7.6589<br>β1 = 2.7873<br>β2 = -3.6387<br>β3 = 1.6122 |
| B9 | NR4A1, TFRC, SAT1 | β0 = -43.3177<br>β1 = 1.5862<br>β2 = 0.111<br>β3 = 2.354 |
| B10 | TFRC, SAT1 | β0 = -40.4475<br>β1 = 1.3975<br>β2 = 2.5618 |
| B11 | SOCS3, TFRC, BRAF | β0 = -42.6409<br>β1 = 1.7603<br>β2 = -0.032<br>β3 = 3.048 |

Table 6. Exemplary Predictive Combinations (Normalized to Four) with Algorithm Values

| Embodiment | Gene Combination (From left to right, each gene represents X1, X2...Xn, as applicable, in the algorithm: Output = β0 + β1X1 + β2X2...βnXn) | Regression Coefficients and Intercept (β0) |
|---|---|---|
| B12 | SOCS3, TFRC | β0 = -34.9546<br>β1 = 1.1505<br>β2 = 2.4557 |
| B13 | SOCS3, SOX4, SAT1, BRAF | β0 = -93.1404<br>β1 = 2.3007<br>β2 = 0.9978<br>β3 = 4.3683<br>β4 = 1.8052 |

FIGS. 3-7 show particular results of the model using the combination of N4RA1 and B4GALT1 as described in this Example to accurately determine whether a sample is properly characterized (*e.g.*, diagnosed) as a nevus or a primary melanoma.

The algorithms disclosed in Table 6 were used to characterize test FFPE skin biopsies as primary melanoma or nevi. The algorithms (aka, fitted model) provide a predicted event probability, which, in this Example, was the probably of a sample being a primary melanoma. A SAS computation method known to those of ordinary skill in the art was used to compute a reduced-bias estimate of the predicted probability (see, support.sas.com/documentation/cdl/en/statug/63347/HTML/default/viewer.htm#statug_logistic_sect044.htm (as of June 22, 2012)).

A series of threshold values, z, where z was between 0 and 1 were set. If the predicted probability calculated for a particular sample exceeded or equaled the pre-set threshold value, z, the sample was assigned to the primary melanoma group; otherwise, it was assigned to the nevi group. The respective group assignments were then cross-checked against the known clinical data to determine, among other things, true positives, true negatives, false positives, and false negative. These results are shown, for example, in Classification Tables such as those set forth in FIGS. 5 and 6.

Threshold values can be determined by the ordinarily skilled artisan based on the desired clinical utility of the model. FIGS. 5 and 6 demonstrate this point using a representative NR4A1 and B4GALT1 (normalized to four) model. A higher threshold can be set for making a primary melanoma call; for example, see the region highlighted in Figure 5 (probability level 0.700-0.780). In this case, the false positive rate was relatively low or, stated otherwise, there was a higher specificity than sensitivity. Conversely, the threshold for calling a sample a primary melanoma can

be lowered; for example see the region highlighted in FIG. 6 (probability level 0.240-0.620). At which threshold levels, the model would provide fewer false negatives or a higher sensitivity but a somewhat lower specificity.

It is noted that the "Output" from the model equations provided in Tables 6 and 8 is a logit. The logit is the log of the odds ratio for a sample being a primary melanoma. As an alternative, an ordinarily skilled artisan could use a logit as a threshold value for calling a sample primary melanoma or nevi. Equivalent results will be obtained under either method. For example, choosing a predicted probability of 0.5 will give the same results as using a logit of 0. This can be understood because an odds ratio of 1 is interpreted as there being an equal probability of a sample being primary melanoma or nevi. The logarithm of 1 is equal to zero and thus the logit of zero is equal to an odds ratio of 1. Given that an odds ratio of 1 or a logit of zero is a 50/50 probability of a sample being primary melanoma or nevi, a predicted probability of 0.5 is an equivalent result; and thus using either the predicted probability from a fitted model or a logit output will lead to the same results.

## Example 3

## Analysis with Two Alternative Normalizers Demonstrates the Robustness of Models for Predicting Nevi versus Primary Melanoma

The robustness of the predictive (*e.g.*, diagnostic) gene combinations described in Example 2 was further shown by using an alternate set of normalization genes to normalize the data. Among other things shown in this Example, such analysis had no meaningful impact on the [N4RA1, B4GALT1] predictive model, which outcome is believed to be representative of all predictive models described in Example 2.

RPL37A, RPS6KB2, SDHA, and TFRC were included on arrays described in Examples 1 and 2 as putative "housekeeper" genes. As discussed above, the concept of a "housekeeper" gene (*e.g.*, a gene whose expression is invariant across sample types) is losing favor and so should be (and was) tested in fact.

The composite average log base 2 expression value (see Example 2) for each of these candidate "housekeeper" genes was normalized to such value for each of the other candidate "housekeeper" genes. Coefficients of variation (CV) and standard deviations (SD) for each normalized "housekeeper" were calculated and are as shown in Table 7.

Table 7: Coefficients of variation (CV) and standard deviations for each normalized "housekeeper".

| Candidate "Housekeeper" | CV | SD |
|---|---|---|
| RPL37A | 0.101384 | 0.894 |
| RPS6KB2 | 0.023556 | 0.365 |
| SDHA | 0.022369 | 0.339 |
| TFRC | 0.070521 | 0.719 |

A candidate "housekeeper" was discarded as a legitimate normalizer if its CV was more than two-fold greater than the CV of the lowest CV of the other candidates. Accordingly, only RPS6KB2 and SDHA were selected as normalizers in this Example.

General information regarding this representative set of two normalization genes for human nevi and primary melanoma skin biopsies is described above in Table 3.

The composite average log base 2 expression values for each of the genes in Table 4 were normalized to the composite average log base 2 expression values for RPS6KB2 and SDHA, and the logistic regression analysis described in Example 2 repeated for each of the predictive gene combinations shown in Table 6.

The corresponding intercepts and coefficients for each predictive algorithm where gene (Xn) was normalized to two (*i.e.*, RPS6KB2 and SDHA) is shown in following Table 8.

Table 8. Exemplary Predictive Combinations (Normalized to Two) with Algorithm Values

| Embodiment | Gene Combination (From left to right, each gene represents X1, X2...Xn, as applicable, in the algorithm: Output = β0 + β1X1 + β2X2...βnXn) | Regression Coefficients and Intercept (β0) |
|---|---|---|
| B1 | NR4A1, B4GALT1 | β0 = -39.9861<br>β1 = 1.9964<br>β2 = 2.1807 |
| B2 | NR4A1, SOX4 | β0 = -25.7153<br>β1 = 2.1472<br>β2 = 0.4994 |
| B3 | NR4A1, SOX4, B4GALT1 | β0 = -39.4785<br>β1 = 1.9795<br>β2 = -0.4937<br>β3 = 2.5948 |
| B4 | NR4A1, SOX4, SQSTM1, B2M | β0 = -12.6489<br>β1 = 2.5444<br>β2 = 0.6808<br>β3 = -3.7649<br>β4 = 1.9304 |

Table 8. Exemplary Predictive Combinations (Normalized to Two) with Algorithm Values

| Embodiment | Gene Combination (From left to right, each gene represents X1, X2...Xn, as applicable, in the algorithm: Output = β0 + β1X1 + β2X2...βnXn) | Regression Coefficients and Intercept (β0) |
|---|---|---|
| B5 | MAGED2, SAT1, SOX4 | β0 = -26.6321<br>β1 = -2.5752<br>β2 = 4.2547<br>β3 = 1.3862 |
| B6 | N4RA1, BRAF | β0 = -38.9049<br>β1 = 2.2051<br>β2 = 2.1615 |
| B7 | NR4A1, RPL37A | β0 = -2.9081<br>β1 = 3.0926<br>β2 = -2.4906 |
| B8 | NR4A1, SQSTM1, TPSAB1 | β0 = -13.1057<br>β1 = 3.004<br>β2 = -3.7264<br>β3 = 2.0336 |
| B9 | NR4A1, TFRC, SAT1 | β0 = -62.1051<br>β1 = 1.6769<br>β2 = 2.2591<br>β3 = 2.6097 |
| B10 | TFRC, SAT1 | β0 = -69.1937<br>β1 = 3.8901<br>β2 = 3.6063 |
| B11 | SOCS3, TFRC, BRAF | β0 = -59.3456<br>β1 = 1.4668<br>β2 = 2.4699<br>β3 = 3.1692 |
| B12 | SOCS3, TFRC | β0 = -35.5872<br>β1 = 0.9893<br>β2 = 3.0369 |
| B13 | SOCS3, SOX4, SAT1, BRAF | β0 = -108.5<br>β1 = 4.1954<br>β2 = -7.3183<br>β3 = 6.3842<br>β4 = 8.8727 |

The algorithms disclosed in Table 8 were determined and tested as described in Example 2.

FIGS. 8 and 9 show (i) the overall B4GALT1 and NR4A1 (2-normalizer) model fit remains highly significant; (ii) the AUC is 97.67 (indicating, *e.g.*, a 97.67% probability that B4GALT1 and NR4A1 (2-normalizer) predictive model will rank a randomly chosen positive instance higher than a randomly chosen negative instance); (iii) 90% plus correct classification over ~50% of possible thresholds after SAS cross validation; (iv) attenuated specificity across the range of possible

thresholds after SAS cross validation; and (v) moderate increase in maximum sensitivity after SAS cross validation as compared to the Example 2 model for this same molecular signature.

In summary, this Example demonstrates that the B4GALT1 and NR4A1 model for characterizing samples as nevi or primary melanoma on a molecular level is repeatable across normalization methods. A similar outcome is expected for the other predictive signatures described in Examples 2 and 3 (and elsewhere herein) using analogous computations. The robustness and utility of these representative and other disclosed models for characterizing a test sample as a nevus or primary melanoma, thus, are clearly demonstrated.

## Example 4

### Classification of Melanoma and Nevi Using Machine Learning Methods

In this Example, mRNA and miRNA expression in a third set of FFPE skin samples biopsied from human subjects was determined using a set of four ArrayPlates, including ArrayPlates Nos. 3 and 4 (see Examples 2 and 3). Such data was used to successfully identify sets of mRNAs and miRNAs significantly differentially expressed in melanoma and nevi samples and to train machine learning (*e.g.*, Random Forest (Breiman, *Machine Learning*, 45 (1): 5–32 (2001)) melanoma-nevi classifiers.

A set of 115 FFPE tissue sections, each approximately 5 um thick and mounted on a microscope slide, was provided by the John Wayne Cancer Institute (JWCI) tissue bank. The set included 56 nevi samples and 59 primary melanoma samples.

The samples were analyzed for expression of 181 (including controls) mRNAs or miRNAs on a set of four ArrayPlates. Assay and detection methods for ArrayPlates Nos. 3-5 (mRNA only) were substantially as described in Example 1. Assay and detection methods for ArrayPlate No. 6 (miRNA and mRNA codetection) were substantially as described in PCT Publication No. WO2013/049231. NPPs for ArrayPlate No. 6 normalizer and negative control (ANT) mRNAs were 25mers corresponding to the 3'-most 25 nucleotides of the respective control NPPs described elsewhere in these Examples. The set of ArrayPlates included ArrayPlates Nos. 3 and 4 (see Table 1), "ArrayPlate No. 5" specific for mRNA targets, and "ArrayPlate No. 6" specific for miRNA targets (plus mRNA controls). The target listings for ArrayPlates Nos. 5 and 6 are shown in Table 9. There were some common mRNA targets on ArrayPlate Nos. 3-6; thus, data was gathered for 101 different mRNAs (including controls) and 42 different miRNAs.

