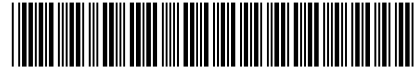
(19)



(11) **EP 3 029 153 A2**

(12) **EUROPEAN PATENT APPLICATION**

(43) Date of publication:
08.06.2016 Bulletin 2016/23

(51) Int Cl.:
C12Q 1/68 *(2006.01)* ***G01N 33/53*** *(2006.01)*
G01N 33/574 *(2006.01)*

(21) Application number: **15189622.2**

(22) Date of filing: **27.09.2011**

(84) Designated Contracting States:
AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR

(30) Priority: **27.09.2010 US 386840 P**
31.03.2011 US 201161470143 P

(62) Document number(s) of the earlier application(s) in accordance with Art. 76 EPC:
11831270.1 / 2 622 104

(71) Applicant: **Somalogic, Inc.**
Boulder, CO 80301 (US)

(72) Inventors:
- **OSTROFF, Rachel M.**
  Westiminster, CO 80031 (US)
- **STEWART, Alex A. E.**
  Waltham, MA 02451 (US)
- **WILLIAMS, Stephen Alaric**
  Boulder, CO 80304 (US)
- **BRODY, Edward N.**
  Boulder, CO 80301 (US)
- **NIKRAD, Malti**
  Boulder, CO 80303 (US)
- **RIEL-MEHAN, Michael**
  Louisville, CO 80027 (US)

(74) Representative: **Clegg, Richard Ian et al**
Mewburn Ellis LLP
City Tower
40 Basinghall Street
London EC2V 5DE (GB)

Remarks:
•Claims filed after the date of filing of the application/after the date of receipt of the divisional application (Rule 68(4) EPC).
•This application was filed on 13-10-2015 as a divisional application to the application mentioned under INID code 62.

(54) **MESOTHELIOMA BIOMARKERS AND USES THEREOF**

(57) The present disclosure includes biomarkers, methods, devices, reagents, systems, and kits for the detection and diagnosis of cancer generally and mesothelioma specifically. In one aspect, the disclosure provides biomarkers that can be used alone or in various combinations to diagnose cancer generally or mesothelioma specifically. In another aspect, methods are provided for diagnosing mesothelioma in an individual, where the methods include detecting, in a biological sample from an individual, at least one biomarker value corresponding to at least one biomarker selected from the group of biomarkers provided in Table 1, wherein the individual is classified as having mesothelioma, or the likelihood of the individual having mesothelioma is determined, based on the at least one biomarker value. In a further aspect, methods are provided for diagnosing cancer generally in an individual, where the methods include detecting, in a biological sample from an individual, at least one biomarker value corresponding to at least one biomarker selected from the group of biomarkers provided in Table 19, wherein the individual is classified as having cancer generally, or the likelihood of the individual having cancer is determined, based on the at least one biomarker value.

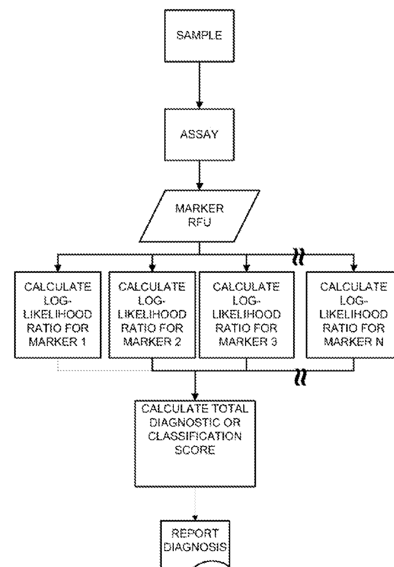


FIG. 1B

Printed by Jouve, 75001 PARIS (FR)

Description

RELATED APPLICATIONS

[0001] This application claims the benefit of U.S. Provisional Application Serial No. 61/386,840, filed September 27, 2010 and U.S. Provisional Application Serial No. 61/470,143, filed March 31, 2011, each of which is incorporated herein by reference in its entirety.

FIELD OF THE INVENTION

[0002] The present application relates generally to the detection of biomarkers and the diagnosis of cancer in an individual and, more specifically, to one or more biomarkers, methods, devices, reagents, systems, and kits for diagnosing cancer, more particularly malignant mesothelioma (mesothelioma), in an individual.

BACKGROUND

[0003] The following description provides a summary of information relevant to the present disclosure and is not an admission that any of the information provided or publications referenced herein is prior art to the present disclosure.

[0004] Mesothelioma is an aggressive, asbestos-related pulmonary cancer that is increasing in incidence. This disease causes an estimated 15,000 to 20,000 deaths per year worldwide. Between 1940 and 1979, approximately 27.5 million people were exposed occupationally to asbestos in the United States. The incidence of mesothelioma in the United States is approximately 3,000 new cases per year and will not peak for another 20 years. Mesothelioma has a latency period of 20-40 years from asbestos exposure, but once diagnosed this aggressive disease is often fatal within 14 months. Because diagnosis is difficult, most patients present at a clinically advanced stage where the possibility of cure is minimal.

[0005] Early diagnosis of mesothelioma in individuals with a history of asbestos exposure is an unmet clinical need. Such exposure may be direct, such as during pipe-laying or installing or removing asbestos-based insulation, or indirect, such as through exposure to vermiculite or coal mining. As the discovery of occupational exposures continues to grow, the need to screen all exposed workers will increase.

[0006] The fact that asbestos exposure is the main causative factor for disease means a high-risk population can be readily identified for clinical screening. Since 1973, the USA Occupational Safety and Health Administration has mandated that individuals with occupational airborne asbestos exposure be monitored for up to 30 years post exposure. Monitouring includes chest X-ray, health history, and spirometry. However, this surveillance has been ineffective in diagnosing early stage mesothelioma or detecting recurrence. As a result, compliance with monitoring is poor, and most disease is detected too late to be cured.

[0007] Currently, most patients are identified due to a pleural effusion, and several consultations are usually necessary before a knowledgeable specialist sees the patient and suspects mesothelioma. A diagnosis is often made through a cytological analysis of a pleural effusion, which has good specificity but is not very sensitive.

[0008] Patients with mesothelioma may present with a variety of symptoms, including:

- Persistent dry or raspy cough (typically non-productive)
- Hemoptysis (coughing up blood)
- Dysphagia (difficulty in swallowing)
- Night sweats or fever
- Unexplained weight loss of 10 percent or more
- Fatigue
- Persistent pain in the chest or rib area, or painful breathing
- Shortness of breath that occurs even when at rest
- The appearance of lumps under the skin on the chest

- Scoliosis towards the side of the malignancy

[0009] These symptoms are non-specific and generally indicate later-stage disease. Many benign pulmonary disease cases undergo invasive procedures because pleural effusion is also a common presentation in patients with asbestosis and pleural plaques.

[0010] Detection of mesothelioma tends to occur during the later stages of the disease. Patient survival from mesothelioma diagnosed at a later stage is poor - less than 15 months for Stage III and worse for Stage IV. Detection at earlier stages, when the disease is resectable and treatable, should increase overall survival and benefit patients.

[0011] Smoking has a strong synergistic effect with asbestos exposure, and the incidence of lung cancer increases 4-5 fold when these two risk factors are combined. Smoking has no effect on the incidence of mesothelioma.

[0012] Biomarker selection for a specific disease state involves first the identification of markers that have a measurable and statistically significant difference in a disease population compared to a control population for a specific medical application. Biomarkers can include secreted or shed molecules that parallel disease development or progression and readily diffuse into the blood stream from mesothelioma or lung cancer tissue or from surrounding tissues and circulating cells in response to a malignancy. The biomarker or set of biomarkers identified are generally clinically validated or shown to be a reliable indicator for the original intended use for which it was selected. Biomarkers can include small molecules, peptides, proteins, and nucleic acids. Some of the key issues that affect the identification of biomarkers include over-fitting of the available data and bias in the data.

[0013] A variety of methods have been utilized in an attempt to identify biomarkers and diagnose disease. For protein-based markers, these include two-dimensional electrophoresis, mass spectrometry, and immunoassay methods. For nucleic acid markers, these include mRNA expression profiles, microRNA profiles, FISH, serial analysis of gene expression (SAGE), methylation profiles, and large scale gene expression arrays.

[0014] The utility of two-dimensional electrophoresis is limited by low detection sensitivity; issues with protein solubility, charge, and hydrophobicity; gel reproducibility, and the possibility of a single spot representing multiple proteins. For mass spectrometry, depending on the format used, limitations revolve around the sample processing and separation, sensitivity to low abundance proteins, signal to noise considerations, and inability to immediately identify the detected protein. Limitations in immunoassay approaches to biomarker discovery are centered on the inability of antibody-based multiplex assays to measure a large number of analytes. One might simply print an array of high-quality antibodies and, without sandwiches, measure the analytes bound to those antibodies. (This would be the formal equivalent of using a whole genome of nucleic acid sequences to measure by hybridization all DNA or RNA sequences in an organism or a cell. The hybridization experiment works because hybridization can be a stringent test for identity. Even very good antibodies are not stringent enough in selecting their binding partners to work in the context of blood or even cell extracts because the protein ensemble in those matrices have extremely different abundances.) Thus, one must use a different approach with immunoassay-based approaches to biomarker discovery - one would need to use multiplexed ELISA assays (that is, sandwiches) to get sufficient stringency to measure many analytes simultaneously to decide which analytes are indeed biomarkers. Sandwich immunoassays do not scale to high content, and thus biomarker discovery using stringent sandwich immunoassays is not possible using standard array formats. Lastly, antibody reagents are subject to substantial lot variability and reagent instability. The instant platform for protein biomarker discovery overcomes this problem.

[0015] Many of these methods rely on or require some type of sample fractionation prior to the analysis. Thus the sample preparation required to run a sufficiently powered study designed to identify and discover statistically relevant biomarkers in a series of well-defined sample populations is extremely difficult, costly, and time consuming. During fractionation, a wide range of variability can be introduced into the various samples. For example, a potential marker could be unstable to the process, the concentration of the marker could be changed, inappropriate aggregation or disaggregation could occur, and inadvertent sample contamination could occur and thus obscure the subtle changes anticipated in early disease.

[0016] It is widely accepted that biomarker discovery and detection methods using these technologies have serious limitations for the identification of diagnostic biomarkers. These limitations include an inability to detect low-abundance biomarkers, an inability to consisgently cover the entire dynamic range of the proteome, irreproducibility in sample processing and fractionation, and overall irreproducibility and lack of robustness of the method. Further, these studies have introduced biases into the data and not adequately addressed the complexity of the sample populations, including appropriate controls, in terms of the distribution and randomization required to identify and validate biomarkers within a target disease population.

[0017] Although efforts aimed at the discovery of new and effective biomarkers have gone on for several decades, the efforts have been largely unsuccessful. Biomarkers for various diseases typically have been identified in academic laboratories, usually through an accidental discovery while doing basic research on some disease process. Based on the discovery and with small amounts of clinical data, papers were published that suggested the identification of a new biomarker. Most of these proposed biomarkers, however, have not been confirmed as real or useful biomarkers; primarily

because the small number of clinical samples tested, provide only weak statistical proof that an effective biomarker has in fact been found. That is, the initial identification was not rigorous with respect to the basic elements of statistics. In each of the years 1994 through 2003, a search of the scientific literature shows that thousands of references directed to biomarkers were published. During that same time frame, however, the FDA approved for diagnostic use, at most, three new protein biomarkers a year, and in several years no new protein biomarkers were approved.

[0018] Based on the history of failed biomarker discovery efforts, mathematical theories have been proposed that further promote the general understanding that biomarkers for disease are rare and difficult to find. Biomarker research based on 2D gels or mass spectrometry supports these notions. Very few useful biomarkers have been identified through these approaches. However, it is usually overlooked that 2D gel and mass spectrometry measure proteins that are present in blood at approximately 1 nM concentrations and higher, and that this ensemble of proteins may well be the least likely to change with disease. Other than the instant biomarker discovery platform, proteomic biomarker discovery platforms that are able to accurately measure protein expression levels at much lower concentrations do not exist.

[0019] Much is known about biochemical pathways for complex human biology. Many biochemical pathways culminate in or are started by secreted proteins that work locally within the pathology, for example growth factors are secreted to stimulate the replication of other cells in the pathology, and other factors are secreted to ward off the immune system, and so on. While many of these secreted proteins work in a paracrine fashion, some operate distally in the body. One skilled in the art with a basic understanding of biochemical pathways would understand that many pathology-specific proteins ought to exist in blood at concentrations below (even far below) the detection limits of 2D gels and mass spectrometry. What must precede the identification of this relatively abundant number of disease biomarkers is a proteomic platform that can analyze proteins at concentrations below those detectable by 2D gels or mass spectrometry.

[0020] Accordingly, a need exists for biomarkers, methods, devices, reagents, systems, and kits that enable: (a) the differentiation of mesothelioma from benign conditions in asbestos exposed individuals; (b) the differentiation of mesothelioma from metastatic disease from other cancers, which may include lung, breast, stomach, kidney, ovary, thymus, and prostate; (c) the differentiation of mesothelioma from lung adenocarcinoma; (d) the detection of mesothelioma biomarkers; and (e) the diagnosis of mesothelioma.

SUMMARY

[0021] The present application includes biomarkers, methods, reagents, devices, systems, and kits for the detection and diagnosis of cancer and more particularly, mesothelioma. The biomarkers of the present application were identified using a multiplex aptamer-based assay which is described in detail in Example 1. By using the aptamer-based biomarker identification method described herein, this application describes a surprisingly large number of mesothelioma biomarkers that are useful for the detection and diagnosis of mesothelioma as well as a large number of cancer biomarkers that are useful for the detection and diagnosis of cancer more generally. In identifying these biomarkers, over 1000 proteins from hundreds of individual samples were measured, some of which were at concentrations in the low femtomolar range. This is about four orders of magnitude lower than biomarker discovery experiments done with 2D gels and/or mass spectrometry.

[0022] While certain of the described mesothelioma biomarkers are useful alone for detecting and diagnosing mesothelioma, methods are described herein for the grouping of multiple subsets of the mesothelioma biomarkers that are useful as a panel of biomarkers. Once an individual biomarker or subset of biomarkers has been identified, the detection or diagnosis of mesothelioma in an individual can be accomplished using any assay platform or format that is capable of measuring differences in the levels of the selected biomarker or biomarkers in a biological sample.

[0023] However, it was only by using the aptamer-based biomarker identification method described herein, wherein over 1000 separate potential biomarker values were individually screened from a large number of individuals having previously been diagnosed either as having or not having mesothelioma that it was possible to identify the mesothelioma biomarkers disclosed herein. This discovery approach is in stark contrast to biomarker discovery from conditioned media or lysed cells as it queries a more patient-relevant system that requires no translation to human pathology.

[0024] Thus, in one aspect of the instant application, one or more biomarkers are provided for use either alone or in various combinations to diagnose mesothelioma or permit the differential diagnosis of mesothelioma from benign conditions such as those found in individuals exposed to asbestos. Exemplary embodiments include the biomarkers provided in Table 1, which as noted above, were identified using a multiplex aptamer-based assay, as described generally in Example 1 and more specifically in Example 2. The markers provided in Table 1 are useful in diagnosing mesothelioma in a high risk population and for distinguishing benign pulmonary diseases in individuals exposed to asbestos from mesothelioma.

[0025] While certain of the described mesothelioma biomarkers are useful alone for detecting and diagnosing mesothelioma, methods are also described herein for the grouping of multiple subsets of the mesothelioma biomarkers that are each useful as a panel of two or more biomarkers. Thus, various embodiments of the instant application provide combinations comprising N biomarkers, wherein N is at least two biomarkers. In other embodiments, N is selected to

be any number from 2-66 biomarkers.

[0026] In yet other embodiments, N is selected to be any number from 2-5, 2-10, 2-15, 2-20, 2-25, 2-30, 2-35, 2-40, 2-45, 2-50, 2-55, 2-60, or 2-66. In other embodiments, N is selected to be any number from 3-5, 3-10, 3-15, 3-20, 3-25, 3-30, 3-35, 3-40, 3-45, 3-50, 3-55, 3-60, or 3-66. In other embodiments, N is selected to be any number from 4-5, 4-10, 4-15, 4-20, 4-25, 4-30, 4-35, 4-40, 4-45, 4-50, 4-55, 4-60, or 4-66. In other embodiments, N is selected to be any number from 5-10, 5-15, 5-20, 5-25, 5-30, 5-35, 5-40, 5-45, 5-50, 5-55, 5-60, or 5-66. In other embodiments, N is selected to be any number from 6-10, 6-15, 6-20, 6-25, 6-30, 6-35, 6-40, 6-45, 6-50, 6-55, 6-60, or 6-66. In other embodiments, N is selected to be any number from 7-10, 7-15, 7-20, 7-25, 7-30, 7-35, 7-40, 7-45, 7-50, 7-55, 7-60, or 7-66. In other embodiments, N is selected to be any number from 8-10, 8-15, 8-20, 8-25, 8-30, 8-35, 8-40, 8-45, 8-50, 8-55, 8-60, or 8-66. In other embodiments, N is selected to be any number from 9-10, 9-15, 9-20, 9-25, 9-30, 9-35, 9-40, 9-45, 9-50, 9-55, 9-60, or 9-66. In other embodiments, N is selected to be any number from 10-15, 10-20, 10-25, 10-30, 10-35, 10-40, 10-45, 10-50, 10-55, 10-60, or 10-66. It will be appreciated that N can be selected to encompass similar, but higher order, ranges.

[0027] In another aspect, a method is provided for diagnosing mesothelioma in an individual, the method including detecting, in a biological sample from an individual, at least one biomarker value corresponding to at least one biomarker selected from the group of biomarkers provided in Table 1, wherein the individual is classified as having mesothelioma based on the at least one biomarker value.

[0028] In another aspect, a method is provided for diagnosing mesothelioma in an individual, the method including detecting, in a biological sample from an individual, biomarker values that each correspond to one of at least N biomarkers selected from the group of biomarkers set forth in Table 1, wherein the likelihood of the individual having mesothelioma is determined based on the biomarker values.

[0029] In another aspect, a method is provided for diagnosing mesothelioma in an individual, the method including detecting, in a biological sample from an individual, biomarker values that each correspond to one of at least N biomarkers selected from the group of biomarkers set forth in Table 1, wherein the individual is classified as having mesothelioma based on the biomarker values, and wherein N = 2-10.

[0030] In another aspect, a method is provided for diagnosing mesothelioma in an individual, the method including detecting, in a biological sample from an individual, biomarker values that each correspond to one of at least N biomarkers selected from the group of biomarkers set forth in Table 1, wherein the likelihood of the individual having mesothelioma is determined based on the biomarker values, and wherein N = 2-10.

[0031] In another aspect, a method is provided for diagnosing that an individual does not have mesothelioma, the method including detecting, in a biological sample from an individual, at least one biomarker value corresponding to at least one biomarker selected from the group of biomarkers set forth in Table 1, wherein the individual is classified as not having mesothelioma based on the at least one biomarker value.

[0032] In another aspect, a method is provided for diagnosing that an individual does not have mesothelioma, the method including detecting, in a biological sample from an individual, biomarker values that each corresponding to one of at least N biomarkers selected from the group of biomarkers set forth in Table 1, wherein the individual is classified as not having mesothelioma based on the biomarker values, and wherein N = 2-10.

[0033] In another aspect, a method is provided for diagnosing mesothelioma, the method including detecting, in a biological sample from an individual, biomarker values that each correspond to a biomarker on a panel of N biomarkers, wherein the biomarkers are selected from the group of biomarkers set forth in Table 1, wherein a classification of the biomarker values indicates that the individual has mesothelioma, and wherein N = 3-10.

[0034] In another aspect, a method is provided for diagnosing mesothelioma, the method including detecting, in a biological sample from an individual, biomarker values that each correspond to a biomarker on a panel of N biomarkers, wherein the biomarkers are selected from the group of biomarkers set forth in Table 1, wherein a classification of the biomarker values indicates that the individual has mesothelioma, and wherein N = 3-10.

[0035] In another aspect, a method is provided for diagnosing mesothelioma, the method including detecting, in a biological sample from an individual, biomarker values that each correspond to a biomarker on a panel of biomarkers selected from the group of panels set forth in Tables 2-11, wherein a classification of the biomarker values indicates that the individual has mesothelioma.

[0036] In another aspect, a method is provided for diagnosing an absence of mesothelioma, the method including detecting, in a biological sample from an individual, biomarker values that each correspond to a biomarker on a panel of N biomarkers, wherein the biomarkers are selected from the group of biomarkers set forth in Table 1, wherein a classification of the biomarker values indicates an absence of mesothelioma in the individual, and wherein N = 3-10.

[0037] In another aspect, a method is provided for diagnosing an absence of mesothelioma, the method including detecting, in a biological sample from an individual, biomarker values that each correspond to a biomarker on a panel of N biomarkers, wherein the biomarkers are selected from the group of biomarkers set forth in Table 1, wherein a classification of the biomarker values indicates an absence of mesothelioma in the individual, and wherein N = 3-10.

[0038] In another aspect, a method is provided for diagnosing an absence of mesothelioma, the method including

detecting, in a biological sample from an individual, biomarker values that each correspond to a biomarker on a panel of biomarkers selected from the group of panels provided in Tables 2-11, wherein a classification of the biomarker values indicates an absence of mesothelioma in the individual.

[0039] In another aspect, a method is provided for diagnosing mesothelioma in an individual, the method including detecting, in a biological sample from an individual, biomarker values that correspond to one of at least N biomarkers selected from the group of biomarkers set forth in Table 1, wherein the individual is classified as having mesothelioma based on a classification score that deviates from a predetermined threshold, and wherein N=2-10.

[0040] In another aspect, a method is provided for diagnosing an absence of mesothelioma in an individual, the method including detecting, in a biological sample from an individual, biomarker values that correspond to one of at least N biomarkers selected from the group of biomarkers set forth in Table 1, wherein said individual is classified as not having mesothelioma based on a classification score that deviates from a predetermined threshold, and wherein N=2-10.

[0041] In another aspect, a computer-implemented method is provided for indicating a likelihood of mesothelioma. The method comprises: retrieving on a computer biomarker information for an individual, wherein the biomarker information comprises biomarker values that each correspond to one of at least N biomarkers, wherein N is as defined above, selected from the group of biomarkers set forth in Table 1; performing with the computer a classification of each of the biomarker values; and indicating a likelihood that the individual has mesothelioma based upon a plurality of classifications.

[0042] In another aspect, a computer-implemented method is provided for classifying an individual as either having or not having mesothelioma. The method comprises: retrieving on a computer biomarker information for an individual, wherein the biomarker information comprises biomarker values that each correspond to one of at least N biomarkers selected from the group of biomarkers provided in Table 1; performing with the computer a classification of each of the biomarker values; and indicating whether the individual has mesothelioma based upon a plurality of classifications.

[0043] In another aspect, a computer program product is provided for indicating a likelihood of mesothelioma. The computer program product includes a computer readable medium embodying program code executable by a processor of a computing device or system, the program code comprising: code that retrieves data attributed to a biological sample from an individual, wherein the data comprises biomarker values that each correspond to one of at least N biomarkers, wherein N is as defined above, in the biological sample selected from the group of biomarkers set forth in Table 1; and code that executes a classification method that indicates a likelihood that the individual has mesothelioma as a function of the biomarker values.

[0044] In another aspect, a computer program product is provided for indicating a Inesothelioma status of an individual. The computer program product includes a computer readable medium embodying program code executable by a processor of a computing device or system, the program code comprising: code that retrieves data attributed to a biological sample from an individual, wherein the data comprises biomarker values that each correspond to one of at least N biomarkers in the biological sample selected from the group of biomarkers provided in Table 1; and code that executes a classification method that indicates a mesothelioma status of the individual as a function of the biomarker values.

[0045] In another aspect, a computer-implemented method is provided for indicating a likelihood of mesothelioma. The method comprises retrieving on a computer biomarker information for an individual, wherein the biomarker information comprises a biomarker value corresponding to a biomarker selected from the group of biomarkers set forth in Table 1; performing with the computer a classification of the biomarker value; and indicating a likelihood that the individual has mesothelioma based upon the classification.

[0046] In another aspect, a computer-implemented method is provided for classifying an individual as either having or not having mesothelioma. The method comprises retrieving from a computer biomarker information for an individual, wherein the biomarker information comprises a biomarker value corresponding to a biomarker selected from the group of biomarkers provided in Table 1; performing with the computer a classification of the biomarker value; and indicating whether the individual has mesothelioma based upon the classification.

[0047] In still another aspect, a computer program product is provided for indicating a likelihood of mesothelioma. The computer program product includes a computer readable medium embodying program code executable by a processor of a computing device or system, the program code comprising: code that retrieves data attributed to a biological sample from an individual, wherein the data comprises a biomarker value corresponding to a biomarker in the biological sample selected from the group of biomarkers set forth in Table 1; and code that executes a classification method that indicates a likelihood that the individual has mesothelioma as a function of the biomarker value.

[0048] In still another aspect, a computer program product is provided for indicating a mesothelioma status of an individual. The computer program product includes a computer readable medium embodying program code executable by a processor of a computing device or system, the program code comprising: code that retrieves data attributed to a biological sample from an individual, wherein the data comprises a biomarker value corresponding to a biomarker in the biological sample selected from the group of biomarkers provided in Table 1; and code that executes a classification method that indicates a mesothelioma status of the individual as a function of the biomarker value.

[0049] While certain of the described biomarkers are also useful alone for detecting and diagnosing general cancer, methods are described herein for the grouping of multiple subsets of the biomarkers that are useful as a panel of

biomarkers for detecting and diagnosing cancer in general. Once an individual biomarker or subset of biomarkers has been identified, the detection or diagnosis of cancer in an individual can be accomplished using any assay platform or format that is capable of measuring differences in the levels of the selected biomarker or biomarkers in a biological sample.

[0050] However, it was only by using the aptamer-based biomarker identification method described herein, wherein over 1000 separate potential biomarker values were individually screened from a large number of individuals having previously been diagnosed either as having or not having cancer that it was possible to identify the cancer biomarkers disclosed herein. This discovery approach is in stark contrast to biomarker discovery from conditioned media or lysed cells as it queries a more patient-relevant system that requires no translation to human pathology.

[0051] Thus, in one aspect of the instant application, one or more biomarkers are provided for use either alone or in various combinations to diagnose cancer. Exemplary embodiments include the biomarkers provided in Table 19, which were identified using a multiplex aptamer-based assay, as described generally in Example 1 and more specifically in Example 5. The markers provided in Table 19 are useful in distinguishing individuals who have cancer from those who do not have cancer.

[0052] While certain of the described cancer biomarkers are useful alone for detecting and diagnosing cancer, methods are also described herein for the grouping of multiple subsets of the cancer biomarkers that are each useful as a panel of three or more biomarkers. Thus, various embodiments of the instant application provide combinations comprising N biomarkers, wherein N is at least three biomarkers. In other embodiments, N is selected to be any number from 3-22 biomarkers.

[0053] In yet other embodiments, N is selected to be any number from 2-5, 2-10, 2-15, 2-20, or 2-22. In other embodiments, N is selected to be any number from 3-5, 3-10, 3-15, 3-20, or 3-22. In other embodiments, N is selected to be any number from 4-5, 4-10, 4-15, 4-20, or 4-22. In other embodiments, N is selected to be any number from 5-10, 5-15, 5-20, or 5-22. In other embodiments, N is selected to be any number from 6-10, 6-15, 6-20, or 6-22. In other embodiments, N is selected to be any number from 7-10, 7-15, 7-20, or 7-22. In other embodiments, N is selected to be any number from 8-10, 8-15, 8-20, or 8-22. In other embodiments, N is selected to be any number from 9-10, 9-15, 9-20, or 9-22. In other embodiments, N is selected to be any number from 10-15, 10-20, or 10-22. It will be appreciated that N can be selected to encompass similar, but higher order, ranges.

[0054] In another aspect, a method is provided for diagnosing cancer in an individual, the method including detecting, in a biological sample from an individual, at least one biomarker value corresponding to at least one biomarker selected from the group of biomarkers provided in Table 19, wherein the individual is classified as having cancer based on the at least one biomarker value.

[0055] In another aspect, a method is provided for diagnosing cancer in an individual, the method including detecting, in a biological sample from an individual, biomarker values that each correspond to one of at least N biomarkers selected from the group of biomarkers set forth in Table 19, wherein the likelihood of the individual having cancer is determined based on the biomarker values.

[0056] In another aspect, a method is provided for diagnosing cancer in an individual, the method including detecting, in a biological sample from an individual, biomarker values that each correspond to one of at least N biomarkers selected from the group of biomarkers set forth in Table 19, wherein the individual is classified as having cancer based on the biomarker values, and wherein N = 3-10.

[0057] In another aspect, a method is provided for diagnosing cancer in an individual, the method including detecting, in a biological sample from an individual, biomarker values that each correspond to one of at least N biomarkers selected from the group of biomarkers set forth in Table 19, wherein the likelihood of the individual having cancer is determined based on the biomarker values, and wherein N = 3-10.

[0058] In another aspect, a method is provided for diagnosing that an individual does not have cancer, the method including detecting, in a biological sample from an individual, at least one biomarker value corresponding to at least one biomarker selected from the group of biomarkers set forth in Table 19, wherein the individual is classified as not having cancer based on the at least one biomarker value.

[0059] In another aspect, a method is provided for diagnosing that an individual does not have cancer, the method including detecting, in a biological sample from an individual, biomarker values that each corresponding to one of at least N biomarkers selected from the group of biomarkers set forth in Table 19, wherein the individual is classified as not having cancer based on the biomarker values, and wherein N = 3-10.

[0060] In another aspect, a method is provided for diagnosing cancer, the method including detecting, in a biological sample from an individual, biomarker values that each corresponds to a biomarker on a panel of N biomarkers, wherein the biomarkers are selected from the group of biomarkers set forth in Table 19, wherein a classification of the biomarker values indicates that the individual has cancer, and wherein N = 3-10.

[0061] In another aspect, a method is provided for diagnosing cancer, the method including detecting, in a biological sample from an individual, biomarker values that each corresponds to a biomarker on a panel of N biomarkers, wherein the biomarkers are selected from the group of biomarkers set forth in Table 19, wherein a classification of the biomarker values indicates that the individual has cancer, and wherein N = 3-10.

[0062] In another aspect, a method is provided for diagnosing cancer, the method including detecting, in a biological sample from an individual, biomarker values that each corresponds to a biomarker on a panel of biomarkers selected from the group of panels set forth in Tables 20-29 wherein a classification of the biomarker values indicates that the individual has cancer.

[0063] In another aspect, a method is provided for diagnosing an absence of cancer, the method including detecting, in a biological sample from an individual, biomarker values that each correspond to a biomarker on a panel of N biomarkers, wherein the biomarkers are selected from the group of biomarkers set forth in Table 19, wherein a classification of the biomarker values indicates an absence of cancer in the individual, and wherein N = 3-10.

[0064] In another aspect, a method is provided for diagnosing an absence of cancer, the method including detecting, in a biological sample from an individual, biomarker values that each correspond to a biomarker on a panel of N biomarkers, wherein the biomarkers are selected from the group of biomarkers set forth in Table 19, wherein a classification of the biomarker values indicates an absence of cancer in the individual, and wherein N = 3-10.

[0065] In another aspect, a method is provided for diagnosing an absence of cancer, the method including detecting, in a biological sample from an individual, biomarker values that each correspond to a biomarker on a panel of biomarkers selected from the group of panels provided in Tables 20-29, wherein a classification of the biomarker values indicates an absence of cancer in the individual.

[0066] In another aspect, a method is provided for diagnosing cancer in an individual, the method including detecting, in a biological sample from an individual, biomarker values that correspond to one of at least N biomarkers selected from the group of biomarkers set forth in Table 19, wherein the individual is classified as having cancer based on a classification score that deviates from a predetermined threshold, and wherein N=3-10.

[0067] In another aspect, a method is provided for diagnosing an absence of cancer in an individual, the method including detecting, in a biological sample from an individual, biomarker values that correspond to one of at least N biomarkers selected from the group of biomarkers set forth in Table 19, wherein said individual is classified as not having cancer based on a classification score that deviates from a predetermined threshold, and wherein N=3-10.

[0068] In another aspect, a computer-implemented method is provided for indicating a likelihood of cancer. The method comprises: retrieving on a computer biomarker information for an individual, wherein the biomarker information comprises biomarker values that each correspond to one of at least N biomarkers, wherein N is as defined above, selected from the group of biomarkers set forth in Table 19; performing with the computer a classification of each of the biomarker values; and indicating a likelihood that the individual has cancer based upon a plurality of classifications.

[0069] In another aspect, a computer-implemented method is provided for classifying an individual as either having or not having cancer. The method comprises: retrieving on a computer biomarker information for an individual, wherein the biomarker information comprises biomarker values that each correspond to one of at least N biomarkers selected from the group of biomarkers provided in Table 19; performing with the computer a classification of each of the biomarker values; and indicating whether the individual has cancer based upon a plurality of classifications.

[0070] In another aspect, a computer program product is provided for indicating a likelihood of cancer. The computer program product includes a computer readable medium embodying program code executable by a processor of a computing device or system, the program code comprising: code that retrieves data attributed to a biological sample from an individual, wherein the data comprises biomarker values that each correspond to one of at least N biomarkers, wherein N is as defined above, in the biological sample selected from the group of biomarkers set forth in Table 19; and code that executes a classification method that indicates a likelihood that the individual has cancer as a function of the biomarker values.

[0071] In another aspect, a computer program product is provided for indicating a cancer status of an individual. The computer program product includes a computer readable medium embodying program code executable by a processor of a computing device or system, the program code comprising: code that retrieves data attributed to a biological sample from an individual, wherein the data comprises biomarker values that each correspond to one of at least N biomarkers in the biological sample selected from the group of biomarkers provided in Table 19; and code that executes a classification method that indicates a cancer status of the individual as a function of the biomarker values.

[0072] In another aspect, a computer-implemented method is provided for indicating a like-lihood of cancer. The method comprises retrieving on a computer biomarker information for an individual, wherein the biomarker information comprises a biomarker value corresponding to a biomarker selected from the group of biomarkers set forth in Table 19; performing with the computer a classification of the biomarker value; and indicating a likelihood that the individual has cancer based upon the classification.

[0073] In another aspect, a computer-implemented method is provided for classifying an individual as either having or not having cancer. The method comprises retrieving from a computer biomarker information for an individual, wherein the biomarker information comprises a biomarker value corresponding to a biomarker selected from the group of biomarkers provided in Table 19; performing with the computer a classification of the biomarker value; and indicating whether the individual has cancer based upon the classification.

[0074] In still another aspect, a computer program product is provided for indicating a likelihood of cancer. The computer

program product includes a computer readable medium embodying program code executable by a processor of a computing device or system, the program code comprising: code that retrieves data attributed to a biological sample from an individual, wherein the data comprises a biomarker value corresponding to a biomarker in the biological sample selected from the group of biomarkers set forth in Table 19; and code that executes a classification method that indicates a likelihood that the individual has cancer as a function of the biomarker value.

[0075] In still another aspect, a computer program product is provided for indicating a cancer status of an individual. The computer program product includes a computer readable medium embodying program code executable by a processor of a computing device or system, the program code comprising: code that retrieves data attributed to a biological sample from an individual, wherein the data comprises a biomarker value corresponding to a biomarker in the biological sample selected from the group of biomarkers provided in Table 19; and code that executes a classification method that indicates a cancer status of the individual as a function of the biomarker value.

BRIEF DESCRIPTION OF THE DRAWINGS

[0076]

Figure 1A is a flowchart for an exemplary method for detecting mesothelioma in a biological sample.

Figure 1B is a flowchart for an exemplary method for detecting mesothelioma in a biological sample using a naïve Bayes classification method.

Figure 2 shows a ROC curve for a single biomarker, CDH1, using a naïve Bayes classifier for a test that detects mesothelioma.

Figure 3 shows ROC curves for biomarker panels of from two to ten biomarkers using naïve Bayes classifiers for a test that detects mesothelioma.

Figure 4 illustrates the increase in the classification score (AUC) as the number of biomarkers is increased from one to ten using naïve Bayes classification for a mesothelioma panel.

Figure 5 shows the measured biomarker distributions for CDH1 as a cumulative distribution function (cdf) in log-transformed RFU for the asbestos exposed individuals combined (solid line) and the mesothelioma disease group (dotted line) along with their curve fits to a formal cdf (dashed lines) used to train the naïve Bayes classifiers.

Figure 6 illustrates an exemplary computer system for use with various computer-implemented methods described herein.

Figure 7 is a flowchart for a method of indicating the likelihood that an individual has mesothelioma in accordance with one embodiment.

Figure 8 is a flowchart for a method of indicating the likelihood that an individual has mesothelioma in accordance with one embodiment.

Figure 9 illustrates an exemplary aptamer assay that can be used to detect one or more mesothelioma biomarkers in a biological sample.

Figure 10 shows a histogram of frequencies for which biomarkers were used in building classifiers to distinguish between mesothelioma and the asbestos exposed individuals from an aggregated set of potential biomarkers.

Figure 11A shows a pair of histograms summarizing all possible single protein naïve Bayes classifier scores (AUC) using the biomarkers set forth in Table 1 (black) and a set of random markers (grey).

Figure 11B shows a pair of histograms summarizing all possible two-protein protein naive Bayes classifier scores (AUC) using the biomarkers set forth in Table 1 (black) and a set of random markers (grey).

Figure 11C shows a pair of histograms summarizing all possible three-protein naïve Bayes classifier scores (AUC) using the biomarkers set forth in Table 1 (black) and a set of random markers (grey).

Figure 12 shows the AUC for naïve Bayes classifiers using from 2-10 markers selected from the full panel and the scores obtained by dropping the best 5, 10, and 15 markers during classifier generation.

Figure 13A shows a set of ROC curves modeled from the data in Table 14 for panels of from two to five markers.

Figure 13B shows a set of ROC curves computed from the training data for panels of from two to five markers as in Figure 12A.

Figures 14A and 14B show a comparison of performance between ten cancer biomarkers selected by a greedy selection procedure described in Example 5 (Table 19) and 1,000 randomly sampled sets of ten "non marker" biomarkers. The mean AUC for the ten cancer biomarkers in Table 19 is shown as a dotted vertical line. In Figure 14A, sets of ten "non markers" were randomly selected that were not selected by the greedy procedure described in Example 5. In Figure 14B, the same procedure as 14A was used; however, the sampling was restricted to the remaining 56 mesothelioma biomarkers from Table 1 that were not selected by the greedy procedure described in Example 5.

Figure 15 shows receiver operating characteristic (ROC) curves for the 3 naïve Bayes classifiers set forth in Table 19. For each study, the area under the curve (AUC) is also displayed next to the legend.

DETAILED DESCRIPTION

[0077] Reference will now be made in detail to representative embodiments of the invention. While the invention will be described in conjunction with the enumerated embodiments, it will be understood that the invention is not intended to be limited to those embodiments. On the contrary, the invention is intended to cover all alternatives, modifications, and equivalents that may be included within the scope of the present invention as defined by the claims.

[0078] One skilled in the art will recognize many methods and materials similar or equivalent to those described herein, which could be used in and are within the scope of the practice of the present invention. The present invention is in no way limited to the methods and materials described.

[0079] Unless defined otherwise, technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any methods, devices, and materials similar or equivalent to those described herein can be used in the practice or testing of the invention, the preferred methods, devices and materials are now described.

[0080] All publications, published patent documents, and patent applications cited in this application are indicative of the level of skill in the art(s) to which the application pertains. All publications, published patent documents, and patent applications cited herein are hereby incorporated by reference to the same extent as though each individual publication, published patent document, or patent application was specifically and individually indicated as being incorporated by reference.

[0081] As used in this application, including the appended claims, the singular forms "a," "an," and "the" include plural references, unless the content clearly dictates otherwise, and are used interchangeably with "at least one" and "one or more." Thus, reference to "an aptamer" includes mixtures of aptamers, reference to "a probe" includes mixtures of probes, and the like.

[0082] As used herein, the term "about" represents an insignificant modification or variation of the numerical value such that the basic function of the item to which the numerical value relates is unchanged.

[0083] As used herein, the terms "comprises," "comprising," "includes," "including," "contains," "containing," and any variations thereof, are intended to cover a non-exclusive inclusion, such that a process, method, product-by-process, or composition of matter that comprises, includes, or contains an element or list of elements does not include only those elements but may include other elements not expressly listed or inherent to such process, method, product-by-process, or composition of matter.