Table 9. ArrayPlate Target List

| ArrayPlate No. 5 | |
|---|---|
| **Identifier** | **GenBank Accesssion No.** |
| MAGEA2 | NM_005361 |
| PAX3 | NM_181457 |
| GALNTL1 | NM_001168368 |
| MAGEA1 | NM_004988 |
| PanMAGEA3-12 | |
| BIRC7 | NM_139317 |
| BIRC5 | NM_001168 |
| XIAP | NM_001167 (v1); NM_001204401 (v2); NR_037916.1 (noncoding v3) |
| PRAME | NM_206953 |
| MET | NM_001127500 |
| MFI2 | NM_005929 |
| MCAM | NM_006500 (GI:71274106) |
| BAD | NM_004322 (v1); NM_032989 (v2) |
| BCL2 | NM_000633 (alpha); NM_000657 (beta) |
| HIF1A | NM_001530 |
| MIB1 | NM_020774 |
| TOP2A | NM_001067 |
| WT1 | *e.g.*, NM_000378 (variant A) or also other variants (*e.g.*, B-F) |
| MCM2 | NM_004526 |
| MCM6 | NM_005915 |
| ALK | NM_004304 |
| S100B | NM_006272 |
| PDIA4 | NM_004911 |
| SOX4 | NM_003107 |
| XRCC5 | NM_021141 |
| DAZAP2 | NM_014764 |
| EVI2B | NM_006495 |
| LDHA | NM_005566 |
| BRAF | NM_004333 |
| ERCC1 | NM_001983 |
| ESR1 | NM_000125 |
| RPL19 | NM_000981 |

| ArrayPlate No. 6 | |
|---|---|
| **Identifier** | **GenBank or miRBase* Identifier (as applicable)** |
| SDHA | NM_004168 |
| RPS6KB2 | NM_003952 |
| RPL37A | NM_000998 |
| TFRC | NM_003234 |
| ANT | NM_119937 |
| 23b | hsa-miR-23b |
| 211 | hsa-miR-211 |
| 1224-3p | hsa-miR-1224-3p |
| 193A-5P | hsa-miR-193a-5p |
| 146A | hsa-miR-146a |
| 513b | hsa-miR-513b |
| 133A | hsa-miR-133a |
| 182 | hsa-miR-182 |
| 205 | hsa-miR-205 |
| 665 | hsa-miR-665 |
| 1254 | hsa-miR-1254 |
| 200C | hsa-miR-200c |
| 292 | mmu-miR-292-3p |
| 200A | hsa-miR-200a |
| 21 | hsa-miR-21 |
| 140-3p | has-miR-140-3p |
| 140-5p | has-miR-140-5p |
| 29C | hsa-miR-29c |
| 142-5P | hsa-miR-142-5p |
| 595 | hsa-miR-595 |
| 207 | mmu-miR-207 |
| 106a | hsa-miR-106a |
| 122 | hsa-miR-122 |
| 1304 | hsa-miR-1304 |
| 155 | hsa-miR-155 |
| 191 | hsa-miR-191 |
| 375 | hsa-miR-375 |

Table 9. ArrayPlate Target List

| ArrayPlate No. 5 | |
|---|---|
| **Identifier** | **GenBank Accesssion No.** |
| SDHA | NM_004168 |
| ALDOA | NM_000034 |
| RPLP0 | NM_001002; NM_053275 |
| PPIA | NM_021130 |
| ANT | See other Examples |
| MAGEB1 | NM_002363 (v1); NM_177404 (v2); NM_177415 (v3); |
| MAGEC2 | NM_016249 |
| MAGED2 | NM_014599 |

| ArrayPlate No. 6 | |
|---|---|
| **Identifier** | **GenBank or miRBase* Identifier (as applicable)** |
| 612 | hsa-miR-612 |
| 650 | hsa-miR-650 |
| 1180 | hsa-miR-1180 |
| 183 | hsa-miR-183 |
| 203 | hsa-miR-203 |
| 1293 | hsa-miR-1293 |
| 342-3p | hsa-miR-342-3p |
| 1294 | hsa-miR-1294 |
| 19b | hsa-miR-19b |
| 557 | hsa-miR-557 |
| 1198-5p | mmu-miR-1198-5p |
| let-7a | hsa-let-7a |
| 1291 | hsa-miR-1291 |
| 29b | hsa-miR-29b |
| 150 | hsa-miR-150 |

* Kozomara and Griffiths-Jones, Nuc. Acids Res., 39(Database Issue):D152 (2011)

Normalizing genes were SDHA and RPS6KB2 on ArrayPlate Nos. 3, 4 and 6, and SDHA, RPL19, RPLP0 and ALDOA on ArrayPlate No. 5 (see, also, Table 3). None of these normalizers showed any significant difference across sample types as described elsewhere in these Examples.

Due to limited sample availability, not all samples were run on each array and all raw data was subject to rigorous quality control (*i.e.*, pre-processing), as follows: Raw data was background subtracted and log 2 transformed. Any samples for which greater than 200 RLU was measured for the negative control gene, ANT, were deemed to have failed, and all data from those particular wells were removed from further consideration. A coefficient of variance (CV) was determined for replicate expression values for each gene. If the CV for sample replicates exceeded 8%, the replicate farthest from the average was removed as an outlier. Replicate reproducibility was measured by pairwise correlation and by pairwise simple linear regression. If the correlation had r >= 0.90 and the intercept of the linear regression was not statistically significantly different from zero, such replicate was accepted; otherwise, it was deemed failed. Any sample with more than two failed replicates was defined as a failed sample. Data failing to meet quality standards were removed from the analysis. A summary is provided in Table 10:

Table 10: Summary of Samples and Genes Analyzed

| ArrayPlate No. | # Targets | Sample Class | # Samples Run | # Samples after Data QC |
|---|---|---|---|---|
| 3 | 47 | Melanoma | 59 | 57 |
| | | Nevus | 56 | 54 |
| 4 | 47 | Melanoma | 59 | 53 |
| | | Nevus | 40 | 34 |
| 5 | 40 | Melanoma | 58 | 46 |
| | | Nevus | 35 | 32 |
| 6 | 47 | Melanoma | 59 | 59 |
| | | Nevus | 50 | 50 |

*Univariate Analysis*

Several univariant analyses of the processed data (*e.g.*, log-fold change, two sample t-test (False Discovery Rate (FDR) adjusted p-value), and AUC logistic regression analysis) were performed to evaluate whether a particular gene was significantly differentially expressed between sample types in each data set.

The results of univariate analyses for the three mRNA arrays (*i.e.*, ArrayPlates 3-5) are shown in FIG. 10A. The values for each of three tests performed ((i) Area under the Receiver Operating Characteristic (ROC) curve (AUC), (ii) log-fold change (fch), and (iii) two sample t-test (FDR adjusted p-value; FDR.pvalue)) are shown. Genes for which expression data was gathered are shown on the x-axis, and the value of the respective univariate statistic is shown on the y-axis. For the AUC analysis, a higher value is desirable. In this case, 0.75 (at dotted line) was assigned as the cut off of statistical significance. Genes with AUCs above that line are candiates for distinguishing nevi from melanoma. For the log-fold change analysis, negative 1 and positive 1 (each of which equals a two-fold difference in expression between nevi and melanoma) were assigned as the cut offs for statistical significance. Genes with log-fold change greater than positive 1 and less than negative 1 are candiates for distinguishing nevi from melanoma. For FDR adjusted p-values, a lower value is desireable, and 0.05 (at dotted line) was assigned as the cut off of statistical significance. Genes with expression below that line are candidates for distinguishing nevi from melanoma.

The results of univariate analysis for the miRNA array are shown in FIG. 10B. As above, AUC under a ROC curve, log-fold change and two sample t-test (FDR adjusted p-value) were determined for each of the miRNAs listed on the x-axis. The value of the respective univariate statistic is on the y-axis. The data labeled "HK" show miRNA expression data normalized to the "housekeepers" on ArrayPlate No. 6 and "NO.HK" show unnormalized miRNA expression data.

Normalized and unnormalized data generally provide similar results. Statistical cut offs for AUC, fold-change and two sample t-test (FDR adjusted p-value) were the same as for the mRNA analysis above. miRNAs with AUC greater than 0.75 , fold-change greater than positive 1, and FDR adjusted p-value below the cut off line are candidates for distinguishing nevi from melanoma.

The positive outcome of the univariate analyses (*i.e.*, identification of mRNA and miRNA significantly differentially expressed between melanoma and nevi) supported the decision to proceed with more resource-consuming multivariate analyses and further melanoma-nevi classifier development.

*Multivariate Analysis*

A multivariant analysis then was performed to determine which subsets of the detected targets most powerfully (from a stastical perspective) distinguished between melanoma and nevi sample types. Multiple feature selection methods (RF, LIMMA, t-test, AUC) were used to evaluate whether a particular gene was significantly differentially expressed between sample types in each data set. Machine learning algorithms (*e.g.*, Logistic Regression (LR), Random Forest (RF), Support Vector Machine (SVM), K-nearest neighbor (KNN)) were used to develop an initial classifier. Both feature selection and classification performance were evaluated in a leave one out cross-validated fashion. Error rate as a function of gene number and Receiver Operating Characteristic (ROC) curve were used to evaluate the performance of the classifier.

For the genes detected in ArrayPlate No. 3, FIG. 11 shows the AUC performance of classifiers based on the top 2 (GN=2) through the top 40 (GN=40) genes on that array. For this type of analysis, AUC increases with higher sensitivity (*i.e.*, true positive rate shown on the y-axis) and lower false positive rate (*i.e.*, "1-Specificity" shown on the x-axis) of the tested classifier. This figure demonstrates that AUC exceeded 0.93 (1.00 is "perfect") with all ArrayPlate No. 3 classifiers greater than 12 genes. One of ordinary skill in the art will appreciate that this result does not mean one could not select a classifier with fewer than 12 genes based on the information disclosed herein; however, such classifier may not have as high sensitivity and specificity. In some settings, high sensitivity or high specificity may not be the greatest priority and classifiers may be accordingly selected. For example, it may be considered worse outcome for a melanoma-nevi classifier to misidentify a melanoma as a nevus rather than to misidentify a nevus as a melanoma; in that case, a classifier may be selected to minimize false negatives while being a bit more lax on false positives (when null hypothesis = melanoma or not).

For the genes in each of ArrayPlate Nos. 3-6, FIG. 12 shows the error rate of classifiers determined by the various statistical methods (*i.e.*, AUC, t-test, Random Forest, LIMMA) as a

function of the number of genes in the classifier. This figure indicates that the misclassification error of a melanoma-nevi classifier is minimized when such a classifier has about 10 or more genes. As above, this result does not mean one could not select a classifier with fewer than about 10 genes based on the information disclosed herein; however, such classifier is likely to be more error prone.

For each of ArrayPlate Nos. 3-5, the genes with the highest occurrence frequency in leave-one-out cross validation (LOOCV) of Random Forest algorithms and the best performance as measured by AUC were selected and consolidated into the gene list shown in Table 11. A similar approach was used to select miRNAs shown in Table 13, and exemplary combinations of genes shown in Table 14.

Based on the above analysis, the mRNAs and miRNAs shown in Tables 11, 13 and 14, as applicable, were selected as useful (in combinations of at least 2, 3, 4, 5, 6, 7, 8, or, as applicable, 9, 10, 11, 12, 13, 14, 15 or more) to accurately classify a test sample as a nevus or melanoma. In particular examples, such classifier utilizes a machine learning (*e.g.*, Random Forest or support vector machine) algorithm. Representative nuclease protection probes used to detect the respective expression product in this Example are also shown in Tables 11 and 13. In some examples, the expression levels of these genes are normalized to one or more housekeepers, such as SDHA, RPS6KB2, RPL37A, and/or TFRC (such as, SDHA and RPS6KB2).

Table 11. Genes (mRNAs) For Nevus-Melanoma Classification

| Symbol | Representative NPP (5' to 3') | SEQ ID NO. |
|---|---|---|
| B4GALT1 | GTCTTGGAACCTGAGCCCAGGCTGGACCTGGCA AAGGCGCTCAGTGGTAG | 125 |
| BAX | CGATGCGCTTGAGACACTCGCTCAGCTTCTTGG TGGACGCATCCTGAGGC | 126 |
| MAGEA2 | CTCAGGCTCTCCACCTGGATGCTTGGCAGATCC TAGAACCACTGCATCTG | 127 |
| NR4A1 | CGCCACAGCTGCCACGTGCTCCTTCAGGCAGCT GGCGATGCGGTTCTGCA | 128 |
| PDIA4 | CACATCAAACCTGCTGGCCAGCACAGACGCTGA GGTTGCATCGATCTTGG | 129 |
| PRAME | GTCTGGCTGTGTCTCCCGTCAAAGGCTGCCATG AAGAGTGGCGGGAAGAG | 130 |
| RUNX1 | GCAGAGTCACACACATGCAAACACGCACTCTTC GGAAGGCAGCCACTGTC | 131 |
| SOCS3 | GTCTTCTCTACCAGGAGCCTGAGGTGAAAGATG TCCCGTCTCCTCCATCC | 132 |
| SAT1 | ATTTCAAACATGCAACAACGCCACTGGTAATAA AGCTTTGGAATGGGTGC | 133 |
| PDLIM7 | CTTCGATGTGTGTGAGGCTACCCGCATTCTCGC CATCGATGCTCAGCACC | 134 |

Table 11. Genes (mRNAs) For Nevus-Melanoma Classification

| Symbol | Representative NPP (5' to 3') | SEQ ID NO. |
|---|---|---|
| BIRC5 | GCACAGGCTCACAGAAGCCGAGATCCACATCA CCGCCTGGCATGCAAAGG | 135 |
| HIF1A | GGCCATTTCTGTGTGTAAGCATTTCTCTCATTTC CTCATGGTCACATGGA | 136 |
| MET | CAAAGAAGTTGATGAACCGGTCCTTTACAGATG AAAGGACTTTGGCTCCC | 137 |
| MAGEC2 | GGACTACTGGGAATGCTCTCGGTAAGATTTGGT ATCACACCAGAGGGCAC | 138 |
| ERCC1 | AGTGGGAAGGCTCTGTGTAGATCGGAATAAGG GCTTGGCCACTCCAGGAG | 139 |
| POLR2J3 | GAGGTTGCAGTGAGCCAAGATCGCGCCAGCCTG GCGACAGAGTGAGACTC | 140 |
| LDHA | TCCATCATCTCTCCCTTCAATTTGTCTTCGATGA CATCAACAAGAGCAAG | 141 |
| PICALM | GACAGGCTGGCTGTATATTAAGGTTGGTTGCGT CATTACAGGAACACTTC | 142 |
| ZFYVE16 | GAAGTTCGCTGTGAGGAAGCCAACTCTGAAGA AACTGAGCAGTGGTTAGA | 143 |
| BEST1 | GTTTCTCCAACTGCTTGTGTTCTGCCGGAGTCAT AAAGCCTGCTTGCACC | 144 |

Additional detail on several of the above-listed genes may be found in Table 4; similar detail for the remaining genes is provided in Table 12 below:

Table 12. Supplement to Table 4

| Symbol | Name | GenBank Ref. No(s). |
|---|---|---|
| BAX | BCL2-Associated X Protein | NM_138761 (GI: 163659848) (alpha); NM_004324.3 (GI: 34335114) (beta); NM_138763 (GI: 163659849) (delta); NM_138764 (GI: 242117892) (sigma); NR_027882 (GI: 242117894) (epsilon, non-coding) |
| BIRC5 | Baculoviral IAP Repeat-Containing Protein 5 (aka, Apoptosis Inhibitor 4; API4 Survivin) | NM_001168 (GI: 59859877) (v1); NM_001012270 (GI: 59859879) (v2); NM_001012271 (GI: 59859881) (v3) |
| HIF1A | Hypoxia-Inducible Factor 1, Alpha Subunit (aka, Hif1-Alpha; Member of PAS Superfamily 1; MOP1) | NM_001530 (GI: 194473733) (v1); NM_181054 (GI: 194473734) (v2); NM_001243084 (GI: 340545530) (v3) |
| MET | MET PROTOONCOGENE (aka, Hepatocyte Growth Factor Receptor; HGFR) | NM_001127500 (GI: 188595715) (v1); NM_000245 (GI: 42741654) (v2) |

Table 12. Supplement to Table 4

| Symbol | Name | GenBank Ref. No(s). |
|---|---|---|
| MAGEC2 | Melanoma Antigen, Family C, 2 (aka, Cancer-Testis Antigen 10; CT10; HCA587; Melanoma Antigen, Family E, 1; MAGEE1 | NM_016249 (GI:262050676) |
| ERCC1 | Excision-Repair, Complementing Defective, in Chinese Hamster, 1 (aka, Complementation of DNA Repair Defect UV-20 of Chinese Hamster Ovary Cells; UV20 | NM_202001 (GI: 260593723) (v1); NM_001983 (GI: 260593722) (v2); NM_001166049 (GI: 260593724) (v3) |
| POLR2J3 | Homo sapiens polymerase (RNA) II (DNA directed) polypeptide J3 (aka, POLR2J2, RPB11b1, RPB11b2) | NM_001097615 (GI:332634983) |
| LDHA | Lactate Dehydrogenase A (aka, LDH, Subunit M) | NM_005566 (GI: 207028465) (v1); NM_001135239 (GI: 207028493) (v2); NM_001165414 (GI: 260099722) (v3); NM_001165415 (GI: 260099724) (v4); NM_001165416 (GI: 260099726) (v5); NR_028500 (GI: 260099728) (v6, noncoding) |
| PICALM | Phosphatidylinositol-Binding Clathrin Assembly Protein (aka, Clathrin Assembly Lymphoid-Myeloid Leukemia Gene; CALM; CLTH; LAP, Homolog of Drosophila LAP | NM_007166 (GI: 332688229) (v1); NM_001008660 (GI: 332688228) (v2); NM_001206946 (GI: 332688230) (v3); NM_001206947 (GI: 332635086) (v4) |
| ZFYVE16 | Zinc Finger FYVE Domain-Containing Protein 16 (aka, Endosome-Associated FYVE Domain Protein; ENDOFIN; KIAA0305) | NM_014733 (GI: 157426863) (v1); NM_001105251 (GI: 157426865) |
| BEST1 | Bestrophin 1 (aka, VMD2 Gene, TU15B) | NM_004183 (GI: 212720874) (v1); NM_001139443 (GI: 212720888) (v2) |

Table 13. miRNAs For Nevus-Melanoma Classification

| Identifier | Representative NPP (5' to 3') | SEQ ID NO | miRBase Accession No. |
|---|---|---|---|
| hsa.miR.122 | CAAACACCATTGTCACACTCCA | 145 | MI0000442 |
| hsa.miR.1291 | ACTGCTGGTCTTCAGTCAGGGCCA | 146 | MI0006353 |
| hsa.miR.191 | CAGCTGCTTTTGGGATTCCGTTG | 147 | MI0004941 |

Table 13. miRNAs For Nevus-Melanoma Classification

| **Identifier** | **Representative NPP (5' to 3')** | **SEQ ID NO** | **miRBase Accession No.** |
|---|---|---|---|
| hsa.miR.19b | TCAGTTTTGCATGGATTTGCACA | 148 | MI0000074 |
| hsa.miR.200a | ACATCGTTACCAGACAGTGTTA | 149 | MI0000737 |
| hsa.miR.200c | TCCATCATTACCCGGCAGTATTA | 150 | MI0000650 |
| hsa.miR.203 | CTAGTGGTCCTAAACATTTCAC | 151 | MI0000283 |
| hsa.miR.205 | CAGACTCCGGTGGAATGAAGGA | 152 | MI0000285 |
| hsa.miR.21 | TCAACATCAGTCTGATAAGCTA | 153 | MI0000077 |
| hsa.miR.23b | GGTAATCCCTGGCAATGTGAT | 154 | MI0000439 |
| hsa.miR.29c | TAACCGATTTCAAATGGTGCTA | 155 | MI0000735 |
| hsa.miR.342.3p | ACGGGTGCGATTTCTGTGTGAGA | 156 | MI0000805 |
| hsa.miR.375 | TCACGCGAGCCGAACGAACAAA | 157 | MI0000783 |
| hsa.miR.665 | AGGGGCCTCAGCCTCCTGGT | 158 | MI0005563 |
| hsa.miR.1304 | CACATCTCACTGTAGCCTCAAA | 159 | MI0006371 |
| hsa.miR.142.5p | AGTAGTGCTTTCTACTTTATG | 160 | MI0000458 |
| hsa.miR.1254 | ACTGCAGGCTCCAGCTTCCAGGCT | 161 | MI0006388 |
| hsa.let.7a | AACTATACAACCTACTACCTCA | 162 | MI0000060 |
| hsa.miR.140.5p | CTACCATAGGGTAAAACCACTG | 163 | MI0000456 |
| hsa.miR.183 | AGTGAATTCTACCAGTGCCATA | 164 | MI0000273 |

Table 14. Exemplary Gene Combinations

| Embodiment | Gene Combination |
|---|---|
| C1 | B4GALT1, BAX, MAGEA2, NR4A1, PDIA4, PRAME, RUNX1, SOCS3, SAT1, PDLIM7, BIRC5, MET, MAGEC2, POLR2J3, ZFYVE16, BEST1 |
| C2 | NR4A1, SOCS3, PRAME, POLR2J3, BEST1, RUNX1, BIRC5, MET, PDLIM7, ZFYVE16, HIF1A, PICALM |
| C3* | MAGEA2, PRAME, PDIA4, NR4A1, PDLIM7, B4GALT1, SAT1, RUNX1, SOCS3 |
| C4 | hsa.miRNA.342.3p, hsa.miRNA.191, hsa.miRNA.29c, hsa.miRNA.183, hsa.miRNA.182, hsa.miRNA.19b, hsa.miRNA.23b, hsa.miRNA.205, hsa.miRNA.122, hsa.miRNA.200a, hsa.miRNA.200c, hsa.miRNA.203 |

* Combination found in each of Table 4 and Table 11

In summary, this Example demonstrates the utility of specified mRNA and miRNA, for example, as used in machine learning (*e.g.*, Random Forest or support vector machine) models, to characterizing samples as nevi or melanoma (*e.g.*, primary melanoma).

In view of the many possible embodiments to which the principles of the disclosure may be applied, it should be recognized that the illustrated embodiments are only examples of the disclosure and should not be taken as limiting the scope of the invention. Rather, the scope of the disclosure is defined by the following claims. We therefore claim as our invention all that comes within the scope and spirit of these claims.

We claim:

1. A method of characterizing a melanocyte-containing sample, comprising:

determining an expression level for

(i) biomarkers MAGEA2, PRAME, PDIA4, NR4A1, PDLIM7, B4GALT1, SAT1, RUNX1, and SOCS3, and

(ii) at least one normalization biomarker(s), in a melanocyte-containing sample obtained from a subject, thereby generating raw expression values for each of the biomarkers and the at least one normalization biomarker(s);

normalizing the raw expression values for each of the biomarkers to the raw expression values for the at least one normalization biomarker(s) to generate normalized expression values for each of the biomarkers;

using the normalized expression values in a regression or machine learning algorithm to generate an output value;

comparing the output value to a cut-off value, wherein the cut-off value was derived from normalized expression values for the biomarkers in a plurality of melanocyte-containing samples known in advance to be benign or malignant; and

characterizing the sample as benign if the output value is on the same side of the cut-off value as the plurality of known benign samples or characterizing the sample as malignant if the output value is on the same side of the cut-off value as the plurality of known malignant samples.

2. The method of claim 1, wherein the at least one normalization biomarker(s) has no statistically significant difference in expression between nevi and primary melanoma samples.

3. The method of claim 1 or 2, wherein the at least one normalization biomarker comprises 1, 2, 3, 4, 5, 6, 7, 8, or all 9 of biomarkers BMP-1, MFI2, NCOR2, RAP2b, RPS6KB2, SDHA, RPL19, RPLP0, and ALDOA.

4. The method of claim 1 or 2, further comprising:

measuring gene expression values for a plurality of biomarkers in the melanocyte-containing sample, wherein the range of expression for the plurality of

biomarkers is representative of the full range of biomarker expression in the sample transcriptome;

calculating a central tendency expression value for such plurality of biomarkers; and

using the central tendency expression value to normalize the raw expression values for each of the biomarkers.

5. The method of claim 4, further comprising removing outliers and calculating a recalculated plurality central tendency expression value without the outlier expression values, and using the recalculated plurality central tendency expression value to normalize the raw expression values for each of the biomarkers.

6. A method of determining gene expression in a melanocyte-containing sample, comprising:

determining in the sample the expression levels of a plurality of genes comprising biomarkers MAGEA2, PRAME, PDIA4, NR4A1, PDLIM7, B4GALT1, SAT1, RUNX1, and SOCS3; and

providing a report of the plurality of genes expression levels in the sample or a characterization of the sample as a nevus or melanoma based on the expression levels of the plurality of genes.

7. The method of any one of claims 1 to 6, wherein the biomarkers further comprise at least two of biomarkers B2M, BRAF, CDK2, CREBBP, CTNNB1, GALNTL1, GNAS, MAGED2, MCM6, PIP4K2A, PPIA, PTEN, RPL37A, SOX4, SP100, SQSTM1, STAT2, TADA3, TEX13A, TFRC, TGFB1, TP53, and TPSAB1, and wherein the output value was generated by a logistic regression algorithm.

8. The method of any one of claims 1 to 7, wherein the algorithm is

Output Value = $\beta 0 + \beta 1X1 + \beta 2X2 + \ldots \beta nXn$

wherein Xn are log expression value for the biomarkers MAGEA2, PRAME, PDIA4, NR4A1, PDLIM7, B4GALT1, SAT1, RUNX1, and SOCS3, wherein $\beta 0$ is greater than -200 and less than 200, wherein all $\beta$ for $n>0$ are greater than -1,000 and less than 1,000.