[0084] The present application includes biomarkers, methods, devices, reagents, systems, and kits for the detection and diagnosis of mesothelioma and cancer more generally.

[0085] In one aspect, one or more biomarkers are provided for use either alone or in various combinations to diagnose mesothelioma, permit the differential diagnosis of mesothelioma from non-malignant conditions found in individuals exposed to asbestos, monitor mesothelioma recurrence, or address other clinical indications. As described in detail below, exemplary embodiments include the biomarkers provided in Table 1, which were identified using a multiplex aptamer-based assay that is described generally in Example 1 and more specifically in Example 2.

[0086] Table 1 sets forth the findings obtained from analyzing hundreds of individual blood samples from mesothelioma cases, and hundreds of equivalent individual blood samples from asbestos exposed individuals. The asbestos exposed individuals group was designed to match the populations with which a mesothelioma diagnostic test can have the most benefit, including asymptomatic individuals and symptomatic individuals. High risk for mesothelioma includes occupational or environmental exposure to asbestos and related fibrous materials including carbon nanotubes and fibrous silicates and exposure to ionizing radiation.

[0087] The potential biomarkers were measured in individual samples rather than pooling the disease and control blood; this allowed a better understanding of the individual and group variations in the phenotypes associated with the presence and absence of disease (in this case mesothelioma). Since more than 1000 protein measurements were made on each sample, and several hundred samples from each of the disease and the control populations were individually measured, Table 1 resulted from an analysis of an uncommonly large set of data. The measurements were analyzed using the methods described in the section, "Classification of Biomarkers and Calculation of Disease Scores" herein. Table 1 lists the 66 biomarkers found to be useful in distinguishing samples obtained from individuals with mesothelioma from "control" samples obtained from asbestos exposed individuals.

[0088] While certain of the described mesothelioma biomarkers are useful alone for detecting and diagnosing mesothelioma, methods are also described herein for the grouping of multiple subsets of the mesothelioma biomarkers, where each grouping or subset selection is useful as a panel of three or more biomarkers, interchangeably referred to herein as a "biomarker panel" and a panel. Thus, various embodiments of the instant application provide combinations comprising N biomarkers, wherein N is at least two biomarkers. In other embodiments, N is selected from 2-66 biomarkers.

[0089] In yet other embodiments, N is selected to be any number from 2-5, 2-10, 2-15, 2-20, 2-25, 2-30, 2-35, 2-40, 2-45, 2-50, 2-55, 2-60, or 2-66. In other embodiments, N is selected to be any number from 3-5, 3-10, 3-15, 3-20, 3-25, 3-30, 3-35, 3-40, 3-45, 3-50, 3-55, 3-60, or 3-66. In other embodiments, N is selected to be any number from 4-5, 4-10,

4-15, 4-20, 4-25, 4-30, 4-35, 4-40, 4-45, 4-50, 4-55, 4-60, or 4-66. In other embodiments, N is selected to be any number from 5-10, 5-15, 5-20, 5-25, 5-30, 5-35, 5-40, 5-45, 5-50, 5-55, 5-60, or 5-66. In other embodiments, N is selected to be any number from 6-10, 6-15, 6-20, 6-25, 6-30, 6-35, 6-40, 6-45, 6-50, 6-55, 6-60, or 6-66. In other embodiments, N is selected to be any number from 7-10, 7-15, 7-20, 7-25, 7-30, 7-35, 7-40, 7-45, 7-50, 7-55, 7-60, or 7-66. In other embodiments, N is selected to be any number from 8-10, 8-15, 8-20, 8-25, 8-30, 8-35, 8-40, 8-45, 8-50, 8-55, 8-60, or 8-66. In other embodiments, N is selected to be any number from 9-10, 9-15, 9-20, 9-25, 9-30, 9-35, 9-40, 9-45, 9-50, 9-55, 9-60, or 9-66. In other embodiments, N is selected to be any number from 10-15, 10-20, 10-25, 10-30, 10-35, 10-40, 10-45, 10-50, 10-55, 10-60, or 10-66. It will be appreciated that N can be selected to encompass similar, but higher order, ranges.

[0090] In one embodiment, the number of biomarkers useful for a biomarker subset or panel is based on the sensitivity and specificity value for the particular combination of biomarker values. The terms "sensitivity" and "specificity" are used herein with respect to the ability to correctly classify an individual, based on one or more biomarker values detected in their biological sample, as having mesothelioma or not having mesothelioma. "Sensitivity" indicates the performance of the biomarker(s) with respect to correctly classifying individuals that have mesothelioma. "Specificity" indicates the performance of the biomarker(s) with respect to correctly classifying individuals who do not have mesothelioma. For example, 85% specificity and 90% sensitivity for a panel of markers used to test a set of control samples and mesothelioma samples indicates that 85% of the control samples were correctly classified as control samples by the panel, and 90% of the mesothelioma samples were correctly classified as mesothelioma samples by the panel. The desired or preferred minimum value can be determined as described in Example 3. Representative panels are set forth in Tables 4-11, which set forth a series of 100 different panels of 3-10 biomarkers, which have the vindicated levels of specificity and sensitivity for each panel. The total number of occurrences of each marker in each of these panels is indicated in Table 12.

[0091] In one aspect, mesothelioma is detected or diagnosed in an individual by conducting an assay on a biological sample from the individual and detecting biomarker values that each correspond to at least one of the biomarkers CDH1, BMPER or F9 and at least N additional biomarkers selected from the list of biomarkers in Table 1, wherein N equals 2, 3, 4, 5, 6, 7, 8, or 9. In a further aspect, mesothelioma is detected or diagnosed in an individual by conducting an assay on a biological sample from the individual and detecting biomarker values that each correspond to the biomarkers CDH11, BMPER or F9 and one of at least N additional biomarkers selected from the list of biomarkers in Table 1, wherein N equals 1, 2, 3, 4, 5, 6, or 7. In a further aspect, mesothelioma is detected or diagnosed in an individual by conducting an assay on a biological sample from the individual and detecting biomarker values that each correspond to the biomarker CDH1 and one of at least N additional biomarkers selected from the list of biomarkers in Table 1, wherein N equals 2, 3, 4, 5, 6, 7, 8, or 9. In a further aspect, mesothelioma is detected or diagnosed in an individual by conducting an assay on a biological sample from the individual and detecting biomarker values that each correspond to the biomarker BMPER and one of at least N additional biomarkers selected from the list of biomarkers in Table 1, wherein N equals 2, 3, 4, 5, 6, 7, 8, or 9. In a further aspect, mesothelioma is detected or diagnosed in an individual by conducting an assay on a biological sample from the individual and detecting biomarker values that each correspond to the biomarker F9 and one of at least N additional biomarkers selected from the list of biomarkers in Table 1, wherein N equals 2, 3, 4, 5, 6, 7, 8, or 9.

[0092] The mesothelioma biomarkers identified herein represent a relatively large number of choices for subsets or panels of biomarkers that can be used to effectively detect or diagnose mesothelioma. Selection of the desired number of such biomarkers depends on the specific combination of biomarkers chosen. It is important to remember that panels of biomarkers for detecting or diagnosing mesothelioma may also include biomarkers not found in Table 1, and that the inclusion of additional biomarkers not found in Table 1 may reduce the number of biomarkers in the particular subset or panel that is selected from Table 1. The number of biomarkers from Table 1 used in a subset or panel may also be reduced if additional biomedical information is used in conjunction with the biomarker values to establish acceptable sensitivity and specificity values for a given assay.

[0093] Another factor that can affect the number of biomarkers to be used in a subset or panel of biomarkers is the procedures used to obtain biological samples from individuals who are being diagnosed for mesothelioma. In a carefully controlled sample procurement environment, the number of biomarkers necessary to meet desired sensitivity and specificity values will be lower than in a situation where there can be more variation in sample collection, handling and storage. In developing the list of biomarkers set forth in Table 1, multiple sample collection sites were utilized to collect data for classifier training. This provides for more robust biomarkers that are less sensitive to variations in sample collection, handling and storage, but can also require that the number of biomarkers in a subset or panel be larger than if the training data were all obtained under very similar conditions.

[0094] One aspect of the instant application can be described generally with reference to Figures 1A and 1B. A biological sample is obtained from an individual or individuals of interest. The biological sample is then assayed to detect the presence of one or more (N) biomarkers of interest and to determine a biomarker value for each of said N biomarkers (referred to in Figure 1B as marker RFU). Once a biomarker has been detected and a biomarker value assigned each marker is scored or classified as described in detail herein. The marker scores are then combined to provide a total diagnostic score, which indicates the likelihood that the individual from whom the sample was obtained has mesothelioma.

[0095] "Biological sample", "sample", and "test sample" are used interchangeable herein to refer to any material, biological fluid, tissue, or cell obtained or otherwise derived from an individual. This includes blood (including whole blood, leukocytes, peripheral blood mononuclear cells, buffy coat, plasma, and serum), sputum, tears, mucus, nasal washes, nasal aspirate, breath, urine, semen, saliva, peritoneal washings, ascites, cystic fluid, meningeal fluid, amniotic fluid, glandular fluid, pancreatic fluid, lymph fluid, pleural fluid, cytologic fluid,nipple aspirate, bronchial aspirate, bronchial brushing, synovial fluid, joint aspirate, organ secretions, cells, a cellular extract, and cerebrospinal fluid. This also includes experimentally separated fractions of all of the preceding. For example, a blood sample can be fractionated into serum, plasma or into fractions containing particular types of blood cells, such as red blood cells or white blood cells (leukocytes). If desired, a sample can be a combination of samples from an individual, such as a combination of a tissue and fluid sample. The term "biological sample" also includes materials containing homogenized solid material, such as from a stool sample, a tissue sample, or a tissue biopsy, for example. The term "biological sample" also includes materials derived from a tissue culture or a cell culture. Any suitable methods for obtaining a biological sample can be employed; exemplary methods include, e.g., phlebotomy, swab (e.g., buccal swab), and a fine needle aspirate biopsy procedure. Exemplary tissues susceptible to fine needle aspiration include lymph node, lung, lung washes, BAL (bronchoalveolar lavage),pleura, thyroid, breast, pancreas and liver. Samples can also be collected, e.g., by micro dissection (e.g., laser capture micro dissection (LCM) or laser micro dissection (LMD)), bladder wash, smear (e.g., a PAP smear), or ductal lavage. A "biological sample" obtained or derived from an individual includes any such sample that has been processed in any suitable manner after being obtained from the individual.

[0096] Further, it should be realized that a biological sample can be derived by taking biological samples from a number of individuals and pooling them or pooling an aliquot of each individual's biological sample. The pooled sample can be treated as a sample from a single individual and if the presence of cancer is established in the pooled sample, then each individual biological sample can be re-tested to determine which individual(s) have mesothelioma.

[0097] For purposes of this specification, the phrase "data attributed to a biological sample from an individual" is intended to mean that the data in some form derived from, or were generated using, the biological sample of the individual. The data may have been reformatted, revised, or mathematically altered to some degree after having been generated, such as by conversion from units in one measurement system to units in another measurement system; but, the data are understood to have been derived from, or were generated using, the biological sample.

[0098] "Target", "target molecule", and "analyte" are used interchangeably herein to refer to any molecule of interest that may be present in a biological sample. A "molecule of interest" includes any minor variation of a particular molecule, such as, in the case of a protein, for example, minor variations in amino acid sequence, disulfide bond formation, glycosylation, lipidation, acetylation, phosphorylation, or any other manipulation or modification, such as conjugation with a labeling component, which does not substantially alter the identity of the molecule. A "target molecule", "target", or "analyte" is a set of copies of one type or species of molecule or multi-molecular structure. "Target molecules", "targets", and "analytes" refer to more than one such set of molecules. Exemplary target molecules include proteins, polypeptides, nucleic acids, carbohydrates, lipids, polysaccharides, glycoproteins, hormones, receptors, antigens, antibodies, autoantibodies, affybodies, antibody mimics, viruses, pathogens, toxic substances, substrates, metabolites, transition state analogs, cofactors, inhibitors, drugs, dyes, nutrients, growth factors, cells, tissues, and any fragment or portion of any of the foregoing.

[0099] As used herein, "polypeptides," "peptide," and "protein" are used interchangeably herein to refer to polymers of amino acids of any length. The polymer may be linear or branched, it may comprise modified amino acids, and it may be interrupted by non-amino acids. The terms also encompass an amino acid polymer that has been modified naturally or by intervention; for example, disulfide bond formation, glycosylation, lipidation, acetylation, phosphorylation, or any other manipulation or modification, such as conjugation with a labeling component. Also included within the definition are, for example, polypeptides containing one or more analogs of an amino acid (including, for example, unnatural amino acids, etc.), as well as other modifications known in the art. Polypeptides can be single chains or associated chains. Also included within the definition are preproteins and intact mature proteins; peptides or polypeptides derived from a mature protein; fragments of a protein; splice variants; recombinant forms of a protein; protein variants with amino acid modifications, deletions, or substitutions; digests; and post-translational modifications, such as glycosylation, acetylation, phosphorylation, and the like.

[0100] As used herein, "marker" and "biomarker" are used interchangeably to refer to a target molecule that indicates or is a sign of a normal or abnormal process in an individual or of a disease or other condition in an individual. More specifically, a "marker" or "biomarker" is an anatomic, physiologic, biochemical, or molecular parameter associated with the presence of a specific physiological state or process, whether normal or abnormal, and, if abnormal, whether chronic or acute. Biomarkers are detectable and measurable by a variety of methods including laboratory assays and medical imaging. When a biomarker is a protein, it is also possible to use the expression of the corresponding gene as a surrogate measure of the amount or presence or absence of the corresponding protein biomarker in a biological sample or methylation state of the gene encoding the biomarker or proteins that control expression of the biomarker.

[0101] As used herein, "biomarker value", "value", "biomarker level", and "level" are used interchangeably to refer to

a measurement that is made using any analytical method for detecting the biomarker in a biological sample and that indicates the presence, absence, absolute amount or concentration, relative amount or concentration, titer, a level, an expression level, a ratio of measured levels, or the like, of, for, or corresponding to the biomarker in the biological sample. The exact nature of the "value" or "level" depends on the specific design and components of the particular analytical method employed to detect the biomarker.

[0102] When a biomarker indicates or is a sign of an abnormal process or a disease or other condition in an individual, that biomarker is generally described as being either over-expressed or under-expressed as compared to an expression level or value of the biomarker that indicates or is a sign of a normal process or an absence of a disease or other condition in an individual. "Up-regulation", "up-regulated", "over-expression", "over-expressed", and any variations thereof are used interchangeably to refer to a value or level of a biomarker in a biological sample that is greater than a value or level (or range of values or levels) of the biomarker that is typically detected in similar biological samples from healthy or normal individuals. The terms may also refer to a value or level of a biomarker in a biological sample that is greater than a value or level (or range of values or levels) of the biomarker that may be detected at a different stage of a particular disease.

[0103] "Down-regulation", "down-regulated", "under-expression", "under-expressed", and any variations thereof are used interchangeably to refer to a value or level of a biomarker in a biological sample that is less than a value or level (or range of values or levels) of the biomarker that is typically detected in similar biological samples from healthy or normal individuals. The terms may also refer to a value or level of a biomarker in a biological sample that is less than a value or level (or range of values or levels) of the biomarker that may be detected at a different stage of a particular disease.

[0104] Further, a biomarker that is either over-expressed or under-expressed can also be referred to as being "differentially expressed" or as having a "differential level" or "differential value" as compared to a "normal" expression level or value of the biomarker that indicates or is a sign of a formal process or an absence of a disease or other condition in an individual. Thus, "differential expression" of a biomarker can also be referred to as a variation from a "normal" expression level of the biomarker.

[0105] The term "differential gene expression" and "differential expression" are used interchangeably to refer to a gene (or its corresponding protein expression product) whose expression is activated to a higher or lower level in a subject suffering from a specific disease, relative to its expression in a formal or control subject. The terms also include genes (or the corresponding protein expression products) whose expression is activated to a higher or lower level at different stages of the same disease. It is also understood that a differentially expressed gene may be either activated or inhibited at the nucleic acid level or protein level, or may be subject to alternative splicing to result in a different polypeptide product. Such differences may be evidenced by a variety of changes including mRNA levels, surface expression, secretion or other partitioning of a polypeptide. Differential gene expression may include a comparison of expression between two or more genes or their gene products; or a comparison of the ratios of the expression between two or more genes or their gene products; or even a comparison of two differently processed products of the same gene, which differ between normal subjects and subjects suffering from a disease; or between various stages of the same disease. Differential expression includes both quantitative, as well as qualitative, differences in the temporal or cellular expression pattern in a gene or its expression products among, for example, formal and diseased cells, or among cells which have undertone different disease events or disease stages.

[0106] As used herein, "individual" refers to a test subject or patient. The individual can be a mammal or a non-mammal. In various embodiments, the individual is a mammal. A mammalian individual can be a human or non-human. In various embodiments, the individual is a human. A healthy or normal individual is an individual in which the disease or condition of interest (including, for example, pleural and peritoneal mesothelium diseases, pleural abnormality-associated diseases, or other pleural abnormality conditions) is not detectable by conventional diagnostic methods.

[0107] "Diagnose", "diagnosing", "diagnosis", and variations thereof refer to the detection, determination, or recognition of a health status or condition of an individual on the basis of one or more signs, symptoms, data, or other information pertaining to that individual. The health status of an individual can be diagnosed as healthy / normal (i.e., a diagnosis of the absence of a disease or condition) or diagnosed as ill / abnormal (i.e., a diagnosis of the presence, or an assessment of the characteristics, of a disease or condition). The terms "diagnose", "diagnosing", "diagnosis", etc., encompass, with respect to a particular disease or condition, the initial detection of the disease; the characterization or classification of the disease; the detection of the progression, remission, or recurrence of the disease; and the detection of disease response after the administration of a treatment or therapy to the individual. The diagnosis of mesothelioma includes distinguishing individuals who have cancer from individuals who do not. It further includes distinguishing asbestos exposed individuals from mesothelioma.

[0108] "Prognose", "prognosing", "prognosis", and variations thereof refer to the prediction of a future course of a disease or condition in an individual who has the disease or condition (e.g., predicting patient survival), and such terms encompass the evaluation of disease response after the administration of a treatment or therapy to the individual.

[0109] "Evaluate", "evaluating", "evaluation", and variations thereof encompass both "diagnose" and "prognose" and also encompass determinations or predictions about the future course of a disease or condition in an individual who

does not have the disease as well as determinations or predictions regarding the likelihood that a disease or condition will recur in an individual who apparently has been cured of the disease. The term "evaluate" also encompasses assessing an individual's response to a therapy, such as, for example, predicting whether an individual is likely to respond favorably to a therapeutic agent or is unlikely to respond to a therapeutic agent (or will experience toxic or other undesirable side effects, for example), selecting a therapeutic agent for administration to an individual, or monitoring or determining an individual's response to a therapy that has been administered to the individual. Thus, "evaluating" mesothelioma can include, for example, any of the following: prognosing the future course of mesothelioma in an individual; predicting the recurrence of mesothelioma in an individual who apparently has been cured of mesothelioma; or determining or predicting an individual's response to a mesothelioma treatment or selecting a mesothelioma treatment to administer to an individual based upon a determination of the biomarker values derived from the individual's biological sample.

[0110] Any of the following examples may be referred to as either "diagnosing" or "evaluating" mesothelioma: initially detecting the presence or absence of mesothelioma; determining a specific stage, type or sub-type, or other classification or characteristic of mesothelioma; determining whether a suspicious pleural abnormality is benign or malignant mesothelioma; or detecting/monitoring mesothelioma progression (e.g., monitoring tumor growth or metastatic spread), remission, or recurrence.