9. The method of any one of claims 1 to 6, wherein the biomarkers further comprise

(a) at least two of biomarkers BAX, BIRC5, HIF1A, MET, MAGEC2, ERCC1, POLR2J3, LDHA, PICALM, ZFYVE16, BEST1, and any miRNA that hybridizes to at least one of the sequences of SEQ ID NOs: 145–164; or

(b) at least two of biomarkers BAX, BIRC5, HIF1A, MET, MAGEC2, ERCC1, POLR2J3, LDHA, PICALM, ZFYVE16, and BEST1, or at least two miRNAs each of which hybridize to one of SEQ ID NO: 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163 or 164;

and wherein the output value was generated by a machine learning algorithm.

10. The method of any one of claims 1 to 9, wherein the biomarkers further comprise:

at least two of BAX, MAGEA2, BIRC5, MET, MAGEC2, POLR2J3, ZFYVE16, and BEST1 or

at least two of, POLR2J3, BEST1, BIRC5, MET, PDLIM7, ZFYVE16, HIF1A, and PICALM, and

wherein the output value was generated by a machine learning algorithm.

11. The method of any one of claims 1 to 10, wherein the biomarkers further comprise:

at least 50%, at least 75%, at least 80%, at least 90%, at least 95%, or at least 98% of genes B2M, BRAF, CDK2, CREBBP, CTNNB1, GALNTL1, GNAS, MAGED2, MCM6, PIP4K2A, PPIA, PTEN, RPL37A, SOX4, SP100, SQSTM1, STAT2, TADA3, TEX13A, TFRC, TGFB1, TP53, and TPSAB1;

at least 50%, at least 75%, at least 80%, at least 90%, at least 95%, or at least 98% of genes BAX, BIRC5, HIF1A, MET, MAGEC2, ERCC1, POLR2J3, LDHA, PICALM, ZFYVE16, and BEST1;

miRNAs that hybridize to at least 50%, at least 75%, at least 80%, at least 90%, at least 95%, or at least 98% of SEQ ID NO: 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163 or 164; or

at least 50%, at least 75%, at least 80%, at least 90%, at least 95%, or at least 98% of genes B2M, BRAF, CDK2, CREBBP, CTNNB1, GALNTL1, GNAS, MAGED2, MCM6, PIP4K2A, PPIA, PTEN, RPL37A, SOX4, SP100, SQSTM1, STAT2, TADA3, TEX13A, TFRC, TGFB1, TP53, TPSAB1, BAX, BIRC5, HIF1A, MET, MAGEC2, ERCC1, POLR2J3, LDHA, PICALM, ZFYVE16, and BEST1.

12. A method of determining malignancy in a melanocyte-containing sample, comprising:

determining, in a melanocyte-containing sample obtained from a subject, an expression level of biomarkers B4GALT1, BAX, MAGEA2, NR4A1, PDIA4, PRAME, RUNX1, SOCS3, SAT1, PDLIM7, BIRC5, MET, MAGEC2, POLR2J3, ZFYVE16, and BEST1;

calculating an output from an algorithm that uses the expression levels of the biomarkers as an input; and

determining from the algorithm output that the sample is or is not malignant by comparing the output to a reference standard from known malignant melanocyte-containing samples.

13. The method of claim 12, further comprising normalizing the expression levels of the selected biomarkers to the expression level of at least one normalization biomarker selected from the group consisting of:

(a) at least one of MFI2, RAP2B, BMP1, and/or NCOR2;

(b) MFI2, NCOR2, RAP2b, and BMP-1;

(c) RPS6KB2 and/or SDHA; or

(d) at least one gene expressed in the melanocyte-containing sample that is not the biomarkers, and the expression of which does not significantly differ in a representative plurality of melanocyte-containing samples.

14. The method of any one of claims 1 to 13, wherein determining an expression level comprises determining nucleic acid expression.

15. The method of claim 14, wherein determining nucleic acid expression comprises contacting the sample with a plurality of nucleic acid probes or paired amplification primers, wherein each probe or paired primers is/are specific and complementary to one of the least two biomarkers, under conditions that permit the plurality of nucleic acid probes or paired primers to hybridize to its/their complementary biomarkers.

16. The method of claim 15, further comprising, after contacting the sample with the plurality of nucleic acid probes, contacting the sample with a nuclease that digests single-stranded nucleic acid molecules.

FIG. 1

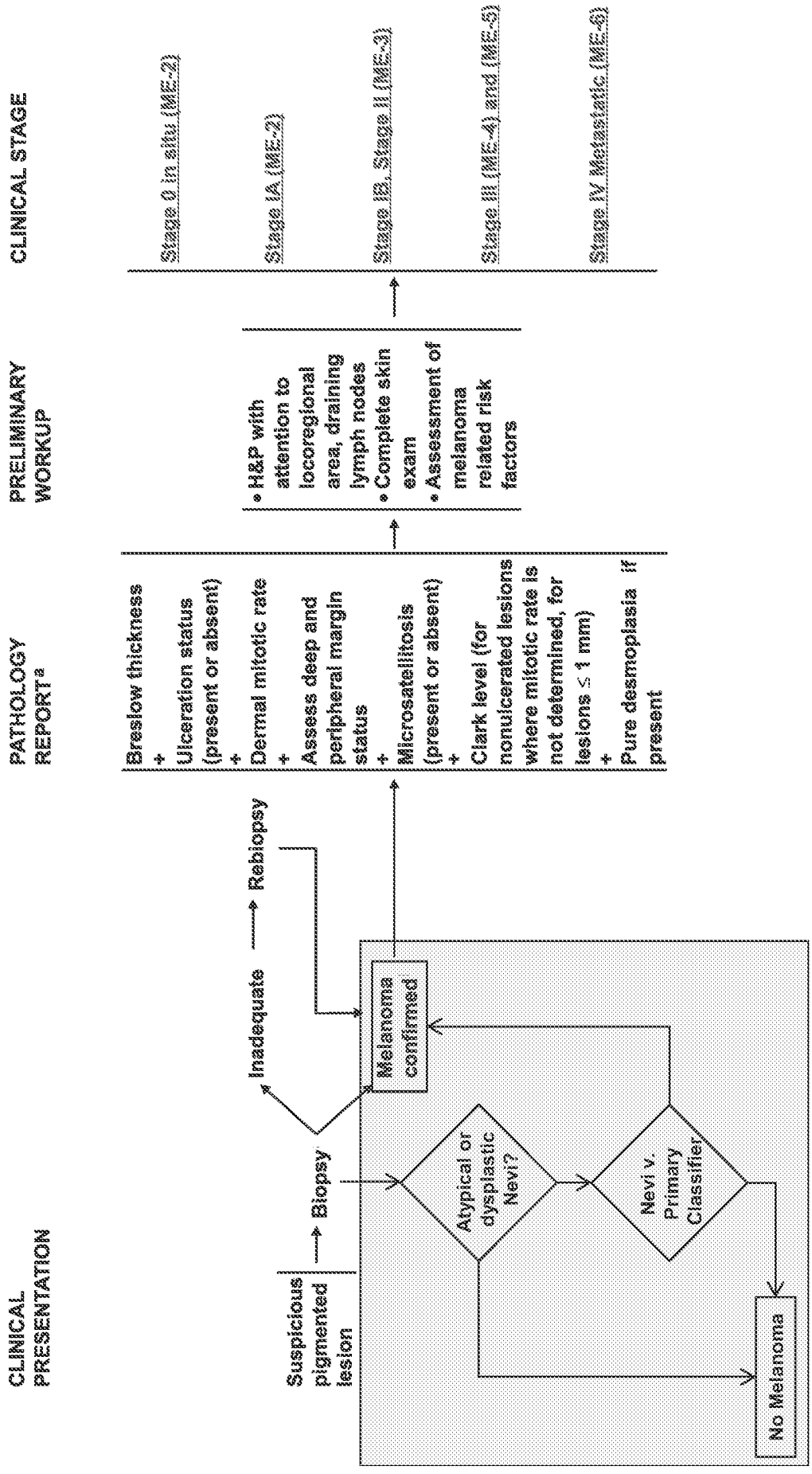
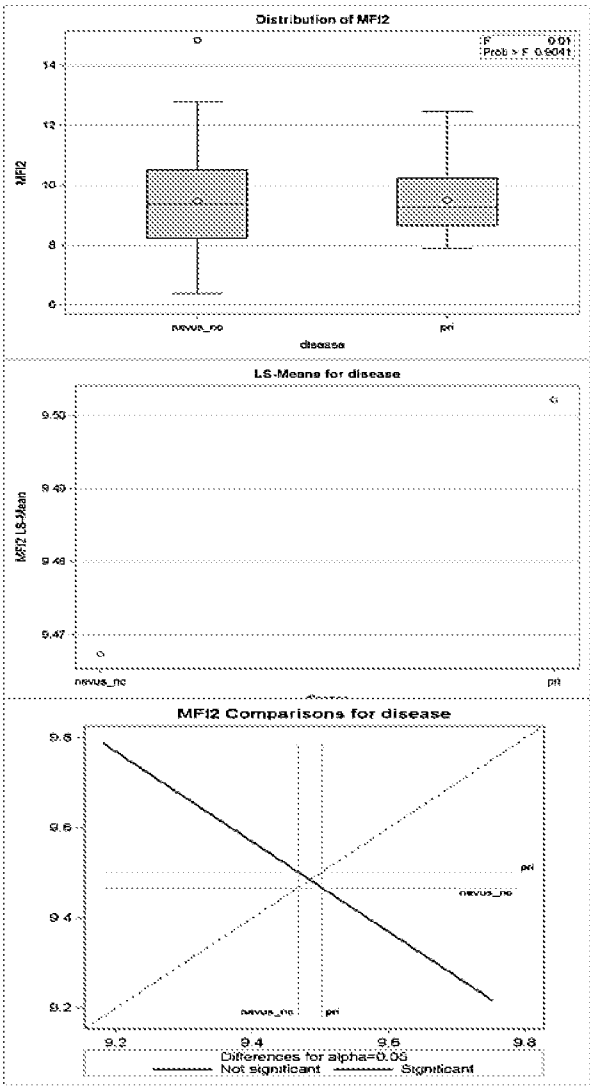

CLINICAL PRESENTATION
PATHOLOGY REPORT[a]
PRELIMINARY WORKUP
CLINICAL STAGE
Suspicious pigmented lesion
Biopsy
Inadequate
Rebiopsy
Atypical or dysplastic Nevi?
Melanoma confirmed
Nevi v. Primary Classifier
No Melanoma
Breslow thickness + Ulceration status (present or absent) + Dermal mitotic rate + Assess deep and peripheral margin status + Microsatellitosis (present or absent) + Clark level (for nonulcerated lesions where mitotic rate is not determined, for lesions ≤ 1 mm) + Pure desmoplasia if present
• H&P with attention to locoregional area, draining lymph nodes
• Complete skin exam
• Assessment of melanoma related risk factors
Stage 0 in situ (ME-2)
Stage IA (ME-2)
Stage IB, Stage II (ME-3)
Stage III (ME-4) and (ME-5)
Stage IV Metastatic (ME-6)

FIG. 2A

MFI2




RAP2B

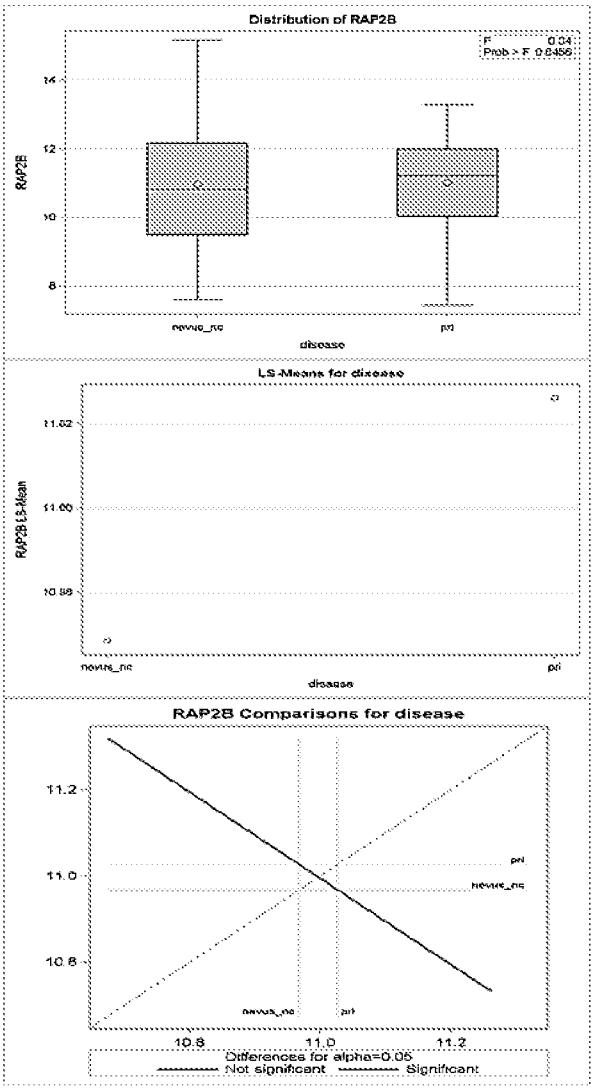
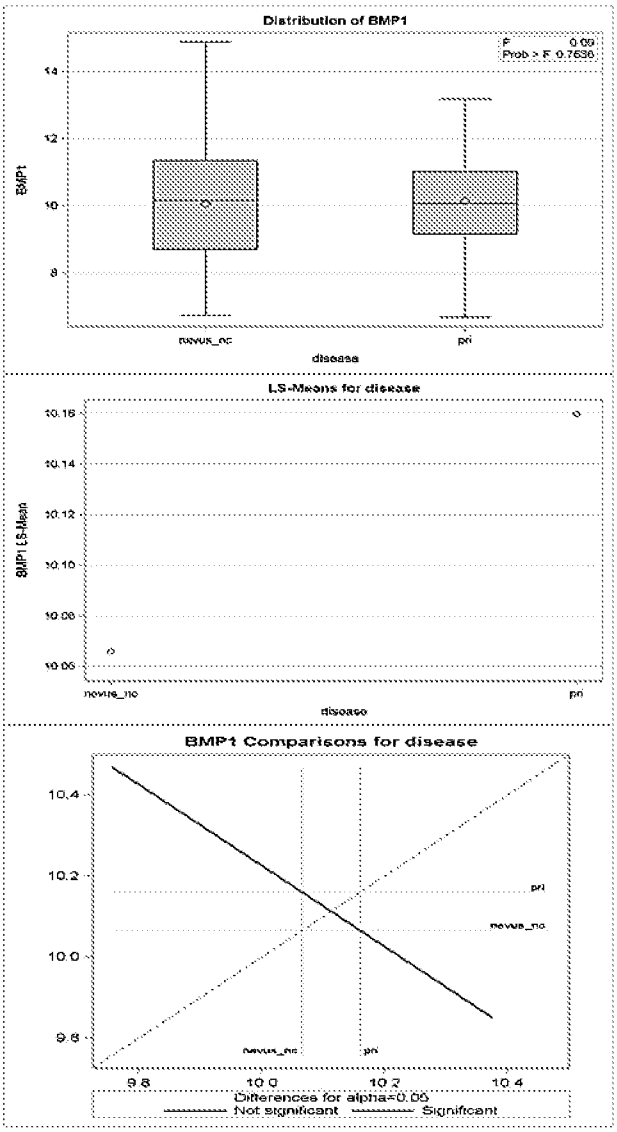


FIG. 2B

BMP1




NCOR2

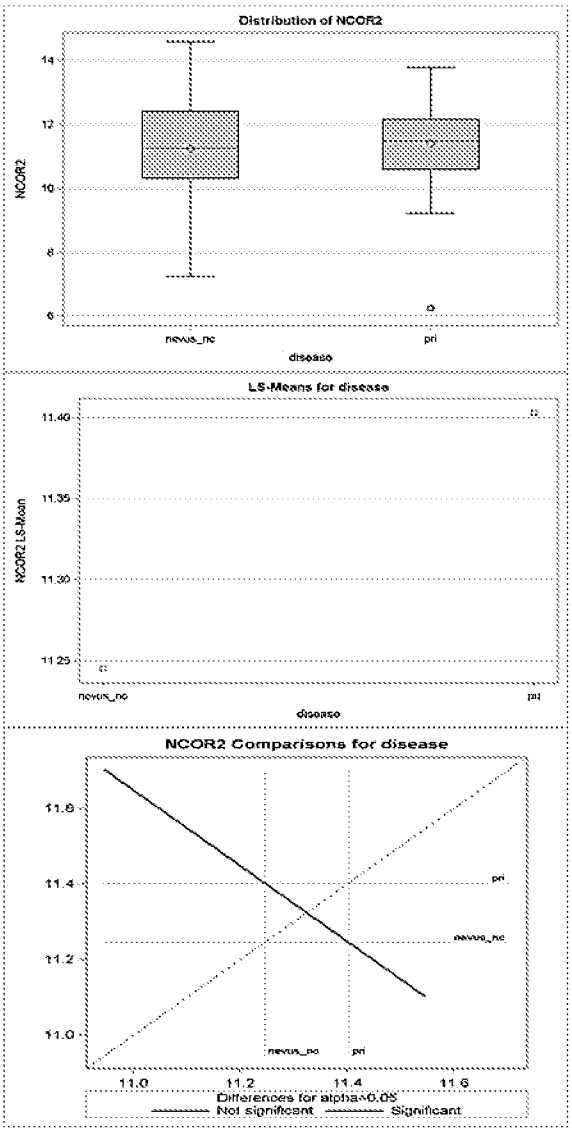


FIG. 3

**B4GALT1 & NR4A1**
(4 normal pairs)

- Logistic Regression estimated using SAS Proc Logistic
  - Binary Logit and iteratively reweighted least squares (Fisher Scoring) optimization
- Very good overall model fit statistics
- No lack of fit indicated by Hosmer and Lemeshow test
- Cross Validation Method:
  - SAS One step approximation based on the full model maximum likelihood estimate versus the maximum likelihood estimate with the $j^{th}$ observation omitted

| Model Convergence Status |
|---|
| Convergence criterion (GCONV=1E-8) satisfied |

| Model Fit Statistics | | |
|---|---|---|
| Criterion | Intercept Only | Intercept and Covariates |
| AIC | 133.851 | 31.908 |
| SC | 136.446 | 39.694 |
| -2 Log L | 131.851 | 25.908 |

| Testing Global Null Hypothesis: BETA=0 | | | |
|---|---|---|---|
| Test | Chi-Square | DF | Pr > ChiSq |
| Likelihood Ratio | 105.9423 | 2 | <.0001 |
| Score | 73.2725 | 2 | <.0001 |
| Wald | 15.8564 | 2 | 0.0004 |

| Hosmer and Lemeshow Goodness-of-Fit Test | | |
|---|---|---|
| Chi-Square | DF | Pr > ChiSq |
| 2.4164 | 8 | 0.9655 |

FIG. 4

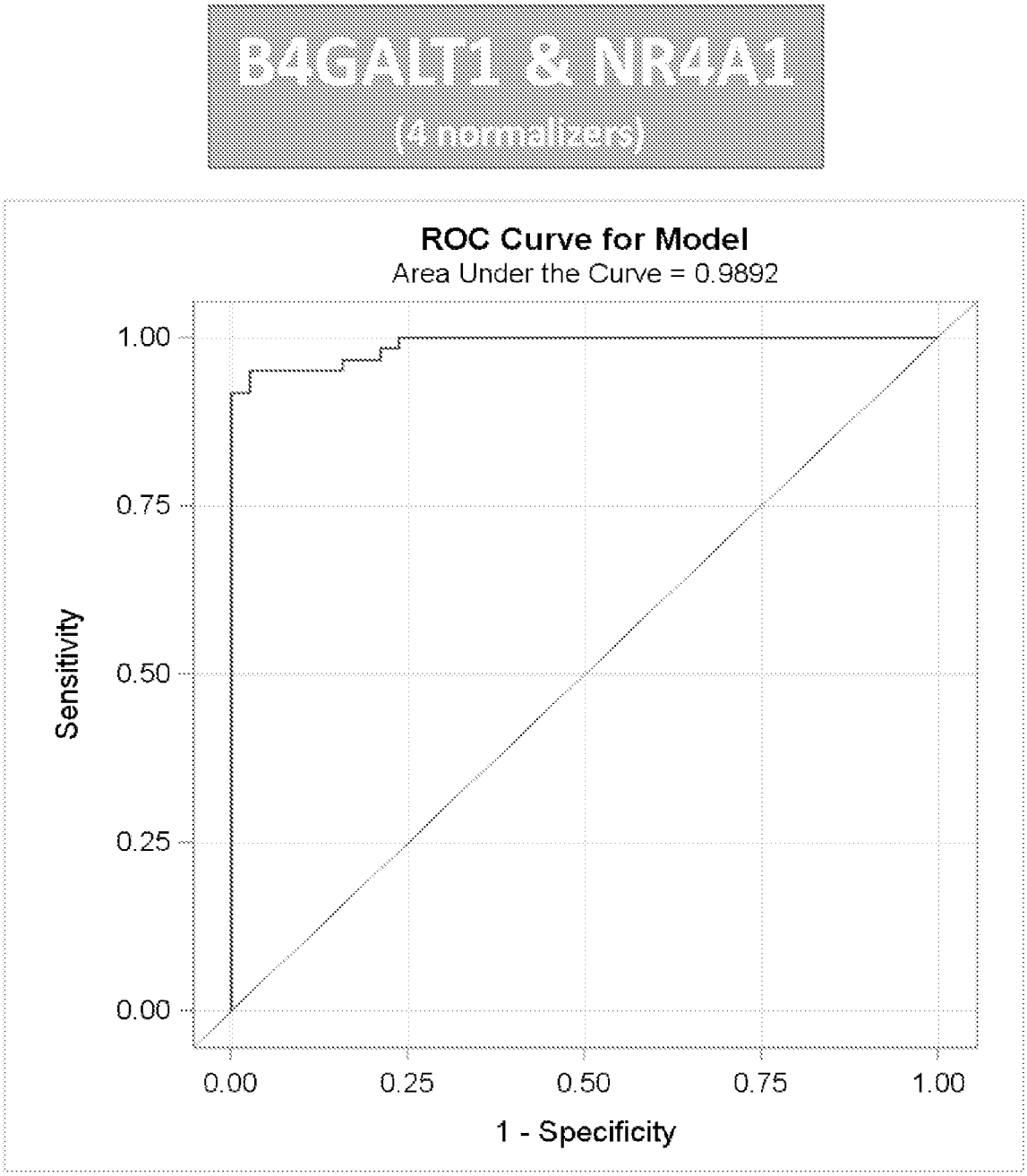

B4GALT1 & NR4A1
(4 normalizers)
ROC Curve for Model
Area Under the Curve = 0.9892
Sensitivity
1.00
0.75
0.50
0.25
0.00
0.00
0.25
0.50
0.75
1.00
1 - Specificity

FIG. 5

Classification Table

| Prob Level | Correct Event | Correct Non-Event | Incorrect Event | Incorrect Non-Event | Percentages Correct | Sensi-tivity | Speci-ficity | False POS | False NEG |
|---|---|---|---|---|---|---|---|---|---|
| 0.640 | 57 | 36 | 2 | 4 | 93.9 | 93.4 | 94.7 | 3.4 | 10.0 |
| 0.660 | 56 | 36 | 2 | 5 | 92.9 | 91.8 | 94.7 | 3.4 | 12.2 |
| 0.680 | 56 | 36 | 2 | 5 | 92.9 | 91.8 | 94.7 | 3.4 | 12.2 |
| 0.700 | 56 | 37 | 1 | 5 | 93.9 | 91.8 | 97.4 | 1.8 | 11.9 |
| 0.720 | 56 | 37 | 1 | 5 | 93.9 | 91.8 | 97.4 | 1.8 | 11.9 |
| 0.740 | 56 | 37 | 1 | 5 | 93.9 | 91.8 | 97.4 | 1.8 | 11.9 |
| 0.760 | 56 | 37 | 1 | 5 | 93.9 | 91.8 | 97.4 | 1.8 | 11.9 |
| 0.780 | 56 | 37 | 1 | 5 | 93.9 | 91.8 | 97.4 | 1.8 | 11.9 |
| 0.800 | 55 | 37 | 1 | 6 | 92.9 | 90.2 | 97.4 | 1.8 | 14.0 |
| 0.820 | 53 | 37 | 1 | 8 | 90.9 | 86.9 | 97.4 | 1.9 | 17.8 |
| 0.840 | 53 | 38 | 0 | 8 | 91.9 | 86.9 | 100.0 | 0.0 | 17.4 |
| 0.860 | 53 | 38 | 0 | 8 | 91.9 | 86.9 | 100.0 | 0.0 | 17.4 |
| 0.880 | 53 | 38 | 0 | 8 | 91.9 | 86.9 | 100.0 | 0.0 | 17.4 |
| 0.900 | 53 | 38 | 0 | 8 | 91.9 | 86.9 | 100.0 | 0.0 | 17.4 |
| 0.920 | 53 | 38 | 0 | 8 | 91.9 | 86.9 | 100.0 | 0.0 | 17.4 |
| 0.940 | 53 | 38 | 0 | 8 | 91.9 | 86.9 | 100.0 | 0.0 | 17.4 |
| 0.960 | 52 | 38 | 0 | 9 | 90.9 | 85.2 | 100.0 | 0.0 | 19.1 |
| 0.980 | 48 | 38 | 0 | 13 | 86.9 | 78.7 | 100.0 | 0.0 | 25.5 |
| 1.000 | 0 | 38 | 0 | 61 | 38.4 | 0.0 | 100.0 | . | 61.6 |