[0111] As used herein, "additional biomedical information" refers to one or more evaluations of an individual, other than using any of the biomarkers described herein, that are associated with cancer risk or, more specifically, mesothelioma risk. "Additional biomedical information" includes any of the following: physical descriptors of an individual, including a peritoneal or pleural abnormality or effusion observed by any of contrast-enhanced multislice (multidetector) helical computed tomography (CT) scanning with three dimensional reconstruction, chest X-ray, PET scan, ultrasound, magnetic resonance imaging (MRI); asbestos exposure history; spirometry measurements; the height and/or weight of an individual; change in weight; the ethnicity of an individual; occupational history; family history of mesothelioma (or other cancer); the presence of a genetic marker(s) correlating with a higher risk of mesothelioma (or other cancer) in the individual or a family member; the presence or absence of a pleural abnormality; size of pleural abnormality; location of pleural abnormality; morphology of pleural abnormality and associated pleural abnormality region (e.g., as observed through imaging); clinical symptoms such as dyspnea, chest pain, palpable chest wall masses, pleural effusion, scoliosis towards the side of the malignancy, weight loss; gene expression values; physical descriptors of an individual, including physical descriptors observed by radiologic imaging; the height and/or weight of an individual; the gender of an individual; the ethnicity of an individual; smoking history; occupational history; exposure to known carcinogens (e.g., exposure to any of asbestos, radon gas, chemicals, smoke from fires, and air pollution, which can include emissions from stationary or mobile sources such as industrial/factory or auto/marine/aircraft emissions); exposure to second-hand smoke; and family history of mesothelioma or other cancer. Testing of biomarker levels in combination with an evaluation of any additional biomedical information, including other laboratory tests (e.g., concentration of mesothelin, soluble mesothelin-related peptide, or osteopontin), may, for example, improve sensitivity, specificity, and/or AUC for detecting mesothelioma

[0112] (or other mesothelioma-related uses) as compared to biomarker testing alone or evaluating any particular item of additional biomedical information alone (e.g., ultrasound imaging alone). Additional biomedical information can be obtained from an individual using routine techniques known in the art, such as from the individual themselves by use of a routine patient questionnaire or health history questionnaire, etc., or from a medical practitioner, etc. Testing of biomarker levels in combination with an evaluation of any additional biomedical information may, for example, improve sensitivity, specificity, and/or AUC for detecting mesothelioma (or other mesothelioma-related uses) as compared to biomarker testing alone or evaluating any particular item of additional biomedical information alone (e.g., CT imaging alone).

[0113] The term "area under the curve" or "AUC" refers to the area under the curve of a receiver operating characteristic (ROC) curve, both of which are well known in the art. AUC measures are useful for comparing the accuracy of a classifier across the complete data range. Classifiers with a greater AUC have a greater capacity to classify unknowns correctly between two groups of interest (e.g., mesothelioma samples and normal or control samples). ROC curves are useful for plotting the performance of a particular feature (e.g., any of the biomarkers described herein and/or any item of additional biomedical information) in distinguishing between two populations (e.g., cases having mesothelioma and controls without mesothelioma). Typically, the feature data across the entire population (e.g., the cases and controls) are sorted in ascending order based on the value of a single feature. Then, for each value for that feature, the true positive and false positive rates for the data are calculated. The true positive rate is determined by counting the number of cases above the value for that feature and then dividing by the total number of cases. The false positive rate is determined by counting the number of controls above the value for that feature and then dividing by the total number of controls. Although this definition refers to scenarios in which a feature is elevated in cases compared to controls, this definition also applies to scenarios in which a feature is lower in cases compared to the controls (in such a scenario, samples below the value for that feature would be counted). ROC curves can be generated for a single feature as well as for other single outputs, for example, a combination of two or more features can be mathematically combined (e.g., added, subtracted, multiplied, etc.) to provide a single sum value, and this single sum value can be plotted in a ROC curve. Additionally, any combination of multiple features, in which the combination derives a single output value, can be

plotted in a ROC curve. These combinations of features may comprise a test. The ROC curve is the plot of the true positive rate (sensitivity) of a test against the false positive rate (1-specificity) of the test.

[0114] As used herein, "detecting" or "determining" with respect to a biomarker value includes the use of both the instrument required to observe and record a signal corresponding to a biomarker value and the material/s required to generate that signal. In various embodiments, the biomarker value is detected using any suitable method, including fluorescence, chemiluminescence, surface plasmon resonance, surface acoustic waves, mass spectrometry, infrared spectroscopy, Roman spectroscopy, atomic force microscopy, scanning tunneling microscopy, electrochemical detection methods, nuclear magnetic resonance, quantum dots, and the like.

[0115] "Solid support" refers herein to any substrate having a surface to which molecules may be attached, directly or indirectly, through either covalent or non-covalent bonds. A "solid support" can have a variety of physical formats, which can include, for example, a membrane; a chip (e.g., a protein chip); a slide (e.g., a glass slide or coverslip); a column; a hollow, solid, semi-solid, pore- or cavity- containing particle, such as, for example, a bead; a gel; a fiber, including a fiber optic material; a matrix; and a sample receptacle. Exemplary sample receptacles include sample wells, tubes, capillaries, vials, and any other vessel, groove or indentation capable of holding a sample. A sample receptacle can be contained on a multi-sample platform, such as a microtiter plate, slide, microfluidics device, and the like. A support can be composed of a natural or synthetic material, an organic or inorganic material. The composition of the solid support on which capture reagents are attached generally depends on the method of attachment (e.g., covalent attachment). Other exemplary receptacles include microdroplets and microfluidic controlled or bulk oil/aqueous emulsions within which assays and related manipulations can occur. Suitable solid supports include, for example, plastics, resins, polysaccharides, silica or silica-based materials, functionalized glass, modified silicon, carbon, metals, inorganic glasses, membranes, nylon, natural fibers (such as, for example, silk, wool and cotton), polymers, and the like. The material coilposing the solid support can include reactive groups such as, for example, carboxy, amino, or hydroxyl groups, which are used for attachment of the capture reagents. Polymeric solid supports can include, e.g., polystyrene, polyethylene glycol tetraphthalate, polyvinyl acetate, polyvinyl chloride, polyvinyl pyrrolidone, polyacrylonitrile, polymethyl methacrylate, polytetrafluoroethylene, butyl rubber, styrenebutadiene rubber, natural rubber, polyethylene, polypropylene, (poly)tetrafluoroethylene, (poly)vinylidenefluoride, polycarbonate, and polymethylpentene. Suitable solid support particles that can be used include, e.g., encoded particles, such as Luminex-type encoded particles, magnetic particles, and glass particles.

Exemplary Uses of Biomarkers

[0116] In various exemplary embodiments, methods are provided for diagnosing mesothelioma in an individual by detecting one or more biomarker values corresponding to one or more biomarkers that are present in the circulation of an individual, such as in serum or plasma, by any number of analytical methods, including any of the analytical methods described herein. These biomarkers are, for example, differentially expressed in individuals with mesothelioma as compared to individuals without mesothelioma. Detection of the differential expression of a biomarker in an individual can be used, for example, to permit the early diagnosis of mesothelioma, to distinguish between a benign and malignant mass (such as, for example, a mass observed on a computed tomography (CT) scan, chest X-ray, MRI or ultrasound), to monitor mesothelioma recurrence, or for differential diagnosis from other clinical conditions such as individuals exposed to asbestos.

[0117] Any of the biomarkers described herein may be used in a variety of clinical indications for mesothelioma, including any of the following: detection of mesothelioma (such as in a high-risk individual or population); characterizing mesothelioma (e.g., determining mesothelioma type, sub-type, or stage), such as by distinguishing between mesothelioma and individuals exposed to asbestos and/or between mesothelioma and adenocarcinoma and other malignant cell types (or otherwise facilitating histopathology); determining whether a pleural abnormality or mass is benign or malignant; determining mesothelioma prognosis; monitoring mesothelioma progression or remission; monitoring for mesothelioma recurrence; monitoring metastasis; treatment selection; monitoring response to a therapeutic agent or other treatment; stratification of individuals for chest CT (e.g., identifying those individuals at greater risk of mesothelioma and thereby most likely to benefit from radiologic screening, thus increasing the positive predictive value of chest CT); combining biomarker testing with additional biomedical information, such as asbestos exposure history, the presence of a genetic marker(s) indicating a higher risk for mesothelioma, etc., or with mass size, morphology, presence of effusion, etc. (such as to provide an assay with increased diagnostic performance compared to other laboratory testing or with mass size, morphology, etc.); facilitating the diagnosis of a pleural abnormality as malignant or benign; facilitating clinical decision making once a pleural abnormality is observed on CT, MRI, PET or US (e.g., ordering repeat radiologic scans if the pleural abnormality is deemed to be low risk, such as if a biomarker-based test is negative, or considering biopsy if the pleural abnormality is deemed medium to high risk, such as if a biomarker-based test is positive, with or without categorization of pleural abnormality or extent of tissue invasion); and facilitating decisions regarding clinical follow-up (e.g., whether to implement repeat radiologic imaging scans, fine needle biopsy, radiation, systemic therapy or surgery after

observing a pleural abnormality on imaging). Biomarker testing may improve positive predictive value (PPV) over CT or chest X-ray screening of high risk individuals alone. In addition to their utilities in conjunction with CT screening, the biomarkers described herein can also be used in conjunction with any other imaging modalities used for mesothelioma, such as chest X-ray, MRI or PET scan. Furthermore, the described biomarkers may also be useful in permitting certain of these uses before indications of mesothelioma are detected by imaging modalities or other clinical correlates, or before symptoms appear. It further includes distinguishing individuals exposed to asbestos from mesothelioma.

[0118] As an example of the manner in which any of the biomarkers described herein can be used to diagnose mesothelioma, differential expression of one or more of the described biomarkers in an individual who is not known to have mesothelioma may indicate that the individual has mesothelioma, thereby enabling detection of mesothelioma at an early stage of the disease when treatment is most effective, perhaps before the mesothelioma is detected by other means or before symptoms appear. Over-expression of one or more of the biomarkers during the course of mesothelioma may be indicative of mesothelioma progression, e.g., a mesothelioma tumor is growing and/or metastasizing (and thus indicate a poor prognosis), whereas a decrease in the degree to which one or more of the biomarkers is differentially expressed (i.e., in subsequent biomarker tests, the expression level in the individual is moving toward or approaching a "normal" expression level) may be indicative of mesothelioma remission, e.g., a mesothelioma tumor is shrinking (and thus indicate a good or better prognosis). Similarly, an increase in the degree to which one or more of the biomarkers is differentially expressed (i.e., in subsequent biomarker tests, the expression level in the individual is moving further away from a "normal" expression level) during the course of mesothelioma treatment may indicate that the mesothelioma is progressing and therefore indicate that the treatment is ineffective, whereas a decrease in differential expression of one or more of the biomarkers during the course of mesothelioma treatment may be indicative of mesothelioma remission and therefore indicate that the treatment is working successfully. Additionally, an increase or decrease in the differential expression of one or more of the biomarkers after an individual has apparently been cured of mesothelioma may be indicative of mesothelioma recurrence. In a situation such as this, for example, the individual can be re-started on therapy (or the therapeutic regimen modified such as to increase dosage amount and/or frequency, if the individual has maintained therapy) at an earlier stage than if the recurrence of mesothelioma was not detected until later. Furthermore, a differential expression level of one or more of the biomarkers in an individual may be predictive of the individual's response to a particular therapeutic agent. In monitoring for mesothelioma recurrence or progression, changes in the biomarker expression levels may indicate the need for repeat imaging, such as to determine mesothelioma activity or to determine the need for changes in treatment.

[0119] Detection of any of the biomarkers described herein may be particularly useful following, or in conjunction with, mesothelioma treatment, such as to evaluate the success of the treatment or to monitor mesothelioma remission, recurrence, and/or progression (including metastasis) following treatment. Mesothelioma treatment may include, for example, administration of a therapeutic agent to the individual, performance of surgery (e.g., surgical resection of at least a portion of a mesothelioma tumor or removal of mesothelioma and surrounding tissue), administration of radiation therapy, or any other type of mesothelioma treatment used in the art, and any combination of these treatments. For example, any of the biomarkers may be detected at least once after treatment or may be detected multiple times after treatment (such as at periodic intervals), or may be detected both before and after treatment. Differential expression levels of any of the biomarkers in an individual over time may be indicative of mesothelioma progression, remission, or recurrence, examples of which include any of the following: an increase or decrease in the expression level of the biomarkers after treatment compared with the expression level of the biomarker before treatment; an increase or decrease in the expression level of the biomarker at a later time point after treatment compared with the expression level of the biomarker at an earlier time point after treatment; and a differential expression level of the biomarker at a single time point after treatment compared with normal levels of the biomarker.

[0120] As a specific example, the biomarker levels for any of the biomarkers described herein can be determined in pre-surgery and post-surgery (e.g., 2-16 weeks after surgery) serum or plasma samples. An increase in the biomarker expression level(s) in the post-surgery sample compared with the pre-surgery sample can indicate progression of mesothelioma (e.g., unsuccessful surgery), whereas a decrease in the biomarker expression level(s) in the post-surgery sample compared with the pre-surgery sample can indicate regression of mesothelioma (e.g., the surgery successfully removed the mesothelium tumor). Similar analyses of the biomarker levels can be carried out before and after other forms of treatment, such as before and after radiation therapy or administration of a therapeutic agent or cancer vaccine.

[0121] In addition to testing biomarker levels as a stand-alone diagnostic test, biomarker levels can also be done in conjunction with determination of SNPs or other genetic lesions or variability that are indicative of increased risk of susceptibility of disease. (See, e.g., Amos et al., Nature Genetics 40, 616-622 (2009)).

[0122] In addition to testing biomarker levels as a stand-alone diagnostic test, biomarker levels can also be done in conjunction with radiologic screening. In addition to testing biomarker levels as a stand-alone diagnostic test, biomarker levels can also be done in conjunction with relevant symptoms or genetic testing. Detection of any of the biomarkers described herein may be useful after a pleural abnormality or mass has been observed through imaging to aid in the diagnosis of mesothelioma and guide appropriate clinical care of the individual, including care by an appropriate surgical

specialist or by palliative therapy in the unresectable patient. In addition to testing biomarker levels in conjunction with relevant symptoms or risk factors, information regarding the biomarkers can also be evaluated in conjunction with other types of data, particularly data that indicates an individual's risk for mesothelioma (e.g., patient clinical history, occupational exposure, symptoms, family history of mesothelioma, history of asbestos exposure, smoking, risk factors such as the presence of a genetic marker(s), and/or status of other biomarkers, etc.). These various data can be assessed by automated methods, such as a computer program/software, which can be embodied in a computer or other apparatus/device.

[0123] In addition to testing biomarker levels in conjunction with radiologic screening in high risk individuals (e.g., assessing biomarker levels in conjunction with size or other characteristics of a pleural abnormality or mass observed on an imaging scan), information regarding the biomarkers can also be evaluated in conjunction with other types of data, particularly data that indicates an individual's risk for mesothelioma (e.g., patient clinical history, occupational exposure history, symptoms, family history of cancer, risk factors such as whether or not the individual was exposed to asbestos, and/or status of other biomarkers, etc.). These various data can be assessed by automated methods, such as a computer program/software, which can be embodied in a computer or other apparatus/device.

[0124] Any of the described biomarkers may also be used in imaging tests. For example, an imaging agent can be coupled to any of the described biomarkers, which can be used to aid in mesothelioma diagnosis, to monitor disease progression/remission or metastasis, to monitor for disease recurrence, or to monitor response to therapy, among other uses.

Detection and Determination of Biomarkers and Biomarker Values

[0125] A biomarker value for the biomarkers described herein can be detected using any of a variety of known analytical methods. In one embodiment, a biomarker value is detected using a capture reagent. As used herein, a "capture agent" or "capture reagent" refers to a molecule that is capable of binding specifically to a biomarker. In various embodiments, the capture reagent can be exposed to the biomarker in solution or can be exposed to the biomarker while the capture reagent is immobilized on a solid support. In other embodiments, the capture reagent contains a feature that is reactive with a secondary feature on a solid support. In these embodiments, the capture reagent can be exposed to the biomarker in solution, and then the feature on the capture reagent can be used in conjunction with the secondary feature on the solid support to immobilize the biomarker on the solid support. The capture reagent is selected based on the type of analysis to be conducted. Capture reagents include but are not limited to aptamers, antibodies, antigens, adnectins, ankyrins, other antibody mimetics and other protein scaffolds, autoantibodies, chimeras, small molecules, an $F(ab')_2$ fragment, a single chain antibody fragment, an Fv fragment, a single chain Fv fragment, a nucleic acid, a lectin, a ligand-binding receptor, affybodies, nanobodies, imprinted polymers, avimers, peptidomimetics, a hormone receptor, a cytokine receptor, and synthetic receptors, and modifications and fragments of these.

[0126] In some embodiments, a biomarker value is detected using a biomarker/capture reagent complex.

[0127] In other embodiments, the biomarker value is derived from the biomarker/capture reagent complex and is detected indirectly, such as, for example, as a result of a reaction that is subsequent to the biomarker/capture reagent interaction, but is dependent on the formation of the biomarker/capture reagent complex.

[0128] In some embodiments, the biomarker value is detected directly from the biomarker in a biological sample.

[0129] In one embodiment, the biomarkers are detected using a multiplexed format that allows for the simultaneous detection of two or more biomarkers in a biological sample. In one embodiment of the multiplexed format, capture reagents are immobilized, directly or indirectly, covalently or non-covalently, in discrete locations on a solid support. In another embodiment, a multiplexed format uses discrete solid supports where each solid support has a unique capture reagent associated with that solid support, such as, for example quantum dots. In another embodiment, an individual device is used for the detection of each one of multiple biomarkers to be detected in a biological sample. Individual devices can be configured to permit each biomarker in the biological sample to be processed simultaneously. For example, a microtiter plate can be used such that each well in the plate is used to uniquely analyze one of multiple biomarkers to be detected in a biological sample.

[0130] In one or more of the foregoing embodiments, a fluorescent tag can be used to label a component of the biomarker/capture complex to enable the detection of the biomarker value. In various embodiments, the fluorescent label can be conjugated to a capture reagent specific to any of the biomarkers described herein using known techniques, and the fluorescent label can then be used to detect the corresponding biomarker value. Suitable fluorescent labels include rare earth chelates, fluorescein and its derivatives, rhodamine and its derivatives, dansyl, allophycocyanin, PBXL-3, Qdot 605, Lissamine, phycoerythrin, Texas Red, and other such compounds.

[0131] In one embodiment, the fluorescent label is a fluorescent dye molecule. In some embodiments, the fluorescent dye molecule includes at least one substituted indolium ring system in which the substituent on the 3-carbon of the indolium ring contains a chemically reactive group or a conjugated substance. In some embodiments, the dye molecule includes an AlexFluor molecule, such as, for example, AlexaFluor 488, AlexaFluor 532, AlexaFluor 647, AlexaFluor 680,

or AlexaFluor 700. In other embodiments, the dye molecule includes a first type and a second type of dye molecule, such as, e.g., two different AlexaFluor molecules. In other embodiments, the dye molecule includes a first type and a second type of dye molecule, and the two dye molecules have different emission spectra.

[0132] Fluorescence can be measured with a variety of instrumentation compatible with a wide range of assay formats. For example, spectrofluorimeters have been designed to analyze microtiter plates, microscope slides, printed arrays, cuvettes, etc. See Principles of Fluorescence Spectroscopy, by J. R. Lakowicz, Springer Science + Business Media, Inc., 2004. See Bioluminescence & Chemiluminescence: Progress & Current Applications; Philip E. Stanley and Larry J. Kricka editors, World Scientific Publishing Company, January 2002.

[0133] In one or more of the foregoing embodiments, a chemiluminescence tag can optionally be used to label a component of the biomarker/capture complex to enable the detection of a biomarker value. Suitable chemiluminescent materials include any of oxalyl chloride, Rodamin 6G, $Ru(bipy)_3^{2+}$, TMAE (tetrakis(dimethylamino)ethylene), Pyrogallol (1,2,3-trihydroxibenzene), Lucigenin, peroxyoxalates, Aryl oxalates, Acridinium esters, dioxetanes, and others.

[0134] In yet other embodiments, the detection method includes an enzyme/substrate combination that generates a detectable signal that corresponds to the biomarker value. Generally, the enzyme catalyzes a chemical alteration of the chromogenic substrate which can be measured using various techniques, including spectrophotometry, fluorescence, and chemiluminescence. Suitable enzymes include, for example, luciferases, luciferin, malate dehydrogenase, urease, horseradish peroxidase (HRPO), alkaline phosphatase, beta-galactosidase, glucoamylase, lysozyme, glucose oxidase, galactose oxidase, and glucose-6-phosphate dehydrogenase, uricase, xanthine oxidase, lactoperoxidase, microperoxidase, and the like.