B4GALT1 & NR4A1
(Annualized)

- Can use the model to obtain high specificity with relatively little impact on sensitivity

FIG. 6

**BAGALT1 & NRAA1**

(4 normalizers)

- Can obtain higher sensitivity with relatively little impact on specificity
- Very large range of possible threshold values for calling a sample primary melanoma that maintain high sensitivity with specificity >80%
  - Robust model

Classification Table

| Prob Level | Correct Event | Correct Non-Event | Incorrect Event | Incorrect Non-Event | Percentages Correct | Sensi. tivity | Speci. ficity | False POS | False NEG |
|---|---|---|---|---|---|---|---|---|---|
| 0.000 | 61 | 0 | 38 | 0 | 61.6 | 100.0 | 0.0 | 38.4 | . |
| 0.020 | 61 | 22 | 16 | 0 | 83.8 | 100.0 | 57.9 | 20.8 | 0.0 |
| 0.040 | 61 | 26 | 12 | 0 | 87.9 | 100.0 | 68.4 | 16.4 | 0.0 |
| 0.060 | 60 | 29 | 9 | 1 | 89.9 | 98.4 | 76.3 | 13.0 | 3.3 |
| 0.080 | 60 | 29 | 9 | 1 | 89.9 | 98.4 | 76.3 | 13.0 | 3.3 |
| 0.100 | 60 | 29 | 9 | 1 | 89.9 | 98.4 | 76.3 | 13.0 | 3.3 |
| 0.120 | 60 | 29 | 9 | 1 | 89.9 | 98.4 | 76.3 | 13.0 | 3.3 |
| 0.140 | 58 | 29 | 9 | 3 | 87.9 | 95.1 | 76.3 | 13.4 | 9.4 |
| 0.160 | 58 | 30 | 8 | 3 | 88.9 | 95.1 | 78.9 | 12.1 | 9.1 |
| 0.180 | 58 | 30 | 8 | 3 | 88.9 | 95.1 | 78.9 | 12.1 | 9.1 |
| 0.200 | 58 | 30 | 8 | 3 | 88.9 | 95.1 | 78.9 | 12.1 | 9.1 |
| 0.220 | 58 | 30 | 8 | 3 | 88.9 | 95.1 | 78.9 | 12.1 | 9.1 |
| 0.240 | 58 | 31 | 7 | 3 | 89.9 | 95.1 | 81.6 | 10.8 | 8.8 |
| 0.260 | 58 | 31 | 7 | 3 | 89.9 | 95.1 | 81.6 | 10.8 | 8.8 |
| 0.280 | 58 | 32 | 6 | 3 | 90.9 | 95.1 | 84.2 | 9.4 | 8.6 |
| 0.300 | 58 | 32 | 6 | 3 | 90.9 | 95.1 | 84.2 | 9.4 | 8.6 |
| 0.320 | 58 | 32 | 6 | 3 | 90.9 | 95.1 | 84.2 | 9.4 | 8.6 |
| 0.340 | 58 | 32 | 6 | 3 | 90.9 | 95.1 | 84.2 | 9.4 | 8.6 |
| 0.360 | 58 | 32 | 6 | 3 | 90.9 | 95.1 | 84.2 | 9.4 | 8.6 |
| 0.380 | 58 | 33 | 5 | 3 | 91.9 | 95.1 | 86.8 | 7.9 | 8.3 |
| 0.400 | 58 | 33 | 5 | 3 | 91.9 | 95.1 | 86.8 | 7.9 | 8.3 |
| 0.420 | 58 | 33 | 5 | 3 | 91.9 | 95.1 | 86.8 | 7.9 | 8.3 |
| 0.440 | 58 | 34 | 4 | 3 | 92.9 | 95.1 | 89.5 | 6.5 | 8.1 |
| 0.460 | 58 | 34 | 4 | 3 | 92.9 | 95.1 | 89.5 | 6.5 | 8.1 |
| 0.480 | 58 | 34 | 4 | 3 | 92.9 | 95.1 | 89.5 | 6.5 | 8.1 |
| 0.500 | 58 | 34 | 4 | 3 | 92.9 | 95.1 | 89.5 | 6.5 | 8.1 |
| 0.520 | 58 | 35 | 3 | 3 | 93.9 | 95.1 | 92.1 | 4.9 | 7.9 |
| 0.540 | 58 | 35 | 3 | 3 | 93.9 | 95.1 | 92.1 | 4.9 | 7.9 |
| 0.560 | 58 | 35 | 3 | 3 | 93.9 | 95.1 | 92.1 | 4.9 | 7.9 |
| 0.580 | 58 | 35 | 3 | 3 | 93.9 | 95.1 | 92.1 | 4.9 | 7.9 |
| 0.600 | 58 | 36 | 2 | 3 | 94.9 | 95.1 | 94.7 | 3.3 | 7.7 |
| 0.620 | 58 | 36 | 2 | 3 | 94.9 | 95.1 | 94.7 | 3.3 | 7.7 |

FIG. 7

Other methods validate the B4GALT1 & NR4A1 (4 normalizer) model:

- Model testing with Empirical Covariance "Sandwich" Estimators
  - Estimated using proc GLIMMIX in SAS
  - Sandwich estimators are:
    - Less sensitive to choice of covariance model
    - Minimize bias due to heteroskedastic variances

**Fit Statistics**

| | |
|---|---|
| -2 Log Likelihood | 25.91 |
| AIC (smaller is better) | 31.91 |
| AICC (smaller is better) | 32.16 |
| BIC (smaller is better) | 39.69 |
| CAIC (smaller is better) | 42.69 |
| HQIC (smaller is better) | 35.06 |
| Pearson Chi-Square | 31.16 |
| Pearson Chi-Square / DF | 0.32 |

**Type III Tests of Fixed Effects**

| Effect | Num DF | Den DF | F Value | Pr > F |
|---|---|---|---|---|
| B4GALT1 | 1 | 96 | 10.87 | 0.0014 |
| NR4A1 | 1 | 96 | 14.73 | 0.0002 |

Firth Bias Reduction Penalized Likelihood

**Model Convergence Status**

Convergence criterion (GCONV=1E-8) satisfied.

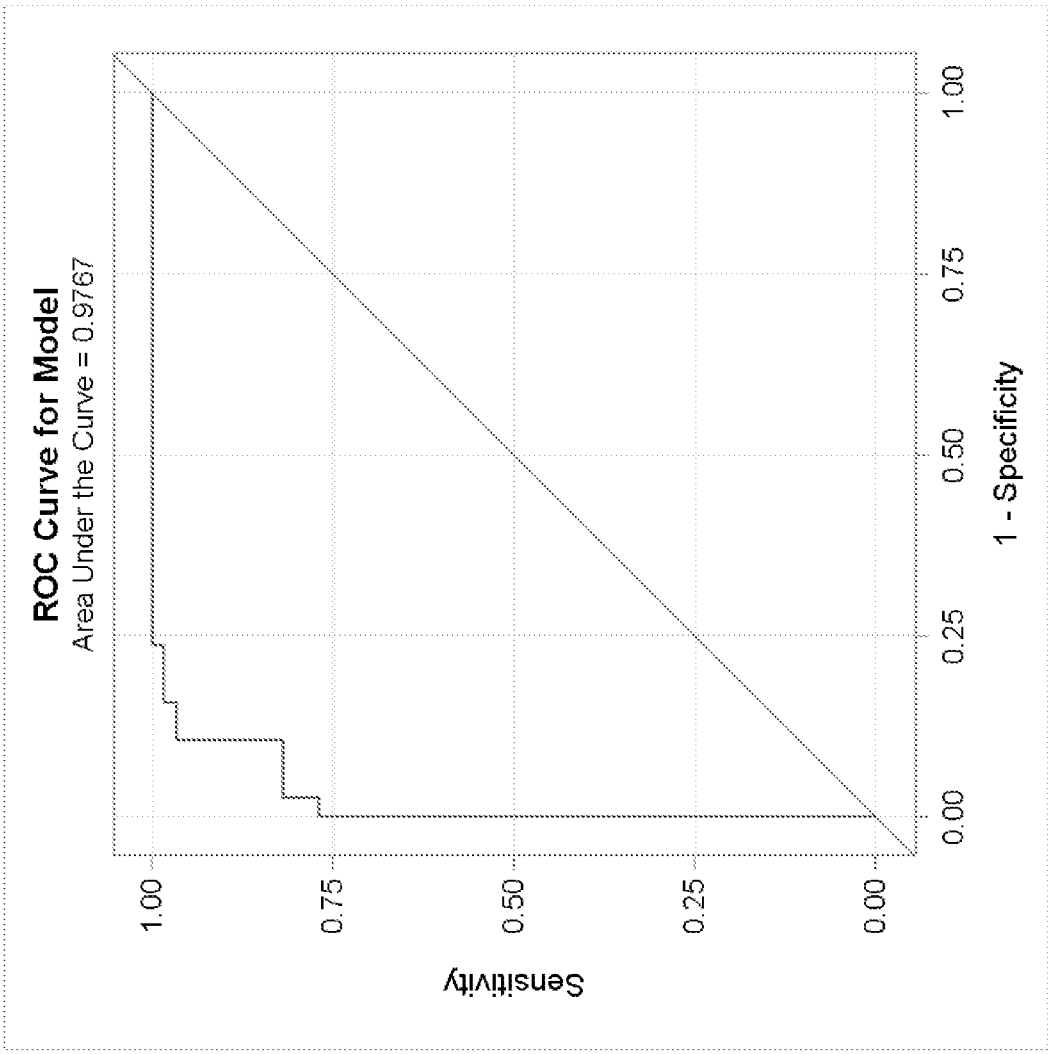
**Model Fit Statistics**

| Criterion | Intercept Only | Intercept and Covariates |
|---|---|---|
| AIC | 124.760 | 29.552 |
| SC | 127.355 | 37.338 |
| -2 Log L | 122.760 | 23.552 |

**Testing Global Null Hypothesis: BETA=0**

| Test | Chi-Square | DF | Pr > ChiSq |
|---|---|---|---|
| Likelihood Ratio | 99.2081 | 2 | <.0001 |
| Score | 73.1213 | 2 | <.0001 |
| Wald | 20.1588 | 2 | <.0001 |

FIG. 8




BAGAIT1 & NR4A1
(2 modified)

| Model Information | |
|---|---|
| Data Set | WORK.MYR_HK |
| Response Variable | disease |
| Number of Response Levels | 2 |
| Model | binary logit |
| Optimization Technique | Fisher's scoring |

| Model Convergence Status |
|---|
| Convergence criterion (GCONV=1E-8) satisfied. |

| Testing Global Null Hypothesis: BETA=0 | | | |
|---|---|---|---|
| Test | Chi-Square | DF | Pr > ChiSq |
| Likelihood Ratio | 93.6780 | 2 | <.0001 |
| Score | 67.4852 | 2 | <.0001 |
| Wald | 22.0216 | 2 | <.0001 |

FIG. 9

BAG4L1 & NR4A1

(Normalized)

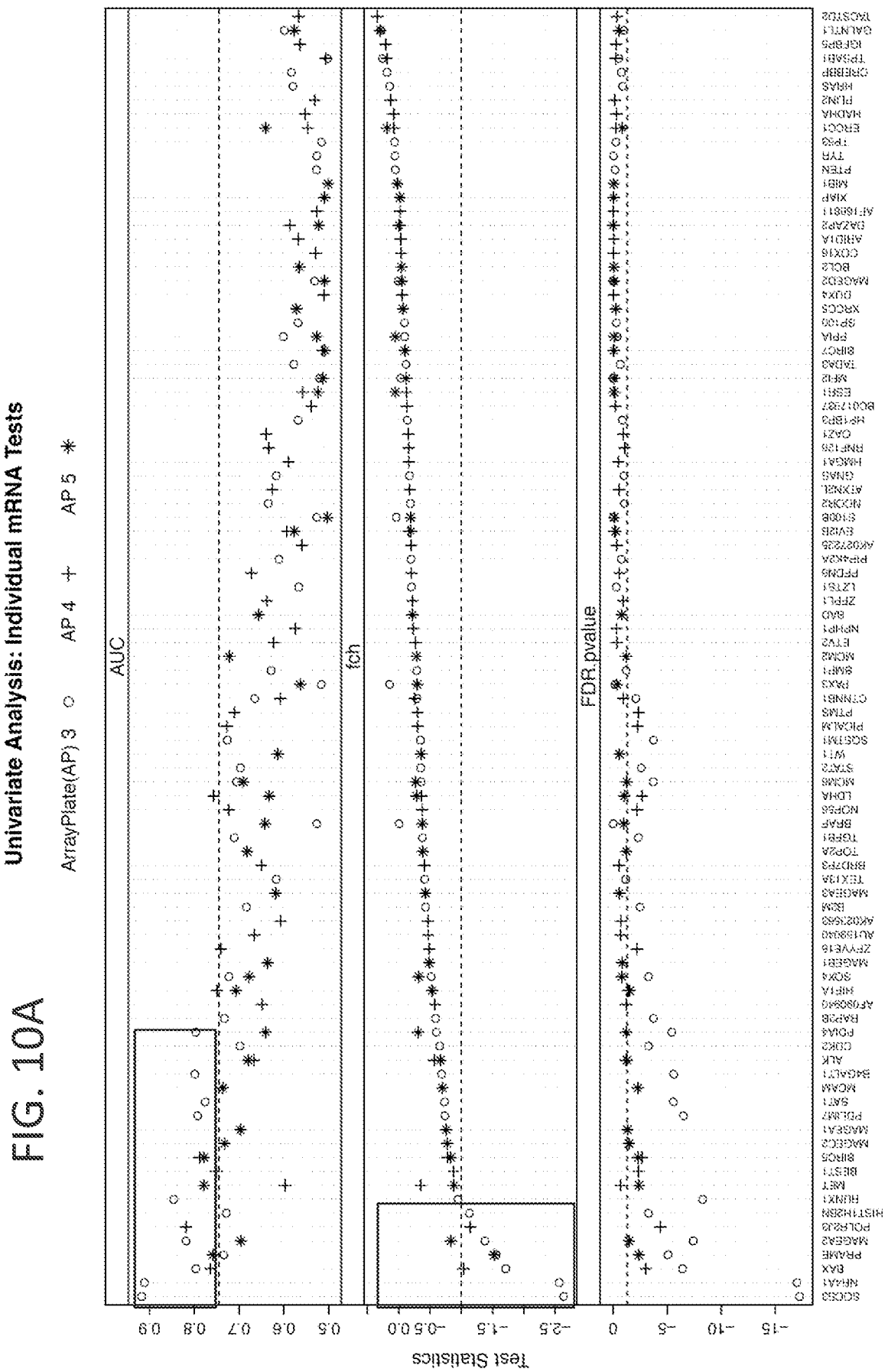
Classification Table

| Prob Level | Correct Event | Correct Non-Event | Incorrect Event | Incorrect Non-Event | Percentages Correct | Percentages Sensitivity | Percentages Specificity | Percentages False POS | Percentages False NEG |
|---|---|---|---|---|---|---|---|---|---|
| 0.000 | 61 | 0 | 38 | 0 | 61.6 | 100.0 | 0.0 | 38.4 | . |
| 0.020 | 61 | 20 | 18 | 0 | 81.8 | 100.0 | 52.6 | 22.8 | 0.0 |
| 0.040 | 61 | 23 | 15 | 0 | 84.8 | 100.0 | 60.5 | 19.7 | 0.0 |
| 0.060 | 61 | 24 | 14 | 0 | 85.9 | 100.0 | 63.2 | 18.7 | 0.0 |
| 0.080 | 61 | 26 | 12 | 0 | 87.9 | 100.0 | 68.4 | 16.4 | 0.0 |
| 0.100 | 60 | 27 | 11 | 1 | 87.9 | 98.4 | 71.1 | 15.5 | 3.6 |
| 0.120 | 60 | 28 | 10 | 1 | 88.9 | 98.4 | 73.7 | 14.3 | 3.4 |
| 0.140 | 60 | 29 | 9 | 1 | 89.9 | 98.4 | 76.3 | 13.0 | 3.3 |
| 0.160 | 60 | 29 | 9 | 1 | 89.9 | 98.4 | 76.3 | 13.0 | 3.3 |
| 0.180 | 60 | 29 | 9 | 1 | 89.9 | 98.4 | 76.3 | 13.0 | 3.3 |
| 0.200 | 60 | 29 | 9 | 1 | 89.9 | 98.4 | 76.3 | 13.0 | 3.3 |
| 0.220 | 60 | 31 | 7 | 1 | 91.9 | 98.4 | 81.6 | 10.4 | 3.1 |
| 0.240 | 60 | 31 | 7 | 1 | 91.9 | 98.4 | 81.6 | 10.4 | 3.1 |
| 0.260 | 60 | 31 | 7 | 1 | 91.9 | 98.4 | 81.6 | 10.4 | 3.1 |
| 0.280 | 60 | 31 | 7 | 1 | 91.9 | 98.4 | 81.6 | 10.4 | 3.1 |
| 0.300 | 59 | 31 | 7 | 2 | 90.9 | 96.7 | 81.6 | 10.6 | 6.1 |
| 0.320 | 59 | 31 | 7 | 2 | 90.9 | 96.7 | 81.6 | 10.6 | 6.1 |
| 0.340 | 59 | 31 | 7 | 2 | 90.9 | 96.7 | 81.6 | 10.6 | 6.1 |
| 0.360 | 58 | 31 | 7 | 3 | 89.9 | 95.1 | 81.6 | 10.8 | 8.8 |
| 0.380 | 58 | 32 | 6 | 3 | 90.9 | 95.1 | 84.2 | 9.4 | 8.6 |
| 0.400 | 58 | 32 | 6 | 3 | 90.9 | 95.1 | 84.2 | 9.4 | 8.6 |
| 0.420 | 58 | 32 | 6 | 3 | 90.9 | 95.1 | 84.2 | 9.4 | 8.6 |
| 0.440 | 58 | 33 | 5 | 3 | 91.9 | 95.1 | 86.8 | 7.9 | 8.3 |
| 0.460 | 58 | 34 | 4 | 3 | 92.9 | 95.1 | 89.5 | 6.5 | 8.1 |
| 0.480 | 58 | 34 | 4 | 3 | 92.9 | 95.1 | 89.5 | 6.5 | 8.1 |
| 0.500 | 57 | 34 | 4 | 4 | 91.9 | 93.4 | 89.5 | 6.6 | 10.5 |

Classification Table

| Prob Level | Correct Event | Correct Non-Event | Incorrect Event | Incorrect Non-Event | Percentages Correct | Percentages Sensitivity | Percentages Specificity | Percentages False POS | Percentages False NEG |
|---|---|---|---|---|---|---|---|---|---|
| 0.500 | 57 | 34 | 4 | 4 | 91.9 | 93.4 | 89.5 | 6.6 | 10.5 |
| 0.520 | 57 | 34 | 4 | 4 | 91.9 | 93.4 | 89.5 | 6.6 | 10.5 |
| 0.540 | 57 | 34 | 4 | 4 | 91.9 | 93.4 | 89.5 | 6.6 | 10.5 |
| 0.560 | 57 | 34 | 4 | 4 | 91.9 | 93.4 | 89.5 | 6.6 | 10.5 |
| 0.580 | 57 | 34 | 4 | 4 | 91.9 | 93.4 | 89.5 | 6.6 | 10.5 |
| 0.600 | 57 | 34 | 4 | 4 | 91.9 | 93.4 | 89.5 | 6.6 | 10.5 |
| 0.620 | 57 | 34 | 4 | 4 | 91.9 | 93.4 | 89.5 | 6.6 | 10.5 |
| 0.640 | 56 | 34 | 4 | 5 | 90.9 | 91.8 | 89.5 | 6.7 | 12.8 |
| 0.660 | 55 | 34 | 4 | 6 | 89.9 | 90.2 | 89.5 | 6.8 | 15.0 |
| 0.680 | 54 | 34 | 4 | 7 | 88.9 | 88.5 | 89.5 | 6.9 | 17.1 |
| 0.700 | 54 | 34 | 4 | 7 | 88.9 | 88.5 | 89.5 | 6.9 | 17.1 |
| 0.720 | 53 | 34 | 4 | 8 | 87.9 | 86.9 | 89.5 | 7.0 | 19.0 |
| 0.740 | 53 | 34 | 4 | 8 | 87.9 | 86.9 | 89.5 | 7.0 | 19.0 |
| 0.760 | 53 | 34 | 4 | 8 | 87.9 | 86.9 | 89.5 | 7.0 | 19.0 |
| 0.780 | 51 | 34 | 4 | 10 | 85.9 | 83.6 | 89.5 | 7.3 | 22.7 |
| 0.800 | 51 | 34 | 4 | 10 | 85.9 | 83.6 | 89.5 | 7.3 | 22.7 |
| 0.820 | 50 | 34 | 4 | 11 | 84.8 | 82.0 | 89.5 | 7.4 | 24.4 |
| 0.840 | 50 | 34 | 4 | 11 | 84.8 | 82.0 | 89.5 | 7.4 | 24.4 |
| 0.860 | 50 | 34 | 4 | 11 | 84.8 | 82.0 | 89.5 | 7.4 | 24.4 |
| 0.880 | 50 | 35 | 3 | 11 | 85.9 | 82.0 | 92.1 | 5.7 | 23.9 |
| 0.900 | 49 | 35 | 3 | 12 | 84.8 | 80.3 | 92.1 | 5.8 | 25.5 |
| 0.920 | 47 | 36 | 2 | 14 | 83.8 | 77.0 | 94.7 | 4.1 | 28.0 |
| 0.940 | 40 | 37 | 1 | 21 | 77.8 | 65.6 | 97.4 | 2.4 | 36.2 |
| 0.960 | 34 | 37 | 1 | 27 | 71.7 | 55.7 | 97.4 | 2.9 | 42.2 |
| 0.980 | 30 | 38 | 0 | 31 | 68.7 | 49.2 | 100.0 | 0.0 | 44.9 |
| 1.000 | 0 | 38 | 0 | 61 | 38.4 | 0.0 | 100.0 | . | 61.6 |

FIG. 10A


Univariate Analysis: Individual mRNA Tests
ArrayPlate(AP) 3 ○ AP 4 + AP 5 ✱
AUC
fch
FDR.pvalue
Test Statistics
0.5 0.6 0.7 0.8 0.9
−2.5 −1.5 −0.5 0.0
−15 −10 −5 0
SOCS3
NR4A1
BAX
PRAME
MAGEA2
POLR2J3
HIST1H2BN
RUNX1
MET
BIRC5
MAGEC2
MAGEA1
POLR1B
SAT1
MCAM
B4GALT1
ALK
CDK2
PDIA4
BAP28
AF090940
HIF1A
SOX4
MAGEB1
ZFYVE16
AU138040
AK092569
IL2M
MAGEA3
TEX13A
BRD7P3
TOP2A
TGFB1
BRAF
NOP56
LDHA
MCM6
STAT3
WT1
SOGTM1
PIK3CA
PTMS
CTNNB1
PAX3
BMP1
MCM2
ETV5
NPHP1
BAD
ZFPL1
LZTS1
PPEF2
Ptpn23A
AK097235
EVI2B
S100B
NCOR2
ATXN3L
GNAS
HMGA1
RNF139
CA21
HPSBP3
BC017387
ESR1
MPG
TADA3
BIRC7
PPIA
SP100
XRCC5
DUX4
MAGED2
BCL2
CDX16
ARID1A
DAZAP2
AF16891
XIAP
MB1
PTEN
TYB
TP53
ERCC1
HADHA
PLIN2
HRAS
CREBBP
TP53BP1
IGFBP5
GALNT1
TMOST02