[0135] In yet other embodiments, the detection method can be a combination of fluorescence, chemiluminescence, radionuclide or enzyme/substrate combinations that generate a measurable signal. Multimodal signaling could have unique and advantageous characteristics in biomarker assay formats.

[0136] More specifically, the biomarker values for the biomarkers described herein can be detected using known analytical methods including, singleplex aptamer assays, multiplexed aptamer assays, singleplex or multiplexed immunoassays, mRNA expression profiling, miRNA expression profiling, mass spectrometric analysis, histological/cytological methods, etc. as detailed below.

Determination of Biomarker Values using Aptamer-Based Assays

[0137] Assays directed to the detection and quantification of physiologically significant molecules in biological samples and other samples are important tools in scientific research and in the health care field. One class of such assays involves the use of a microarray that includes one or more aptamers immobilized on a solid support. The aptamers are each capable of binding to a target molecule in a highly specific manner and with very high affinity. See, e.g., U. S. Patent No. 5,475,096 entitled "Nucleic Acid Ligands"; see also, e.g., U. S. Patent No. 6,242,246, U. S. Patent No. 6,458,543, and U. S. Patent No. 6,503,715, each of which is entitled "Nucleic Acid Ligand Diagnostic Biochip". Once the microarray is contacted with a sample, the aptamers bind to their respective target molecules present in the sample and thereby enable a determination of a biomarker value corresponding to a biomarker.

[0138] As used herein, an "aptamer" refers to a nucleic acid that has a specific binding affinity for a target molecule. It is recognized that affinity interactions are a matter of degree; however, in this context, the "specific binding affinity" of an aptamer for its target means that the aptamer binds to its target generally with a much higher degree of affinity than it binds to other components in a test sample. An "aptamer" is a set of copies of one type or species of nucleic acid molecule that has a particular nucleotide sequence. An aptamer can include any suitable number of nucleotides, including any number of chemically Modified nucleotides. "Aptamers" refers to more than one such set of molecules. Different aptamers can have either the same or different numbers of nucleotides. Aptamers can be DNA or RNA or chemically modified nucleic acids and can be single stranded, double stranded, or contain double stranded regions, and can include higher ordered structures. An aptamer can also be a photoaptamer, where a photoreactive or chemically reactive functional group is included in the aptamer to allow it to be covalently linked to its corresponding target. Any of the aptamer methods disclosed herein can include the use of two or more aptamers that specifically bind the same target molecule. As further described below, an aptamer may include a tag. If an aptamer includes a tag, all copies of the aptamer need not have the same tag. Moreover, if different aptamers each include a tag, these different aptamers can have either the same tag or a different tag.

[0139] An aptamer can be identified using any known method, including the SELEX process. Once identified, an aptamer can be prepared or synthesized in accordance with any known method, including chemical synthetic methods and enzymatic synthetic methods.

[0140] As used herein, a "SOMAmer" or Slow Off-Rate Modified Aptamer refers to an aptamer having improved off-rate characteristics. SOMAmers can be generated using the improved SELEX methods described in U. S. Publication No. 2009/0004667, entitled "Method for Generating Aptamers with Improved Off-Rates."

[0141] The terms "SELEX" and "SELEX process" are used interchangeably herein to refer generally to a combination

of (1) the selection of aptamers that interact with a target molecule in a desirable manner, for example binding with high affinity to a protein, with (2) the amplification of those selected nucleic acids. The SELEX process can be used to identify aptamers with high affinity to a specific target or biomarker.

[0142] SELEX generally includes preparing a candidate mixture of nucleic acids, binding of the candidate mixture to the desired target molecule to form an affinity complex, separating the affinity complexes from the unbound candidate nucleic acids, separating and isolating the nucleic acid from the affinity complex, purifying the nucleic acid, and identifying a specific aptamer sequence. The process may include multiple rounds to further refine the affinity of the selected aptamer. The process can include amplification steps at one or more points in the process. See, e.g., U. S. Patent No. 5,475,096, entitled "Nucleic Acid Ligands". The SELEX process can be used to generate an aptamer that covalently binds its target as well as an aptamer that non-covalently binds its target. See, e.g., U. S. Patent No. 5, 105,331 entitled "Systematic Evolution of Nucleic Acid Ligands by Exponential Enrichment: Chemi-SELEX."

[0143] The SELEX process can be used to identify high-affinity aptamers containing modified nucleotides that confer improved characteristics on the aptamer, such as, for example, improved in vivo stability or improved delivery characteristics. Examples of such modifications include chemical substitutions at the ribose and/or phosphate and/or base positions. SELEX process-identified aptamers containing modified nucleotides are described in U. S. Patent No. 5,660,985, entitled "High Affinity Nucleic Acid Ligands Containing Modified Nucleotides", which describes oligonucleotides containing nucleotide derivatives chemically modified at the 5'- and 2'-positions of pyrimidines. U. S. Patent No. 5,580,737, see supra, describes highly specific aptamers containing one or more nucleotides modified with 2'-amino (2'-NH_2), 2'-fluoro (2'-F), and/or 2'-O-methyl (2'-OMe). See also, U. S. Patent Application Publication 2009/0098549, entitled "SELEX and PHOTOSELEX", which describes nucleic acid libraries having expanded physical and chemical properties and their use in SELEX and photoSELEX.

[0144] SELEX can also be used to identify aptamers that have desirable off-rate characteristics. See U. S. Patent Application Publication 2009/0004667, entitled "Method for Generating Aptamers with Improved Off-Rates", which describes improved SELEX methods for generating aptamers that can bind to target molecules. Methods for producing aptamers and photoaptamers having slower rates of dissociation from their respective target molecules are described. The methods involve contacting the candidate mixture with the target molecule, allowing the formation of nucleic acid-target complexes to occur, and performing a slow off-rate enrichment process wherein nucleic acid-target complexes with fast dissociation rates will dissociate and not reform, while complexes with slow dissociation rates will remain intact. Additionally, the methods include the use of modified nucleotides in the production of candidate nucleic acid mixtures to generate aptamers with improved off-rate performance.

[0145] A variation of this assay employs aptamers that include photoreactive functional groups that enable the aptamers to covalently bind or "photocrosslink" their target molecules. See, e.g., U. S. Patent No. 6,544,776 entitled "Nucleic Acid Ligand Diagnostic Biochip". These photoreactive aptamers are also referred to as photoaptamers. See, e.g., U. S. Patent No. 5,763,177, U. S. Patent No. 6,001,577, and U. S. Patent No. 6,291,184, each of which is entitled "Systematic Evolution of Nucleic Acid Ligands by Exponential Enrichment: Photoselection of Nucleic Acid Ligands and Solution SELEX"; see also, e.g., U. S. Patent No. 6,458,539, entitled "Photoselection of Nucleic Acid Ligands". After the microarray is contacted with the sample and the photoaptamers have had an opportunity to bind to their target molecules, the photoaptamers are photoactivated, and the solid support is washed to remove any non-specifically bound molecules. Harsh wash conditions may be used, since target molecules that are bound to the photoaptamers are generally not removed, due to the covalent bonds created by the photoactivated functional group(s) on the photoaptamers. In this manner, the assay enables the detection of a biomarker value corresponding to a biomarker in the test sample.

[0146] In both of these assay formats, the aptamers are immobilized on the solid support prior to being contacted with the sample. Under certain circumstances, however, immobilization of the aptamers prior to contact with the sample may not provide an optimal assay. For example, pre-immobilization of the aptamers may result in inefficient mixing of the aptamers with the target molecules on the surface of the solid support, perhaps leading to lengthy reaction times and, therefore, extended incubation periods to permit efficient binding of the aptamers to their target molecules. Further, when photoaptamers are employed in the assay and depending upon the material utilized as a solid support, the solid support may tend to scatter or absorb the light used to effect the formation of covalent bonds between the photoaptamers and their target molecules. Moreover, depending upon the method employed, detection of target molecules bound to their aptamers can be subject to imprecision, since the surface of the solid support may also be exposed to and affected by any labeling agents that are used. Finally, immobilization of the aptamers on the solid support generally involves an aptamer-preparation step (i.e., the immobilization) prior to exposure of the aptamers to the sample, and this preparation step may affect the activity or functionality of the aptamers.

[0147] Aptamer assays that permit an aptamer to capture its target in solution and then employ separation steps that are designed to remove specific components of the aptamertarget mixture prior to detection have also been described (see U. S. Patent Application Publication 2009/0042206, entitled "Multiplexed Analyses of Test Samples"). The described aptamer assay methods enable the detection and quantification of a non-nucleic acid target (e.g., a protein target) in a test sample by detecting and quantifying a nucleic acid (i.e., an aptamer). The described methods create a nucleic acid

surrogate (i.e, the aptamer) for detecting and quantifying a non-nucleic acid target, thus allowing the wide variety of nucleic acid technologies, including amplification, to be applied to a broader range of desired targets, including protein targets.

[0148] Aptamers can be constructed to facilitate the separation of the assay components from an aptamer biomarker complex (or photoaptamer biomarker covalent complex) and permit isolation of the aptamer for detection and/or quantification. In one embodiment, these constructs can include a cleavable or releasable element within the aptamer sequence. In other embodiments, additional functionality can be introduced into the aptamer, for example, a labeled or detectable component, a spacer component, or a specific binding tag or immobilization element. For example, the aptamer can include a tag connected to the aptamer via a cleavable moiety, a label, a spacer component separating the label, and the cleavable moiety. In one embodiment, a cleavable element is a photocleavable linker. The photocleavable linker can be attached to a biotin moiety and a spacer section, can include an NHS group for derivatization of amines, and can be used to introduce a biotin group to an aptamer, thereby allowing for the release of the aptamer later in an assay method.

[0149] Homogenous assays, done with all assay components in solution, do not require separation of sample and reagents prior to the detection of signal. These methods are rapid and easy to use. These methods generate signal based on a molecular capture or binding reagent that reacts with its specific target. For mesothelioma, the molecular capture reagents would be an aptamer or an antibody or the like and the specific target would be a mesothelioma biomarker of Table 1.

[0150] In one embodiment, a method for signal generation takes advantage of anisotropy signal change due to the interaction of a fluorophore-labeled capture reagent with its specific biomarker target. When the labeled capture reacts with its target, the increased molecular weight causes the rotational motion of the fluorophore attached to the complex to become much slower changing the anisotropy value. By monitoring the anisotropy change, binding events may be used to quantitatively measure the biomarkers in solutions. Other methods include fluorescence polarization assays, molecular beacon methods, time resolved fluorescence quenching, chemiluminescence, fluorescence resonance energy transfer, and the like.

[0151] An exemplary solution-based aptamer assay that can be used to detect a biomarker value corresponding to a biomarker in a biological sample includes the following: (a) preparing a mixture by contacting the biological sample with an aptamer that includes a first tag and has a specific affinity for the biomarker, wherein an aptamer affinity complex is formed when the biomarker is present in the sample; (b) exposing the mixture to a first solid support including a first capture element, and allowing the first tag to associate with the first capture element; (c) removing any components of the mixture not associated with the first solid support; (d) attaching a second tag to the biomarker component of the aptamer affinity complex; (e) releasing the aptamer affinity complex from the first solid support; (f) exposing the released aptamer affinity complex to a second solid support that includes a second capture element and allowing the second tag to associate with the second capture element; (g) removing any non-complexed aptamer from the mixture by partitioning the non-complexed aptamer from the aptamer affinity complex; (h) eluting the aptamer from the solid support; and (i) detecting the biomarker by detecting the aptamer component of the aptamer affinity complex.

[0152] Any means known in the art can be used to detect a biomarker value by detecting the aptamer component of an aptamer affinity complex. A number of different detection methods can be used to detect the aptamer component of an affinity complex, such as, for example, hybridization assays, mass spectroscopy, or QPCR. In some embodiments, nucleic acid sequencing methods can be used to detect the aptamer component of an aptamer affinity complex and thereby detect a biomarker value. Briefly, a test sample can be subjected to any kind of nucleic acid sequencing method to identify and quantify the sequence or sequences of one or more aptamers present in the test sample. In some embodiments, the sequence includes the entire aptamer molecule or any portion of the molecule that may be used to uniquely identify the molecule. In other embodiments, the identifying sequencing is a specific sequence added to the aptamer; such sequences are often referred to as "tags," "barcodes," or "zipcodes." In some embodiments, the sequencing method includes enzymatic steps to amplify the aptamer sequence or to convert any kind of nucleic acid, including RNA and DNA that contain chemical modifications to any position, to any other kind of nucleic acid appropriate for sequencing.

[0153] In some embodiments, the sequencing method includes one or more cloning steps. In other embodiments the sequencing method includes a direct sequencing method without cloning.

[0154] In some embodiments, the sequencing method includes a directed approach with specific primers that target one or more aptamers in the test sample. In other embodiments, the sequencing method includes a shotgun approach that targets all aptamers in the test sample.

[0155] In some embodiments, the sequencing method includes enzymatic steps to amplify the molecule targeted for sequencing. In other embodiments, the sequencing method directly sequences single molecules. An exemplary nucleic acid sequencing-based method that can be used to detect a biomarker value corresponding to a biomarker in a biological sample includes the following: (a) converting a mixture of aptamers that contain chemically modified nucleotides to unmodified nucleic acids with an enzymatic step; (b) shotgun sequencing the resulting unmodified nucleic acids with a

massively parallel sequencing platform such as, for example, the 454 Sequencing System (454 Life Sciences/Roche), the Illumina Sequencing System (Illumina), the ABI SOLiD Sequencing System (Applied Biosystems), the HeliScope Single Molecule Sequencer (Helicos Biosciences), or the Pacific Biosciences Real Time Single-Molecule Sequencing System (Pacific BioSciences) or the Polonator G Sequencing System (Dover Systems); and (c) identifying and quantifying the aptamers present in the mixture by specific sequence and sequence count.

Determination of Biomarker Values using Immunoassays

[0156] Immunoassay methods are based on the reaction of an antibody to its corresponding target or analyte and can detect the analyte in a sample depending on the specific assay format. To improve specificity and sensitivity of an assay method based on immuno-reactivity, monoclonal antibodies are often used because of their specific epitope recognition. Polyclonal antibodies have also been successfully used in various immunoassays because of their increased affinity for the target as compared to monoclonal antibodies. Immunoassays have been designed for use with a wide range of biological sample matrices. Immunoassay formats have been designed to provide qualitative, semi-quantitative, and quantitative results.

[0157] Quantitative results are generated through the use of a standard curve created with known concentrations of the specific analyte to be detected. The response or signal from an unknown sample is plotted onto the standard curve, and a quantity or value corresponding to the target in the unknown sample is established.

[0158] Numerous immunoassay formats have been designed. ELISA or EIA can be quantitative for the detection of an analyte. This method relies on attachment of a label to either the analyte or the antibody and the label component includes, either directly or indirectly, an enzyme. ELISA tests may be formatted for direct, indirect, competitive, or sandwich detection of the analyte. Other methods rely on labels such as, for example, radioisotopes (I125) or fluorescence. Additional techniques include, for example, agglutination, nephelometry, turbidimetry, Western blot, immunoprecipitation, immunocytochemistry, immunohistochemistry, flow cytometry, Luminex assay, and others (see ImmunoAssay: A Practical Guide, edited by Brian Law, published by Taylor & Francis, Ltd., 2005 edition).

[0159] Exemplary assay formats include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay, fluorescent, chemiluminescence, and fluorescence resonance energy transfer (FRET) or time resolved-FRET (TR-FRET) immunoassays. Examples of procedures for detecting biomarkers include biomarker immunoprecipitation followed by quantitative methods that allow size and peptide level discrimination, such as gel electrophoresis, capillary electrophoresis, planar electrochromatography, and the like.

[0160] Methods of detecting and/or quantifying a detectable label or signal generating material depend on the nature of the label. The products of reactions catalyzed by appropriate enzymes (where the detectable label is an enzyme; see above) can be, without limitation, fluorescent, luminescent, or radioactive or they may absorb visible or ultraviolet light. Examples of detectors suitable for detecting such detectable labels include, without limitation, x-ray film, radioactivity counters, scintillation counters, spectrophotometers, colorimeters, fluorometers, luminometers, and densitometers.

[0161] Any of the methods for detection can be performed in any format that allows for any suitable preparation, processing, and analysis of the reactions. This can be, for example, in multi-well assay plates (e.g., 96 wells or 384 wells) or using any suitable array or microarray. Stock solutions for various agents can be made manually or robotically, and all subsequent pipetting, diluting, mixing, distribution, washing, incubating, sample readout, data collection and analysis can be done robotically using commercially available analysis software, robotics, and detection instrumentation capable of detecting a detectable label.

Determination of Biomarker Values using Gene Expression Profiling

[0162] Measuring mRNA in a biological sample may be used as a surrogate for detection of the level of the corresponding protein in the biological sample. Thus, any of the biomarkers or biomarker panels described herein can also be detected by detecting the appropriate RNA.

[0163] mRNA expression levels are measured by reverse transcription quantitative polymerase chain reaction (RT-PCR followed with qPCR). RT-PCR is used to create a cDNA from the InRNA. The cDNA may be used in a qPCR assay to produce fluorescence as the DNA amplification process progresses. By comparison to a standard curve, qPCR can produce an absolute measurement such as number of copies of mRNA per cell. Northern blots, microarrays, Invader assays, and RT-PCR combined with capillary electrophoresis have all been used to measure expression levels of mRNA in a sample. See Gene Expression Profiling: Methods and Protocols, Richard A. Shimkets, editor, Humana Press, 2004.

[0164] miRNA molecules are small RNAs that are non-coding but may regulate gene expression. Any of the methods suited to the measurement of mRNA expression levels can also be used for the corresponding miRNA. Recently many laboratories have investigated the use of miRNAs as biomarkers for disease. Many diseases involve wide-spread transcriptional regulation, and it is not surprising that miRNAs might find a role as biomarkers. The connection between miRNA concentrations and disease is often even less clear than the connections between protein levels and disease,

yet the value of miRNA biomarkers might be substantial. Of course, as with any RNA expressed differentially during disease, the problems facing the development of an in vitro diagnostic product will include the requirement that the miRNAs survive in the diseased cell and are easily extracted for analysis, or that the miRNAs are released into blood or other matrices where they must survive long enough to be measured. Protein biomarkers have similar requirements, although many potential protein biomarkers are secreted intentionally at the site of pathology and function, during disease, in a paracrine fashion. Many potential protein biomarkers are designed to function outside the cells within which those proteins are synthesized.

Detection of Biomarkers Using In Vivo Molecular Imaging Technologies

[0165] Any of the described biomarkers (see Table 1) may also be used in molecular imaging tests. For example, an imaging agent can be coupled to any of the described biomarkers, which can be used to aid in mesothelioma diagnosis, to monitor disease progression/remission or metastasis, to monitor for disease recurrence, or to monitor response to therapy, among other uses.

[0166] In vivo imaging technologies provide non-invasive methods for determining the state of a particular disease in the body of an individual. For example, entire portions of the body, or even the entire body, may be viewed as a three dimensional image, thereby providing valuable information concerning morphology and structures in the body. Such technologies may be combined with the detection of the biomarkers described herein to provide information concerning the cancer status, in particular the mesothelioma status, of an individual.

[0167] The use of in vivo molecular imaging technologies is expanding due to various advances in technology. These advances include the development of new contrast agents or labels, such as radiolabels and/or fluorescent labels, which can provide strong signals within the body; and the development of powerful new imaging technology, which can detect and analyze these signals from outside the body, with sufficient sensitivity and accuracy to provide useful information. The contrast agent can be visualized in an appropriate imaging system, thereby providing an image of the portion or portions of the body in which the contrast agent is located. The contrast agent may be bound to or associated with a capture reagent, such as an aptamer or an antibody, for example, and/or with a peptide or protein, or an oligonucleotide (for example, for the detection of gene expression), or a complex containing any of these with one or more macromolecules and/or other particulate forms.

[0168] The contrast agent may also feature a radioactive atom that is useful in imaging. Suitable radioactive atoms include technetium-99m or iodine-123 for scintigraphic studies. Other readily detectable moieties include, for example, spin labels for magnetic resonance imaging (MRI) such as, for example, iodine-123 again, iodine-131, indium-111, fluorine-19, carbon-13, nitrogen-15, oxygen-17, gadolinium, manganese or iron. Such labels are well known in the art and could easily be selected by one of ordinary skill in the art.

[0169] Standard imaging techniques include but are not limited to magnetic resonance imaging, computed tomography scanning, positron emission tomography (PET), single photon emission computed tomography (SPECT), and the like. For diagnostic in vivo imaging, the type of detection instrument available is a major factor in selecting a given contrast agent, such as a given radionuclide and the particular biomarker that it is used to target (protein, mRNA, and the like). The radionuclide chosen typically has a type of decay that is detectable by a given type of instrument. Also, when selecting a radionuclide for in vivo diagnosis, its half-life should be long enough to enable detection at the time of maximum uptake by the target tissue but short enough that deleterious radiation of the host is minimized.

[0170] Exemplary imaging techniques include but are not limited to PET and SPECT, which are imaging techniques in which a radionuclide is synthetically or locally administered to an individual. The subsequent uptake of the radiotracer is measured over time and used to obtain information about the targeted tissue and the biomarker. Because of the high-energy (gamma-ray) emissions of the specific isotopes employed and the sensitivity and sophistication of the instruments used to detect them, the two-dimensional distribution of radioactivity may be inferred from outside of the body.

[0171] Commonly used positron-emitting nuclides in PET include, for example, carbon-11, nitrogen-13, oxygen-15, and fluorine-18. Isotopes that decay by electron capture and/or gamma-emission are used in SPECT and include, for example iodine-123 and technetium-99m. An exemplary method for labeling amino acids with technetium-99m is the reduction of pertechnetate ion in the presence of a chelating precursor to form the labile technetium-99m-precursor complex, which, in turn, reacts with the metal binding group of a bifunctionally modified chemotactic peptide to form a technetium-99m-chemotactic peptide conjugate.

[0172] Antibodies are frequently used for such in vivo imaging diagnostic methods. The preparation and use of antibodies for in vivo diagnosis is well known in the art. Labeled antibodies which specifically bind any of the biomarkers in Table 1 can be injected into an individual suspected of having a certain type of cancer (e.g., mesothelioma), detectable according to the particular biomarker used, for the purpose of diagnosing or evaluating the disease status of the individual. The label used will be selected in accordance with the imaging modality to be used, as previously described. Localization of the label permits determination of the spread of the cancer. The amount of label within an organ or tissue also allows determination of the presence or absence of cancer in that organ or tissue.

[0173] Similarly, aptamers may be used for such in vivo imaging diagnostic methods. For example, an aptamer that was used to identify a particular biomarker described in Table 1 (and therefore binds specifically to that particular biomarker) may be appropriately labeled and injected into an individual suspected of having mesothelioma, detectable according to the particular biomarker, for the purpose of diagnosing or evaluating the mesothelioma status of the individual. The label used will be selected in accordance with the imaging modality to be used, as previously described. Localization of the label permits determination of the spread of the cancer. The amount of label within an organ or tissue also allows determination of the presence or absence of cancer in that organ or tissue. Aptamer-directed imaging agents could have unique and advantageous characteristics relating to tissue penetration, tissue distribution, kinetics, elimination, potency, and selectivity as compared to other imaging agents.

[0174] Such techniques may also optionally be performed with labeled oligonucleotides, for example, for detection of gene expression through imaging with antisense oligonucleotides. These methods are used for in situ hybridization, for example, with fluorescent molecules or radionuclides as the label. Other methods for detection of gene expression include, for example, detection of the activity of a reporter gene.

[0175] Another general type of imaging technology is optical imaging, in which fluorescent signals within the subject are detected by an optical device that is external to the subject. These signals may be due to actual fluorescence and/or to bioluminescence. Improvements in the sensitivity of optical detection devices have increased the usefulness of optical imaging for in vivo diagnostic assays.

[0176] The use of in vivo molecular biomarker imaging is increasing, including for clinical trials, for example, to more rapidly measure clinical efficacy in trials for new cancer therapies and/or to avoid prolonged treatment with a placebo for those diseases, in which such prolonged treatment may be considered to be ethically questionable.

[0177] For a review of other techniques, see N. Blow, Nature Methods, 6, 465-469, 2009.

Determination of Biomarker Values using Histology/Cytology Methods

[0178] For evaluation of mesothelioma, a variety of tissue samples may be used in histological or cytological methods. Sample selection depends on the primary tumor location and sites of metastases. For example, tissue or effusion samples (forceps biopsy, fine needle aspiration (FNA), and/or brush cytology) collected at the time of CT or US-guided FNA can be used for histology. Ascites or peritoneal washings, pleural effusions or mesothelium fluid can be used for cyotology. Any of the biomarkers identified herein that were shown to be up-regulated (Table 1) in the individuals with pleural abnormality can be used to stain a histological specimen as an indication of disease.

[0179] In one embodiment, one or more capture reagents specific to the corresponding biomarker(s) are used in a cytological evaluation of a mesothelium cell sample and may include one or more of the following: collecting a cell sample, fixing the cell sample, dehydrating, clearing, immobilizing the cell sample on a microscope slide, permeabilizing the cell sample, treating for analyte retrieval, staining, destaining, washing, blocking, and reacting with one or more capture reagent/s in a buffered solution. In another embodiment, the cell sample is produced from a cell block.

[0180] In another embodiment, one or more capture reagent(s) specific to the corresponding biomarker(s) are used in a histological evaluation of a peritoneal or pleural abnormality tissue sample and may include one or more of the following: collecting a tissue specimen, fix-ing the tissue sample, dehydrating, clearing, immobilizing the tissue sample on a microscope slide, permeabilizing the tissue sample, treating for analyte retrieval, staining, destaining, washing, blocking, dehydrating, and reacting with capture reagent(s) in a buffered solution. In another embodiment, fixing and dehydrating are replaced with freezing.

[0181] In another embodiment, the one or more aptamer(s) specific to the corresponding biomarker(s) are reacted with the histological or cytological sample and can serve as the nu-cleic acid target in a nucleic acid amplification method. Suitable nucleic acid amplification methods include, for example, PCR, q-beta replicase, rolling circle amplification, strand dis-placement, helicase dependent amplification, loop mediated isothermal amplification, ligase chain reaction, and restriction and circularization aided rolling circle amplification.

[0182] In one embodiment, the one or more capture reagent(s) specific to the corresponding biomarkers for use in the histological or cytological evaluation are mixed in a buffered solution that can include any of the following: blocking materials, competitors, detergents, stabilizers, carrier nucleic acid, polyanionic materials, etc.

[0183] A "cytology protocol" generally includes sample collection, sample fixation, sample immobilization, and staining. "Cell preparation" can include several processing steps after sample collection, including the use of one or more slow off-rate aptamers for the staining of the prepared cells.

[0184] Sample collection can include directly placing the sample in an untreated transport container, placing the sample in a transport container containing some type of media, or placing the sample directly onto a slide (immobilization) without any treatment or fixation.

[0185] Sample immobilization can be improved by applying a portion of the collected specimen to a glass slide that is treated with polylysine, gelatin, or a silane. Slides can be prepared by smearing a thin and even layer of cells across the slide. Care is generally taken to minimize mechanical distortion and drying artifacts. Liquid specimens can be proc-

essed in a cell block method. Or, alternatively, liquid specimens can be mixed 1:1 with the fixative solution for about 10 minutes at room temperature.

[0186] Cell blocks can be prepared from residual effusions, sputum, urine sediments, gastrointestinal fluids, cell scraping, or fine needle aspirates. Cells are concentrated or packed by centrifugation or membrane filtration. A number of methods for cell block preparation have been developed. Representative procedures include the fixed sediment, bacterial agar, or membrane filtration methods. In the fixed sediment method, the cell sediment is mixed with a fixative like Bouins, picric acid, or buffered formalin and then the mixture is centrifuged to pellet the fixed cells. The supernatant is removed, drying the cell pellet as completely as possible. The pellet is collected and wrapped in lens paper and then placed in a tissue cassette. The tissue cassette is placed in a jar with additional fixative and processed as a tissue sample. Agar method is very similar but the pellet is removed and dried on paper towel and then cut in half. The cut side is placed in a drop of melted agar on a glass slide and then the pellet is covered with agar making sure that no bubbles form in the agar. The agar is allowed to harden and then any excess agar is trimmed away. This is placed in a tissue cassette and the tissue process completed. Alternatively, the pellet may be directly suspended in 2% liquid agar at 65°C and the sample centrifuged. The agar cell pellet is allowed to solidify for an hour at 4°C. The solid agar may be removed from the centrifuge tube and sliced in half. The agar is wrapped in filter paper and then the tissue cassette. Processing from this point forward is as described above. Centrifugation can be replaced in any these procedures with membrane filtration. Any of these processes may be used to generate a "cell block sample".

[0187] Cell blocks can be prepared using specialized resin including Lowicryl resins, LR White, LR Gold, Unicryl, and MonoStep. These resins have low viscosity and can be poly-merized at low temperatures and with ultra violet (UV) light. The embedding process relies on progressively cooling the sample during dehydration, transferring the sample to the resin, and polymerizing a block at the final low temperature at the appropriate UV wavelength.

[0188] Cell block sections can be stained with hematoxylin-eosin for cytomorphological examination while additional sections are used for examination for specific markers.

[0189] Whether the process is cytologoical or histological, the sample may be fixed prior to additional processing to prevent sample degradation. This process is called "fixation" and describes a wide range of materials and procedures that may be used interchangeably. The sample fixation protocol and reagents are best selected empirically based on the targets to be detected and the specific cell/tissue type to be analyzed. Sample fixation relies on reagents such as ethanol, polyethylene glycol, methanol, formalin, or isopropanol. The samples should be fixed as soon after collection and affixation to the slide as possible. However, the fixative selected can introduce structural changes into various molecular targets making their subse-quent detection more difficult. The fixation and immobilization processes and their sequence can modify the appearance of the cell and these changes must be anticipated and recognized by the cytotechnologist. Fixatives can cause shrinkage of certain cell types and cause the cytoplasm to appear granular or reticular. Many fixative function by crosslinking cellular components. This can damage or modify specific epitopes, generate new epitopes, cause molecular associations, and reduce membrane permeability. Formalin fixation is one of the most common cytological/histological approaches. Formalin forms methyl bridges between neighboring proteins or within proteins. Precipitation or coagulation is also used for fixation and ethanol is frequently used in this type of fixation. A combination of crosslinking and precipitation can also be used for fixation. A strong fixation process is best at preserving morphological information while a weaker fixation process is best for the preservation of molecular targets.

[0190] A representative fixative is 50% absolute ethanol, 2 mM polyethylene glycol (PEG), 1.85% formaldehyde. Variations on this formulation include ethanol (50% to 95%), methanol (20% - 50%), and formalin (formaldehyde) only. Another common fixative is 2% PEG 1500, 50% ethanol, and 3% methanol. Slides are place in the fixative for about 10 to 15 minutes at room temperature and then removed and allowed to dry. Once slides are fixed they can be rinsed with a buffered solution like PBS.

[0191] A wide range of dyes can be used to differentially highlight and contrast or "stain" cellular, sub-cellular, and tissue features or morphological structures. Hematoylin is used to stain nuclei a blue or black color. Orange G-6 and Eosin Azure both stain the cell's cytoplasm. Orange G stains keratin and glycogen containing cells yellow. Eosin Y is used to stain nucleoli, cilia, red blood cells, and superficial epithelial squamous cells. Romanowsky stains are used for air dried slides and are useful in enhancing pleomorphism and distinguishing extracellular from intracytoplasmic material.

[0192] The staining process can include a treatment to increase the permeability of the cells to the stain. Treatment of the cells with a detergent can be used to increase permeability. To increase cell and tissue permeability, fixed samples can be further treated with solvents, saponins, or non-ionic detergents. Enzymatic digestion can also improve the accessibility of specific targets in a tissue sample.

[0193] After staining, the sample is dehydrated using a succession of alcohol rinses with increasing alcohol concentration. The final wash is done with xylene or a xylene substitute, such as a citrus terpene, that has a refractive index close to that of the coverslip to be applied to the slide. This final step is referred to as clearing. Once the sample is dehydrated and cleared, a mounting medium is applied. The mounting medium is selected to have a refractive index close to the glass and is capable of bonding the coverslip to the slide. It will also inhibit the additional drying, shrinking, or fading of the cell sample.

[0194] Regardless of the stains or processing used, the final evaluation of the mesothelium cytological specimen is made by some type of microscopy to permit a visual inspection of the morphology and a determination of the marker's presence or absence. Exemplary microscopic methods include brightfield, phase contrast, fluorescence, and differential interference contrast.

[0195] If secondary tests are required on the sample after examination, the coverslip may be removed and the slide destained. Destaining involves using the original solvent systems used in staining the slide originally without the added dye and in a reverse order to the original staining procedure. Destaining may also be completed by soaking the slide in an acid alcohol until the cells are colorless. Once colorless the slides are rinsed well in a water bath and the second staining procedure applied.

[0196] In addition, specific molecular differentiation may be possible in conjunction with the cellular morphological analysis through the use of specific molecular reagents such as antibodies or nucleic acid probes or aptamers. This improves the accuracy of diagnostic cytology. Micro-dissection can be used to isolate a subset of cells for additional evaluation, in particular, for genetic evaluation of abnormal chromosomes, gene expression, or mutations.

[0197] Preparation of a tissue sample for histological evaluation involves fixation, dehy-dration, infiltration, embedding, and sectioning. The fixation reagents used in histology are very similar or identical to those used in cytology and have the same issues of preserving morphological features at the expense of molecular ones such as individual proteins. Time can be saved if the tissue sample is not fixed and dehydrated but instead is frozen and then sectioned while frozen. This is a more gentle processing procedure and can preserve more individual markers. However, freezing is not acceptable for long term storage of a tissue sample as subcellular information is lost due to the introduction of ice crystals. Ice in the frozen tissue sample also prevents the sectioning process from producing a very thin slice and thus some microscopic resolution and imaging of subcellular structures can be lost. In addition to formalin fixation, osmium tetroxide is used to fix and stain phospholipids (membranes).

[0198] Dehydration of tissues is accomplished with successive washes of increasing alcohol concentration. Clearing employs a material that is miscible with alcohol and the embedding material and involves a stepwise process starting at 50:50 alcohol:clearing reagent and then 100% clearing agent (xylene or xylene substitute). Infiltration involves incubating the tissue with a liquid form of the embedding agent (warm wax, nitrocellulose solution) first at 50:50 embedding agent: clearing agent and the 100% embedding agent. Embedding is completed by placing the tissue in a mold or cassette and filling with melted embedding agent such as wax, agar, or gelatin. The embedding agent is allowed to harden. The hardened tissue sample may then be sliced into thin section for staining and subsequent examination.

[0199] Prior to staining, the tissue section is dewaxed and rehydrated. Xylene is used to dewax the section, one or more changes of xylene may be used, and the tissue is rehydrated by successive washes in alcohol of decreasing concentration. Prior to dewax, the tissue section may be heat immobilized to a glass slide at about 80°C for about 20 minutes.

[0200] Laser capture micro-dissection allows the isolation of a subset of cells for further analysis from a tissue section.

[0201] As in cytology, to enhance the visualization of the microscopic features, the tissue section or slice can be stained with a variety of stains. A large menu of commercially available stains can be used to enhance or identify specific features.

[0202] To further increase the interaction of molecular reagents with cytological/histological samples, a number of techniques for "analyte retrieval" have been developed. The first such technique uses high temperature heating of a fixed sample. This method is also referred to as heat-induced epitope retrieval or HIER. A variety of heating techniques have been used, including steam heating, microwaving, autoclaving, water baths, and pressure cooking or a combination of these methods of heating. Analyze retrieval solutions include, for example, water, citrate, and formal saline buffers. The key to analyte retrieval is the time at high temperature but lower temperatures for longer times have also been successfully used. Another key to analyte retrieval is the pH of the heating solution. Low pH has been found to provide the best immunostaining but also gives rise to backgrounds that frequently require the use of a second tissue Section as a negative control. The most consistent ben-efit (increased immunostaining without increase in background) is generally obtained with a high pH solution regardless of the buffer composition. The analyte retrieval process for a specific target is empirically optimized for the target using heat, time, pH, and buffer composition as variables for process optimization. Using the microwave analyte retrieval method allows for sequential staining of different targets with antibody reagents. But the time required to achieve antibody and enzyme complexes between staining steps has also been shown to degrade cell membrane analytes. Microwave heating methods have improved in situ hybridization methods as well.

[0203] To initiate the analyte retrieval process, the section is first dewaxed and hydrated. The slide is then placed in 10mM sodium citrate buffer pH 6.0 in a dish or jar. A representative procedure uses an 1100W microwave and microwaves the slide at 100% power for 2 minutes followed by microwaving the slides using 20% power for 18 minutes after checking to be sure the slide remains covered in liquid. The slide is then allowed to cool in the uncovered contained and then rinsed with distilled water. HIER may be used in combination with an enzymatic digestion to improve the reactivity of the target to immunochemical reagents.

[0204] One such enzymatic digestion protocol uses proteinase K. A 20 g/ml concentration of proteinase K is prepared

in 50 mM Tris Base, 1mM EDTA, 0.5% Triton X-100, pH 8.0 buffer. The process first involves dewaxing sections in 2 changes of xylene, 5 minutes each. Then the sample is hydrated in 2 changes of 100% ethanol for 3 minutes each, 95% and 80% ethanol for 1 minute each, and then rinsed in distilled water. Sections are covered with Proteinase K working solution and incubated 10-20 minutes at 37°C in humidified chamber (optimal incubation time may vary depending on tissue type and degree of fixation). The sections are cooled at room temperature for 10 minutes and then rinsed in PBS Tween 20 for 2x2 min. If desired, sections can be blocked to eliminate potential interference from endogenous compounds and enzymes. The Section is then incubated with primary antibody at appropriate dilution in primary antibody dilution buffer for 1 hour at room temperature or overnight at 4°C. The Section is then rinsed with PBS Tween 20 for 2x2 min. Additional blocking can be performed, if required for the specific application, followed by additional rinsing with PBS Tween 20 for 3x2 min and then finally the immunostaining protocol completed.

[0205] A simple treatment with 1% SDS at room temperature has also been demonstrated to improve immunohistochemical staining. Analyze retrieval methods have been applied to slide mounted sections as well as free floating sections. Another treatment option is to place the slide in ajar containing citric acid and 0.1 Nonident P40 at pH 6.0 and heating to 95°C. The slide is then washed with a buffer solution like PBS.

[0206] For immunological staining of tissues it may be useful to block non-specific association of the antibody with tissue proteins by soaking the Section in a protein solution like serum or non-fat dry milk.

[0207] Blocking reactions may include the need to reduce the level of endogenous biotin; eliminate endogenous charge effects; inactivate endogenous nucleases; and/ or inactivate endogenous enzymes like peroxidase and alkaline phosphatase. Endogenous nucleases may be inactivated by degradation with proteinase K, by heat treatment, use of a chelating agent such as EDTA or EGTA, the introduction of carrier DNA or RNA, treatment with a chaotrope such as urea, thiourea, guanidine hydrochloride, guanidine thiocyanate, lithium perchlorate, etc, or diethyl pyrocarbonate. Alkaline phosphatase may be inactivated by treated with 0.1N HCl for 5 minutes at room temperature or treatment with 1 mM lev-amisole. Peroxidase activity may be eliminated by treatment with 0.03% hydrogen peroxide. Endogenous biotic may be blocked by soaking the slide or Section in an avidin (streptavidin, neutravidin may be substituted) solution for at least 15 minutes at room temperature. The slide or Section is then washed for at least 10 minutes in buffer. This may be repeated at least three times. Then the slide or section is soaked in a biotic solution for 10 minutes. This may be repeated at least three times with a fresh biotic solution each time. The buffer wash procedure is repeated. Blocking protocols should be minimized to prevent damaging either the cell or tissue structure or the target or targets of interest but one or more of these protocols could be combined to "block" a slide or Section prior to reaction with one or more slow off-rate aptamers. See Basic Medical Histology: the Biology of Cells, Tissues and Organs, authored by Richard G. Kessel, Oxford University Press, 1998.