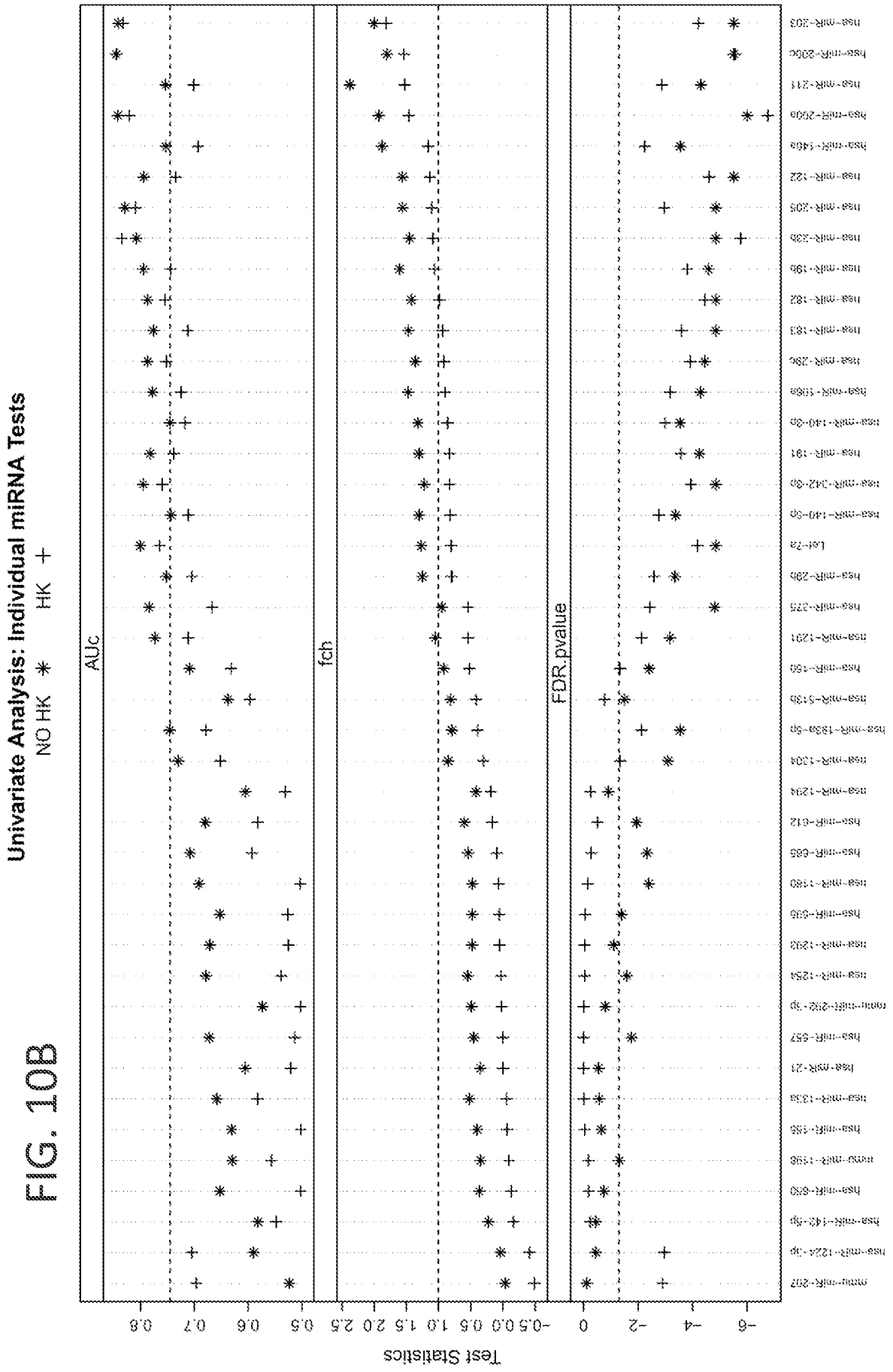

Univariate Analysis: Individual miRNA Tests
NO HK
HK
AUc
fch
FDR.pvalue
Test Statistics
hsa-miR-202
hsa-miR-1224-3p
hsa-miR-142-5p
hsa-miR-650
mmu-miR-1198
hsa-miR-155
hsa-miR-133a
hsa-miR-21
hsa-miR-657
mmu-miR-292-3p
hsa-miR-1254
hsa-miR-1292
hsa-miR-636
hsa-miR-1180
hsa-miR-885
hsa-miR-612
hsa-miR-1294
hsa-miR-1204
hsa-miR-1538-5p
hsa-miR-513b
hsa-miR-150
hsa-miR-1231
hsa-miR-223
hsa-miR-29b
let-7a
hsa-miR-140-5p
hsa-miR-342-3p
hsa-miR-191
hsa-miR-140-3p
hsa-miR-106a
hsa-miR-29c
hsa-miR-183
hsa-miR-182
hsa-miR-19b
hsa-miR-33b
hsa-miR-205
hsa-miR-122
hsa-miR-146a
hsa-miR-200a
hsa-miR-211
hsa-miR-200c
hsa-miR-203

FIG. 10B

FIG. 11

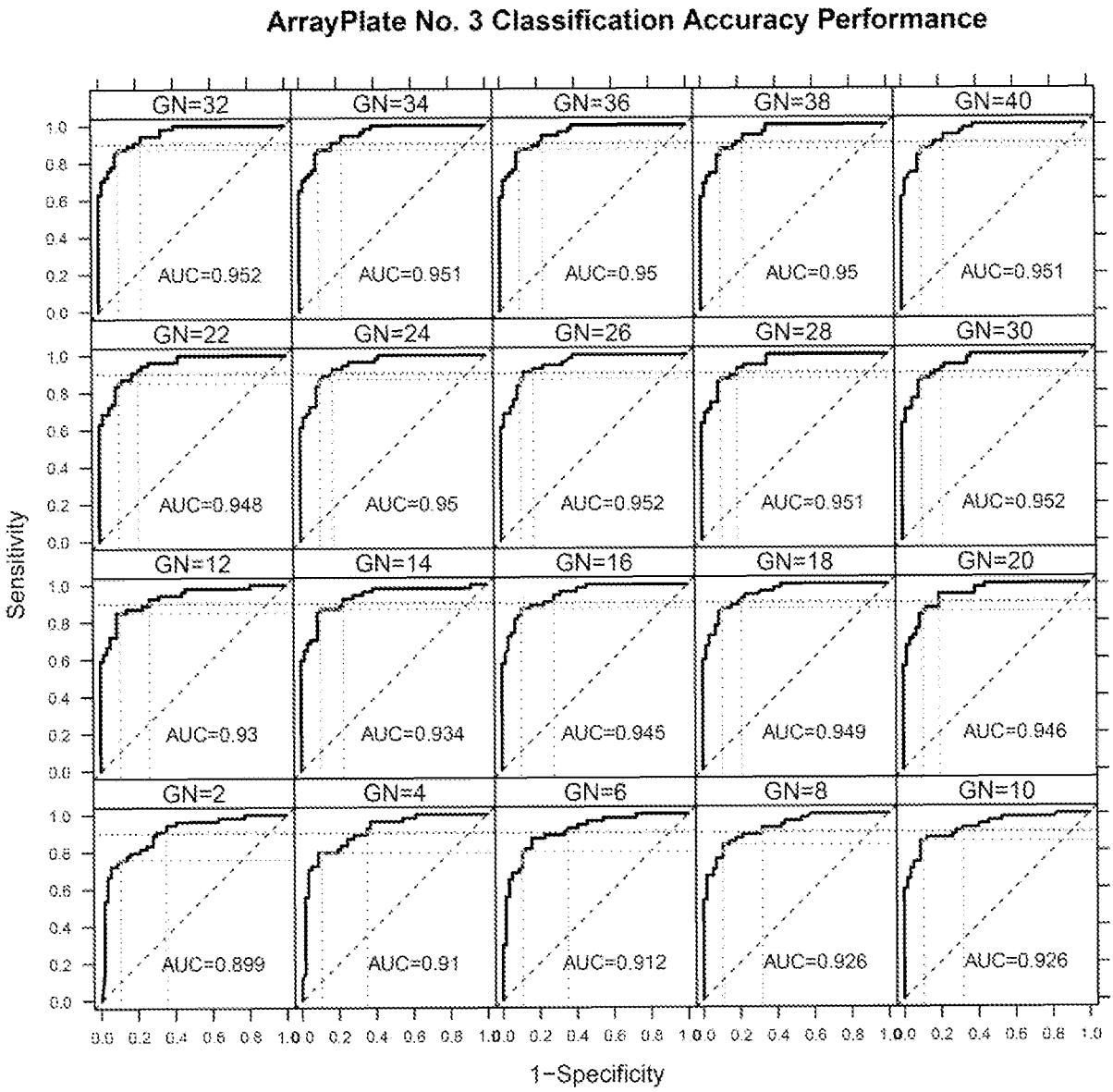
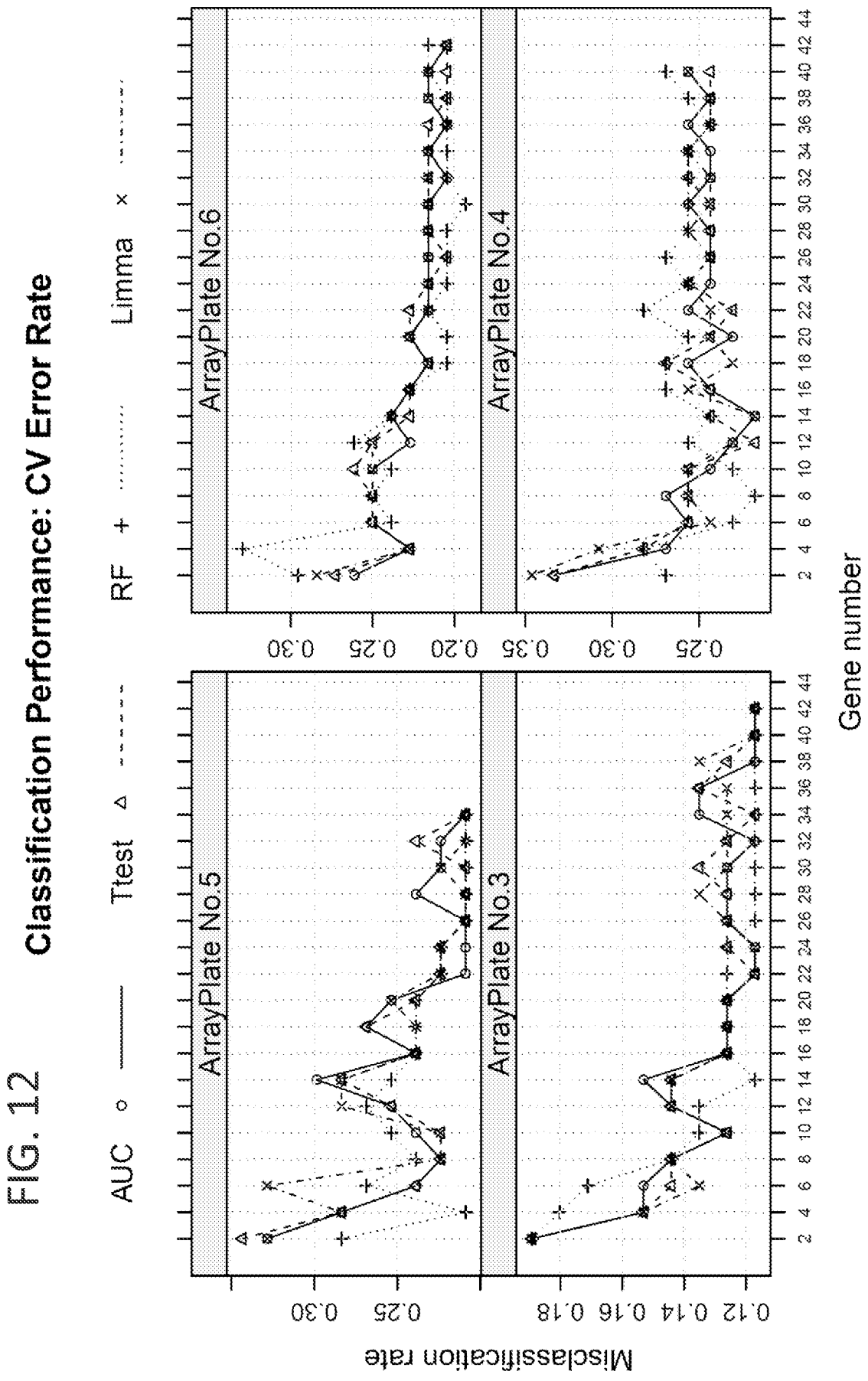

ArrayPlate No. 3 Classification Accuracy Performance
GN=32
GN=34
GN=36
GN=38
GN=40
AUC=0.952
AUC=0.951
AUC=0.95
AUC=0.95
AUC=0.951
GN=22
GN=24
GN=26
GN=28
GN=30
AUC=0.948
AUC=0.95
AUC=0.952
AUC=0.951
AUC=0.952
GN=12
GN=14
GN=16
GN=18
GN=20
AUC=0.93
AUC=0.934
AUC=0.945
AUC=0.949
AUC=0.946
GN=2
GN=4
GN=6
GN=8
GN=10
AUC=0.899
AUC=0.91
AUC=0.912
AUC=0.926
AUC=0.926
Sensitivity
1−Specificity

FIG. 12


Classification Performance: CV Error Rate
AUC
Ttest
RF
Limma
ArrayPlate No.5
ArrayPlate No.3
ArrayPlate No.6
ArrayPlate No.4
Misclassification rate
Gene number

# B4GALT1 & NR4A1
(4 normalizers)