Determination of Biomarker Values using Mass Spectrometry Meth-ods

[0208] A variety of configurations of mass spectrometers can be used to detect biomarker values. Several types of mass spectrometers are available or can be produced with various configurations. In general, a mass spectrometer has the following major components: a sample inlet, an ion source, a mass analyzer, a detector, a vacuum system, and instrument-control system, and a data system. Difference in the sample inlet, ion source, and mass analyzer generally define the type of instrument and its capabilities. For example, an inlet can be a capillary-column liquid chromatography source or can be a direct probe or stage such as used in matrix-assisted laser desorption. Common ion sources are, for example, electrospray, including nanospray and microspray or matrix-assisted laser desorption. Common mass analyzers include a quadrupole mass filter, ion trap mass analyzer and time-of-night mass analyzer. Additional mass spectrometry methods are well known in the art (see Burlingame et al. Anal. Chem. 70:647 R-716R (1998); Kinter and Sherman, New York (2000)).

[0209] Protein biomarkers and biomarker values can be detected and measured by any of the following: electrospray ionization mass spectrometry (ESI-MS), ESI-MS/MS, ESI-MS/(MS)n, matrix-assisted laser desorption ionization time-of-night mass spectrometry (MALI-TOF-MS), surface-enhanced laser desorption/ionization time-of-night mass spectrometry (SELDI-TOF-MS), desorption/ionization on silicon (DIOS), secondary ion mass spectrometry (SIMS), quadrupole time-of-night (Q-TOF), tandem time-of-flight (TOF/TOF) technology, called ultraflex III TOF/TOF, atmospheric pressure chemical ionization mass spectrometry (APCI-MS), APCI-MS/MS, APCI-(MS)N, atmospheric pressure photoionization mass spectrometry (APPI-MS), APPI-MS/MS, and APPI-(MS)N, quadrupole mass spectrometry, Fourier transform mass spectrometry (FTMS), quantitative mass spectrometry, and ion trap mass spectrometry.

[0210] Sample preparation strategies are used to label and enrich samples before mass spectroscopic characterization of protein biomarkers and determination biomarker values. Labeling methods include but are not limited to isobaric tag for relative and absolute quantitation (iTRAQ) and stable isotope labeling with amino acids in cell culture (SILAC). Capture reagents used to selectively enrich samples for candidate biomarker proteins prior to mass spectroscopic analysis include but are not limited to aptamers, antibodies, nucleic acid probes, chimeras, small molecules, an F(ab')2 fragment, a single chain antibody fragment, an Fv fragment, a single chain Fv fragment, a nucleic acid, a lectin, a ligand-binding receptor,

affybodies, nanobodies, ankyrins, domain antibodies, alternative antibody scaffolds (e.g. diabodies etc) imprinted polymers, avimers, peptidomimetics, peptoids, peptide nucleic acids, threose nucleic acid, a hormone receptor, a cytokine receptor, and synthetic receptors, and modifications and fragments of these.

Determination of Biomarker Values using a Proximity Ligation Assay

[0211] A proximity ligation assay can be used to determine biomarker values. Briefly, a test sample is contacted with a pair of affinity probes that may be a pair of antibodies or a pair of aptamers, with each member of the pair extended with an oligonucleotide. The targets for the pair of affinity probes may be two distinct determinates on one protein or one determinate on each of two different proteins, which may exist as homo- or hetero-multimeric complexes. When probes bind to the target determinates, the free ends of the oligonucleotide extensions are brought into sufficiently close proximity to hybridize together. The hybridization of the oligonucleotide extensions is facilitated by a common connector oligonucleotide which serves to bridge together the oligonucleotide extensions when they are positioned in sufficient proximity. Once the oligonucleotide extensions of the probes are hybridized, the ends of the extensions are joined together by enzymatic DNA ligation.

[0212] Each oligonucleotide extension comprises a primer site for PCR amplification. Once the oligonucleotide extensions are ligated together, the oligonucleotides form a continuous DNA sequence which, through PCR amplification, reveals information regarding the identity and amount of the target protein, as well as, information regarding protein-protein interactions where the target determinates are on two different proteins. Proximity ligation can provide a highly sensitive and specific assay for real-time protein concentration and interaction information through use of real-time PCR. Probes that do not bind the determinates of interest do not have the corresponding oligonucleotide extensions brought into proximity and no ligation or PCR amplification can proceed, resulting in no signal being produced.

[0213] The foregoing assays enable the detection of biomarker values that are useful in methods for diagnosing mesothelioma, where the methods comprise detecting, in a biological sample from an individual, at least N biomarker values that each correspond to a biomarker selected from the group consisting of the biomarkers provided in Table 1, wherein a classification, as described in detail below, using the biomarker values indicates whether the individual has mesothelioma. While certain of the described mesothelioma biomarkers are useful alone for detecting and diagnosing mesothelioma, methods are also described herein for the grouping of multiple subsets of the mesothelioma biomarkers that are each useful as a panel of three or more biomarkers. Thus, various embodiments of the instant application provide combinations comprising N biomarkers, wherein N is at least three biomarkers. In other embodiments, N is selected to be any number from 2-66 biomarkers. It will be appreciated that N can be selected to be any number from any of the above described ranges, as well as similar, but higher order, ranges. In accordance with any of the methods described herein, biomarker values can be detected and classified individually or they can be detected and classified collectively, as for example in a multiplex assay format.

[0214] In another aspect, methods are provided for detecting an absence of mesothelioma, the methods comprising detecting, in a biological sample from an individual, at least N biomarker values that each correspond to a biomarker selected from the group consisting of the biomarkers provided in Table 1, wherein a classification, as described in detail below, of the biomarker values indicates an absence of mesothelioma in the individual. While certain of the described mesothelioma biomarkers are useful alone for detecting and diagnosing the absence of mesothelioma, methods are also described herein for the grouping of multiple subsets of the mesothelioma biomarkers that are each useful as a panel of three or more biomarkers. Thus, various embodiments of the instant application provide combinations comprising N biomarkers, wherein N is at least three biomarkers. In other embodiments, N is selected to be any number from 2-66 biomarkers. It will be appreciated that N can be selected to be any number from any of the above described ranges, as well as similar, but higher order, ranges. In accordance with any of the methods described herein, biomarker values can be detected and classified individually or they can be detected and classified collectively, as for example in a multiplex assay format.

Classification of Biomarkers and Calculation of Disease Scores

[0215] A biomarker "signature" for a given diagnostic test contains a set of markers, each marker having different levels in the populations of interest. Different levels, in this context, may refer to different means of the marker levels for the individuals in two or more groups, or different variances in the two or more groups, or a combination of both. For the simplest form of a diagnostic test, these markers can be used to assign an unknown sample from an individual into one of two groups, either diseased or not diseased. The assignment of a sample into one of two or more groups is known as classification, and the procedure used to accomplish this assignment is known as a classifier or a classification method. Classification methods may also be referred to as scoring methods. There are many classification methods that can be used to construct a diagnostic classifier from a set of biomarker values. In general, classification methods are most easily performed using supervised learning techniques where a data set is collected using samples obtained

from individuals within two (or more, for multiple classification states) distinct groups one wishes to distinguish. Since the class (group or population) to which each sample belongs is known in advance for each sample, the classification method can be trained to give the desired classification response. It is also possible to use unsupervised learning techniques to produce a diagnostic classifier.

[0216] Common approaches for developing diagnostic classifiers include decision trees; bagging, boosting, forests and random forests; rule inference based learning; Parzen Windows; linear models; logistic; neural network methods; unsupervised clustering; K-means; hierarchical ascending/ descending; semi-supervised learning; prototype methods; nearest neighbor; kernel density estimation; support vector machines; hidden Markov models; Boltzmann Learning; and classifiers may be combined either simply or in ways which minimize particular objective functions. For a review, see, e.g., Pattern Classification, R. O. Duda, et al., editors, John Wiley & Sons, 2nd edition, 2001; see also, The Elements of Statistical Learning - Data Mining, Inference, and Prediction, T. Hastie, et al., editors, Springer Science+Business Media, LLC, 2nd edition, 2009; each of which is incorporated by reference in its entirety.

[0217] To produce a classifier using supervised learning techniques, a set of samples called training data are obtained. In the context of diagnostic tests, training data includes samples from the distinct groups (classes) to which unknown samples will later be assigned. For example, samples collected from individuals in a control population and individuals in a particular disease population can constitute training data to develop a classifier that can classify unknown samples (or, more particularly, the individuals from whom the samples were obtained) as either having the disease or being free from the disease. The development of the classifier from the training data is known as training the classifier. Specific details on classifier training depend on the nature of the supervised learning technique. For purposes of illustration, an example of training a naïve Bayesian classifier will be described below (see, e.g., Pattern Classification, R. O. Duda, et al., editors, John Wiley & Sons, 2nd edition, 2001; see also, The Elements of Statistical Learning - Data Mining, Inference, and Prediction, T. Hastie, et al., editors, Springer Science+Business Media, LLC, 2nd edition, 2009).

[0218] Since typically there are many more potential biomarker values than samples in a training set, care must be used to avoid over-fitting. Over-fitting occurs when a statistical model describes random error or noise instead of the underlying relationship. Over-fitting can be avoided in a variety of way, including, for example, by limiting the number of markers used in developing the classifier, by assuming that the marker responses are independent of one another, by limiting the complexity of the underlying statistical model employed, and by ensuring that the underlying statistical model conforms to the data.

[0219] An illustrative example of the development of a diagnostic test using a set of biomarkers includes the application of a naïve Bayes classifier, a simple probabilistic classifier based on Bayes theorem with strict independent treatment of the biomarkers. Each biolarker is described by a class-dependent, probability density function (pdf) for the measured RFU values or log RFU (relative fluorescence units) values in each class. The joint pdfs for the set of markers in one class is assumed to be the product of the individual class-dependent pdfs for each biomarkers. Training a naïve Bayes classifier in this context, amounts to assigning parameters ("parameterization") to characterize the class dependent pdfs. Any underlying model for the class-dependent pdfs may be used, but the model should generally conform to the data observed in the training set.

[0220] Specifically, the class-dependent probability of measuring a value x_i for biomarker i in the disease class is written as $p(x_i|d)$ and the overall naïve Bayes probability of observing n markers with values $\tilde{x} = (x_1, x_2, \ldots x_n)$ is written as $p(\tilde{x}|d) = \prod_{i=1}^{n} p(x_i|d)$ where the individual x_is are the measured biomarker levels in RFU or log RFU. The classification assignment for an unknown is facilitated by calculating the probability of being diseased $p(d|\tilde{x})$ having measured $\tilde{x}$ compared to the probability of being disease free (control) $p(c|\tilde{x})$ for the same measured values. The ratio of these probabilities is computed from the class-dependant pdfs by application of Bares theorem, i.e., $\frac{p(d|\tilde{x})}{p(c|\tilde{x})} = \frac{p(\tilde{x}|d)p(d)}{p(\tilde{x}|c)(1-p(d))}$ where $p(d)$ is the prevalence of the disease in the population appropriate to the test. Taking the logarithm of both sides of this ratio and substituting the naïve Bayes class-dependent probabilities from above gives $ln\left(\frac{p(d|\tilde{x})}{p(c|\tilde{x})}\right) = \sum_{i-1}^{n} ln\left(\frac{p(x_i|d)}{p(x_i|c)}\right) + ln\left(\frac{p(d)}{1-p(d)}\right)$. This form is known as the log likelihood ratio and simply states that the log likelihood of being free of the particular disease versus having the disease and is primarily composed of the sum of individual log likelihood ratios of the n individual biomarkers. In its simplest form, an unknown sample (or, more particularly, the individual from whom the sample was obtained) is classified as being free of the disease if the above ratio is greater than zero and having the disease if the ratio is less than zero.

[0221] In one exemplary embodiment, the class-dependent biomarker pdfs $p(x_i|C)$ and $p(x_i|d)$ are assumed to be normal or log-normal distributions in the measured RFU values x_i, i.e.

$$p(x_i|c) = \frac{1}{\sqrt{2\pi}\sigma_{c,i}} exp\left(-\frac{(x_i - \mu_{c,i})^2}{2\sigma_{c,i}^2}\right),$$ with a similar expression for $p(x_i|d)$ with μ_d and σ_d. Parameterization of the model requires estimation of two parameters for each class-dependent pdf, a mean μ and a variance σ^2, from the training data. This may be accomplished in a number of ways, including, for example, by maximum likelihood estimates, by least-squares, and by any other methods known to one skilled in the art. Substituting the normal distributions for and into the log-likelihood ratio defined above gives the following expression:

$$ln\left(\frac{p(d|\tilde{x})}{p(c|\tilde{x})}\right) = \sum_{i=1}^{n} ln\left(\frac{\sigma_{c,i}}{\sigma_{d,i}}\right) - \frac{1}{2}\sum_{i=1}^{n}\left[\left(\frac{x_i - \mu_{d,i}}{\sigma_{d,i}}\right)^2 - \left(\frac{x_i - \mu_{c,i}}{\sigma_{c,i}}\right)^2\right] + ln\left(\frac{p(d)}{1 - p(d)}\right)$$

[0222] Once a set of μs and σ^2s have been defined for each pdf in each class from the training data and the disease prevalence in the population is specified, the Bayes classifier is fully determined and may be used to classify unknown samples with measured values $\tilde{x}$.

[0223] The performance of the naïve Bayes classifier is dependent upon the number and quality of the biomarkers used to construct and train the classifier. A single biomarker will perform in accordance with its KS-distance (Kolmogorov-Smirnov), as defined in Example 3, below. If a classifier performance metric is defined as the area under the receiver operator characteristic curve (AUC), a perfect classifier will have a score of 1 and a random classifier, on average, will have a score of 0.5. The definition of the KS-distance between two sets *A* and *B* of sizes *n* and *m* is the value, $D_{n,m} = sup_x|F_{A,n}(x) - F_{B,m}(x)|$, which is the largest difference between two empirical cumulative distribution functions (cdf). The empirical cdf for a set *A* of *n* observations X_i is defined as, $F_{A,n}(x) = \frac{1}{n}\sum_{i=1}^{n} I_{X_i \leq x}$, where $IX_i \leq x$ is the indicator function which is equal to 1 if $X_i < x$ and is otherwise equal to 0. By definition, this value is bounded between 0 and 1, where a KS-distance of 1 indicates that the emperical distributions do not overlap.

[0224] The addition of subsequent markers with good KS distances (>0.3, for example) will, in general, improve the classification performance if the subsequently added markers are independent of the first marker. Using the area under the ROC curve (AUC) as a classifier score, it is straightforward to generate many high scoring classifiers with a variation of a greedy algorithm. (A greedy algorithm is any algorithm that follows the problem solving metaheuristic of making the locally optimal choice at each stage with the hope of finding the global optimum.)

[0225] The algorithm approach used here is described in detail in Example 4. Briefly, all single analyte classifiers are generated from a table of potential biomarkers and added to a list. Next, all possible additions of a second analyte to each of the stored single analyte classifiers is then performed, saving a predetermined number of the best scoring pairs, say, for example, a thousand, on a new list. All possible three marker classifiers are explored using this new list of the best two-marker classifiers, again saving the best thousand of these. This process continues until the score either plateaus or begins to deteriorate as additional markers are added. Those high scoring classifiers that remain after convergence can be evaluated for the desired performance for an intended use. For example, in one diagnostic application, classifiers with a high sensitivity and modest specificity may be more desirable than modest sensitivity and high specificity. In another diagnostic application, classifiers with a high specificity and a modest sensitivity may be more desirable. The desired level of performance is generally selected based upon a trade-off that must be made between the number of false positives and false negatives that can each be tolerated for the particular diagnostic application. Such trade-offs generally depend on the medical consequences of an error, either false positive or false negative.

[0226] Various other techniques are known in the art and may be employed to generate many potential classifiers from a list of biomarkers using a naive Bayes classifier. In one embodiment, what is referred to as a genetic algorithm can be used to combine different markers using the fitness score as defined above. Genetic algorithms are particularly well suited to exploring a large diverse population of potential classifiers. In another embodiment, so-called ant colony optimization can be used to generate sets of classifiers. Other strategies that are known in the art can also be employed, including, for example, other evolutionary strategies as well as simulated annealing and other stochastic search methods. Metaheuristic methods, such as, for example, harmony search may also be employed.

[0227] Exemplary embodiments use any number of the mesothelioma biomarkers listed in Table 1 in various combinations to produce diagnostic tests for detecting mesothelioma (see Example 2 for a detailed description of how these biomarkers were identified). In one embodiment, a method for diagnosing mesothelioma uses a naïve Bayes classification method in conjunction with any number of the mesothelioma biomarkers listed in Table 1. In an illustrative example

(Example 3), the simplest test for detecting mesothelioma from a population of asbestos exposed individuals can be constructed using a single biomarker, for example, CDH1 which is differentially expressed in mesothelioma with a KS-distance of 0.63. Using the parameters, $\mu_{c,i}$, $\sigma_{c,i}$, $\mu_{d,i}$, and, $\sigma_{d,i}$ for CDH1 from Table 16 and the equation for the log-likelihood described above, a diagnostic test with an AUC of 0.884 can be derived, see Table 15. The ROC curve for this test is displayed in Figure 2.

[0228] Addition of biomarker BMPER, for example, with a KS-distance of 0.60, significantly improves the classifier performance to an AUC of 0.947. Note that the score for a classifier constructed of two biomarkers is not a simple sum of the KS-distances; KS-distances are not additive when combining biomarkers and it takes many more weak markers to achieve the same level of performance as a strong marker. Adding a third marker, F9, for example, boosts the classifier performance to an AUC of 0.951. Adding additional biomarkers, such as, for example, CCL23, CRK, BMP1, TPT1, FRZB, MDK, and ICAM2, produces a series of mesothelioma tests summarized in Table 15 and displayed as a series of ROC curves in Figure 3. The score of the classifiers as a function of the number of analytes used in classifier construction is displayed in Figure 4. The AUC of this exemplary ten-marker classifier is 0.993.

[0229] The markers listed in Table 1 can be combined in many ways to produce classifiers for diagnosing mesothelioma. In some embodiments, panels of biomarkers are comprised of different numbers of analytes depending on a specific diagnostic performance criterion that is selected. For example, certain combinations of biomarkers will produce tests that are more sensitive (or more specific) than other combinations.

[0230] Once a panel is defined to include a particular set of biomarkers from Table 1 and a classifier is constructed from a set of training data, the definition of the diagnostic test is complete. In one embodiment, the procedure used to classify an unknown sample is outlined in Figure 1A. In another embodiment the procedure used to classify an unknown sample is outlined in Figure 1B. The biological sample is appropriately diluted and then run in one or more assays to produce the relevant quantitative biomarker levels used for classification. The measured biomarker levels are used as input for the classification method that outputs a classification and an optional score for the sample that reflects the confidence of the class assignment.

[0231] Table 1 identifies 66 biomarkers that are useful for diagnosing mesothelioma. This is a surprisingly larger number than expected when compared to what is typically found during biomarker discovery efforts and may be attributable to the scale of the described study, which encompassed over 1000 proteins measured in hundreds of individual samples, in some cases at concentrations in the low femtomolar range. Presumably, the large number of discovered biomarkers reflects the diverse biochemical pathways implicated in both tumor biology and the body's response to the tumor's presence; each pathway and process involves many proteins. The results show that no single protein of a small group of proteins is uniquely informative about such complex processes; rather, that multiple proteins are involved in relevant processes, such as apoptosis or extracellular matrix repair, for example.

[0232] Given the numerous biomarkers identified during the described study, one would expect to be able to derive large numbers of high-performing classifiers that can be used in various diagnostic methods. To test this notion, tens of thousands of classifiers were evaluated using the biomarkers in Table 1. As described in Example 4, many subsets of the biomarkers presented in Table 1 can be combined to generate useful classifiers. By way of example, descriptions are provided for classifiers containing 1, 2, and 3 biomarkers for detection of mesothelioma. As described in Example 4, all classifiers that were built using the biomarkers in Table 1 perform distinctly better than classifiers that were built using "non-markers".

[0233] The performance of classifiers obtained by randomly excluding some of the markers in Table 1, which resulted in smaller subsets from which to build the classifiers, was also tested. As described in Example 4, Part 3, the classifiers that were built from random subsets of the markers in Table 1 performed similarly to optimal classifiers that were built using the full list of markers in Table 1.

[0234] The performance of ten-marker classifiers obtained by excluding the "best" individual markers from the ten-marker aggregation was also tested. As described in Example 4, Part 3, classifiers constructed without the "best" markers in Table 1 also performed well. Many subsets of the biomarkers listed in Table 1 performed close to optimally, even after removing the top 15 of the markers listed in the Table. This implies that the performance characteristics of any particular classifier are likely not due to some small core group of biomarkers and that the disease process likely impacts numerous biochemical pathways, which alters the expression level of many proteins.

[0235] The results from Example 4 suggest certain possible conclusions: First, the identification of a large number of biomarkers enables their aggregation into a vast number of classifiers that offer similarly high performance. Second, classifiers can be constructed such that particular biomarkers may be substituted for other biomarkers in a manner that reflects the redundancies that undoubtedly pervade the complexities of the underlying disease processes. That is to say, the information about the disease contributed by any individual biomarker identified in Table 1 overlaps with the information contributed by other biomarkers, such that it may be that no particular biomarker or small group of biomarkers in Table 1 must be included in any classifier.

[0236] Exemplary embodiments use naive Bayes classifiers constructed from the data in Table 16 to classify an unknown sample. The procedure is outlined in Figures 1A and 1B. In one embodiment, the biological sample is optionally

diluted and run in a multiplexed aptamer assay. The data from the assay are normalized and calibrated as outlined in Example 3, and the resulting biomarker levels are used as input to a Bayes classification scheme. The log-likelihood ratio is computed for each measured biomarker individually and then summed to produce a final classification score, which is also referred to as a diagnostic score. The resulting assignment as well as the overall classification score can be reported. Optionally, the individual log-likelihood risk factors computed for each biomarker level can be reported as well. The details of the classification score calculation are presented in Example 3.

Kits

[0237] Any combination of the biomarkers of Table 1 (as well as additional biomedical information) can be detected using a suitable kit, such as for use in performing the methods disclosed herein. Furthermore, any kit can contain one or more detectable labels as described herein, such as a fluorescent moiety, etc.

[0238] In one embodiment, a kit includes (a) one or more capture reagents (such as, for example, at least one aptamer or antibody) for detecting one or more biomarkers in a biological sample, wherein the biomarkers include any of the biomarkers set forth in Table 1, and optionally (b) one or more software or computer program products for classifying the individual from whom the biological sample was obtained as either having or not having mesothelioma or for determining the likelihood that the individual has mesothelioma, as further described herein. Alternatively, rather than one or more computer program products, one or more instructions for manually performing the above steps by a human can be provided.

[0239] The combination of a solid support with a corresponding capture reagent and a signal generating material is referred to herein as a "detection device" or "kit". The kit can also include instructions for using the devices and reagents, handling the sample, and analyzing the data. Further the kit may be used with a computer system or software to analyze and report the result of the analysis of the biological sample.

[0240] The kits can also contain one or more reagents (e.g., solubilization buffers, detergents, washes, or buffers) for processing a biological sample. Any of the kits described herein can also include, e.g., buffers, blocking agents, mass spectrometry matrix materials, antibody capture agents, positive control samples, negative control samples, software and information such as protocols, guidance and reference data.

[0241] In one aspect, the invention provides kits for the analysis of mesothelioma status. The kits include PCR primers for one or more biomarkers selected from Table 1. The kit may further include instructions for use and correlation of the biomarkers with mesothelioma. The kit may also include a DNA array containing the complement of one or more of the biomarkers selected from Table 1, reagents, and/or enzymes for amplifying or isolating sample DNA. The kits may include reagents for real-time PCR, for example, TaqMan probes and/or primers, and enzymes.

[0242] For example, a kit can comprise (a) reagents comprising at least capture reagent for quantifying one or more biomarkers in a test sample, wherein said biomarkers comprise the biomarkers set forth in Table 1, or any other biomarkers or biomarkers panels described herein, and optionally (b) one or more algorithms or computer programs for performing the steps of comparing the amount of each biomarker quantified in the test sample to one or more predetermined cutoffs and assigning a score for each biomarker quantified based on said comparison, combining the assigned scores for each biomarker quantified to obtain a total score, comparing the total score with a predetermined score, and using said comparison to determine whether an individual has mesothelioma. Alternatively, rather than one or more algorithms or computer programs, one or more instructions for manually performing the above steps by a human can be provided.

Computer Methods and Software

[0243] Once a biomarker or biomarker panel is selected, a method for diagnosing an individual can comprise the following: 1) collect or otherwise obtain a biological sample; 2) perform an analytical method to detect and measure the biomarker or biomarkers in the panel in the biological sample; 3) perform any data normalization or standardization required for the method used to collect biomarker values; 4) calculate the marker score; 5) combine the marker scores to obtain a total diagnostic score; and 6) report the individual's diagnostic score. In this approach, the diagnostic score may be a single number determined from the sum of all the marker calculations that is compared to a preset threshold value that is an indication of the presence or absence of disease. Or the diagnostic score may be a series of bars that each represent a biomarker value and the pattern of the responses may be compared to a pre-set pattern for determination of the presence or absence of disease.

[0244] At least some embodiments of the methods described herein can be implemented with the use of a computer. An example of a computer system 100 is shown in Figure 6. With reference to Figure 6, system 100 is shown comprised of hardware elements that are electrically coupled via bus 108, including a processor 101, input device 102, output device 103, storage device 104, computer-readable storage media reader 105a, communications system 106 processing acceleration (e.g., DSP or special-purpose processors) 107 and memory 109. Computer-readable storage media reader 105a is further coupled to computer-readable storage media 105b, the combination comprehensively representing re-

mote, local, fixed and/or removable storage devices plus storage media, memory, etc. for temporarily and/or more permanently containing computer-readable information, which can include storage device 104, memory 109 and/or any other such accessible system 100 resource. System 100 also comprises software elements (shown as being currently located within working memory 191) including an operating system 192 and other code 193, such as programs, data and the like.

[0245] With respect to Figure 6, system 100 has extensive flexibility and configurability. Thus, for example, a single architecture might be utilized to implement one or more servers that can be further configured in accordance with currently desirable protocols, protocol variations, extensions, etc. However, it will be apparent to those skilled in the art that embodiments may well be utilized in accordance with more specific application requirements. For example, one or more system elements might be implemented as sub-elements within a system 100 component (e.g., within communications system 106). Customized hardware might also be utilized and/or particular elements might be implemented in hardware, software or both. Further, while connection to other computing devices such as network input/output devices (not shown) may be employed, it is to be understood that wired, wireless, modem, and/or other connection or connections to other computing devices might also be utilized.

[0246] In one aspect, the system can comprise a database containing features of biomarkers characteristic of mesothelioma. The biomarker data (or biomarker information) can be utilized as an input to the computer for use as part of a computer implemented method. The biomarker data can include the data as described herein.

[0247] In one aspect, the system further comprises one or more devices for providing input data to the one or more processors.

[0248] The system further comprises a memory for storing a data set of ranked data elements.

[0249] In another aspect, the device for providing input data comprises a detector for detecting the characteristic of the data element, e.g., such as a mass spectrometer or gene chip reader.

[0250] The system additionally may comprise a database management system. User requests or queries can be formatted in an appropriate language understood by the database management system that processes the query to extract the relevant information from the database of training sets.

[0251] The system may be connectable to a network to which a network server and one or more clients are connected. The network may be a local area network (LAN) or a wide area network (WAN), as is known in the art. Preferably, the server includes the hardware necessary for running computer program products (e.g., software) to access database data for processing user requests.

[0252] The system may include an operating system (e.g., UNIX or Linux) for executing instructions from a database management system. In one aspect, the operating system can operate on a global communications network, such as the internet, and utilize a global communications network server to connect to such a network.

[0253] The system may include one or more devices that comprise a graphical display interface comprising interface elements such as buttons, pull down menus, scroll bars, fields for entering text, and the like as are routinely found in graphical user interfaces known in the art. Requests entered on a user interface can be transmitted to an application program in the system for formatting to search for relevant information in one or more of the system databases. Requests or queries entered by a user may be constructed in any suitable database language.

[0254] The graphical user interface may be generated by a graphical user interface code as part of the operating system and can be used to input data and/or to display inputted data. The result of processed data can be displayed in the interface, printed on a printer in communication with the system, saved in a memory device, and/or transmitted over the network or can be provided in the form of the computer readable medium.

[0255] The system can be in communication with an input device for providing data regarding data elements to the system (e.g., expression values). In one aspect, the input device can include a gene expression profiling system including, e.g., a mass spectrometer, gene chip or array reader, and the like.

[0256] The methods and apparatus for analyzing mesothelioma biomarker information according to various embodiments may be implemented in any suitable manner, for example, using a computer program operating on a computer system. A conventional computer system comprising a processor and a random access memory, such as a remotely-accessible application server, network server, personal computer or workstation may be used. Additional computer system components may include memory devices or information storage systems, such as a mass storage system and a user interface, for example a conventional monitor, keyboard and tracking device. The computer system may be a stand-alone system or part of a network of computers including a server and one or more databases.

[0257] The mesothelioma biomarker analysis system can provide functions and operations to complete data analysis, such as data gathering, processing, analysis, reporting and/or diagnosis. For example, in one embodiment, the computer system can execute the computer program that may receive, store, search, analyze, and report information relating to the mesothelioma biomarkers. The computer program may comprise multiple modules performing various functions or operations, such as a processing module for processing raw data and generating supplemental data and an analysis module for analyzing raw data and supplemental data to generate a mesothelioma status and/or diagnosis. Diagnosing mesothelioma status may comprise generating or collecting any other information, including additional biomedical in-

formation, regarding the condition of the individual relative to the disease, identifying whether further tests may be desirable, or otherwise evaluating the health status of the individual.

[0258] Referring now to Figure 7, an example of a method of utilizing a computer in accordance with principles of a disclosed embodiment can be seen. In Figure 7, a flowchart 3000 is shown. In block 3004, biomarker information can be retrieved for an individual. The biomarker information can be retrieved from a computer database, for example, after testing of the individual's biological sample is performed. The biomarker information can comprise biomarker values that each correspond to one of at least N biomarkers selected from a group consisting of the biomarkers provided in Table 1, wherein N = 2-66. In block 3008, a computer can be utilized to classify each of the biomarker values. And, in block 3012, a determination can be made as to the likelihood that an individual has mesothelioma based upon a plurality of classifications. The indication can be output to a display or other indicating device so that it is viewable by a person. Thus, for example, it can be displayed on a display screen of a computer or other output device.

[0259] Referring now to Figure 8, an alternative method of utilizing a computer in accordance with another embodiment can be illustrated via flowchart 3200. In block 3204, a computer can be utilized to retrieve biomarker information for an individual. The biomarker information comprises a biomarker value corresponding to a biomarker selected from the group of biomarkers provided in Table 1. In block 3208, a classification of the biomarker value can be performed with the computer. And, in block 3212, an indication can be made as to the likelihood that the individual has mesothelioma based upon the classification. The indication can be output to a display or other indicating device so that it is viewable by a person. Thus, for example, it can be displayed on a display screen of a computer or other output device.

[0260] Some embodiments described herein can be implemented so as to include a computer program product. A computer program product may include a computer readable medium having computer readable program code embodied in the medium for causing an application program to execute on a computer with a database.

[0261] As used herein, a "computer program product" refers to an organized set of instructions in the form of natural or programming language statements that are contained on a physical media of any nature (e.g., written, electronic, magnetic, optical or otherwise) and that may be used with a computer or other automated data processing system. Such programming language statements, when executed by a computer or data processing system, cause the computer or data processing system to act in accordance with the particular content of the statements. Computer program products include without limitation: programs in source and object code and/or test or data libraries embedded in a computer readable medium. Furthermore, the computer program product that enables a computer system or data processing equipment device to act in pre-selected ways may be provided in a number of forms, including, but not limited to, original source code, assembly code, object code, machine language, encrypted or compressed versions of the foregoing and any and all equivalents.

[0262] In one aspect, a computer program product is provided for indicating a likelihood of mesothelioma. The computer program product includes a computer readable medium embodying program code executable by a processor of a computing device or system, the program code comprising: code that retrieves data attributed to a biological sample from an individual, wherein the data comprises biomarker values that each correspond to one of at least N biomarkers in the biological sample selected from the group of biomarkers provided in Table 1, wherein N = 2-66; and code that executes a classification method that indicates a mesothelioma status of the individual as a function of the biomarker values.

[0263] In still another aspect, a computer program product is provided for indicating a likelihood of mesothelioma. The computer program product includes a computer readable medium embodying program code executable by a processor of a computing device or system, the program code comprising: code that retrieves data attributed to a biological sample from an individual, wherein the data comprises a biomarker value corresponding to a biomarker in the biological sample selected from the group of biomarkers provided in Table 1; and code that executes a classification method that indicates a mesothelioma status of the individual as a function of the biomarker value.

[0264] While various embodiments have been described as methods or apparatuses, it should be understood that embodiments can be implemented through code coupled with a computer, e.g., code resident on a computer or accessible by the computer. For example, software and databases could be utilized to implement many of the methods discussed above. Thus, in addition to embodiments accomplished by hardware, it is also noted that these embodiments can be accomplished through the use of an article of manufacture comprised of a computer usable medium having a computer readable program code embodied therein, which causes the enablement of the functions disclosed in this description. Therefore, it is desired that embodiments also be considered protected by this patent in their program code means as well. Furthermore, the embodiments may be embodied as code stored in a computer-readable memory of virtually any kind including, without limitation, RAM, ROM, magnetic media, optical media, or magneto-optical media. Even more generally, the embodiments could be implemented in software, or in hardware, or any combination thereof including, but not limited to, software running on a general purpose processor, microcode, PLAs, or ASICs.

[0265] It is also envisioned that embodiments could be accomplished as computer signals embodied in a carrier wave, as well as signals (e.g., electrical and optical) propagated through a transmission medium. Thus, the various types of information discussed above could be formatted in a structure, such as a data structure, and transmitted as an electrical

signal through a transmission medium or stored on a computer readable medium.

[0266] It is also noted that many of the structures, materials, and acts recited herein can be recited as means for performing a function or step for performing a function. Therefore, it should be understood that such language is entitled to cover all such structures, materials, or acts disclosed within this specification and their equivalents, including the matter incorporated by reference.

[0267] The biomarker identification process, the utilization of the biomarkers disclosed herein, and the various methods for determining biomarker values are described in detail above with respect to mesothelioma. However, the application of the process, the use of identified biomarkers, and the methods for determining biomarker values are fully applicable to other specific types of cancer, to cancer generally, to any other disease or medical condition, or to the identification of individuals who may or may not be benefited by an ancillary medical treatment. Except when referring to specific results related to mesothelioma, as is clear from the context, references herein to mesothelioma may be understood to include other types of cancer, cancer generally, or any other disease or medical condition.

EXAMPLES

[0268] The following examples are provided for illustrative purposes only and are not intended to limit the scope of the application as defined by the appended claims. All examples described herein were carried out using standard techniques, which are well known and routine to those of skill in the art. Routine molecular biology techniques described in the following examples can be carried out as described in standard laboratory manuals, such as Sambrook et al., Molecular Cloning: A Laboratory Manual, 3rd. ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., (2001).

Example 1. Multiplexed Aptamer Analysis of Samples

[0269] This example describes the multiplex aptamer assay used to analyze the samples and controls for the identification of the biomarkers set forth in Table 1 (see Figure 9) and the identification of the cancer biomarkers set forth in Table 19. For the mesothelioma, lung cancer, and renal cell carcinoma studies, the multiplexed analysis utilized 1045 aptamers, each unique to a specific target.

[0270] In this method, pipette tips were changed for each solution addition.

[0271] Also, unless otherwise indicated, most solution transfers and wash additions used the 96-well head of a Beckman Biomek FxP. Method steps manually pipetted used a twelve channel P200 Pipetteman (Rainin Instruments, LLC, Oakland, CA), unless otherwise indicated. A custom buffer referred to as SB17 was prepared in-house, comprising 40 mM HEPES, 100 mM NaCl, 5 mM KCl, 5 mM $MgCl_2$, 1 mM EDTA at pH 7.5. A custom buffer referred to as SB18 was prepared in-house, comprising 40 mM HEPES, 100 mM NaCl, 5 mM KCl, 5 mM $MgCl_2$ at pH 7.5. All steps were performed at room temperature unless otherwise indicated.

1. Preparation of Aptamer Stock Solution

[0272] Custom stock aptamer solutions for 5%, 0.316% and 0.01% serum were prepared at 2x concentration in 1x SB17, 0.05% Tween-20.

[0273] These solutions are stored at -20°C until use. The day of the assay, each aptamer mix was thawed at 37°C for 10 minutes, placed in a boiling water bath for 10 minutes and allowed to cool to 25°C for 20 minutes with vigorous mixing in between each heating step. After heat-cool, 55 μL of each 2x aptamer mix was manually pipetted into a 96-well Hybaid plate and the plate foil sealed. The final result was three, 96-well, foil-sealed Hybaid plates with 5%, 0.316% or 0.01% aptamer mixes. The individual aptamer concentration was 2x final or 1 nM.

2. Assay Sample Preparation

[0274] Frozen aliquots of 100% serum or plasma, stored at -80°C, were placed in 25°C water bath for 10 minutes. Thawed samples were placed on ice, gently vortexed (set on 4) for 8 seconds and then replaced on ice.

[0275] A 10% sample solution (2x final) was prepared by transferring 8 μL of sample using a 50 μL 8-channel spinning pipettor into 96-well Hybaid plates, each well containing 12 μL of the appropriate sample diluent at 4°C (1x SB17 for serum or 0.8x SB18 for plasma, plus 0.06% Tween-20, 11.1 μM Z-block2, 0.44 mM $MgCl_2$, 2.2mM AEBSF, 1.1mM EGTA, 55.6 μM EDTA). This plate was stored on ice until the next sample dilution steps were initiated on the BiomekFxP robot.

[0276] To commence sample and aptamer equilibration, the 10% sample plate was briefly centrifuged and placed on the Beckman FX where it was mixed by pipetting up and down with the 96-well pipettor. A 0.632% sample plate (2x final) was then prepared by diluting 6 μL of the 10% sample into 89 μL of 1xSB17, 0.05% Tween-20 with 2 mM AEBSF. Next, dilution of 6 μL of the resultant 0.632% sample into 184 μL of 1xSB17, 0.05% Tween-20 made a 0.02% sample

plate (2x final). Dilutions were done on the Beckman Biomek FxP. After each transfer, the solutions were mixed by pipetting up and down. The 3 sample dilution plates were then transferred to their respective aptamer solutions by adding 55 μL of the sample to 55 μL of the appropriate 2x aptamer mix. The sample and aptamer solutions were mixed on the robot by pipetting up and down.

3. Sample Equilibration binding

[0277] The sample/aptamer plates were foil sealed and placed into a 37°C incubator for 3.5 hours before proceeding to the Catch 1 step.

4. Preparation of Catch 2 bead plate

[0278] An 11 mL aliquot of MyOne (Invitrogen Corp., Carlsbad, CA) Streptavidin C1 beads (10 mg/mL) was washed 2 times with equal volumes of 20 mM NaOH (5 minute incubation for each wash), 3 times with equal volumes of 1x SB17, 0.05% Tween-20 and resuspended in 11 mL 1x SB17, 0.05% Tween-20. Using a 12-span multichannel pipettor, 50 μL of this solution was manually pipetted into each well of a 96-well Hybaid plate. The plate was then covered with foil and stored at 4°C for use in the assay.

5. Preparation of Catch 1 bead plates

[0279] Three 0.45 μm Millipore HV plates (Durapore membrane, Cat# MAHVN4550) were equilibrated with 100 μL of 1x SB17, 0.05% Tween-20 for at least 10 minutes. The equilibration buffer was then filtered through the plate and 133.3 μL of a 7.5% streptavidin-agarose bead slurry (in 1x SB17, 0.05% Tween-20) was added into each well. To keep the streptavidin-agarose beads suspended while transferring them into the filter plate, the bead solution was manually mixed with a 200 μL, 12-channel pipettor, at least 6 times between pipetting events. After the beads were distributed across the 3 filter plates, a vacuum was applied to remove the bead supernatant. Finally, the beads were washed in the filter plates with 200 μL 1x SB17, 0.05% Tween-20 and then resuspended in 200 μL 1x SB17, 0.05% Tween-20. The bottoms of the filter plates were blotted and the plates stored for use in the assay.

6. Loading the Cytomat

[0280] The cytomat was loaded with all tips, plates, all reagents in troughs (except NHS-biotin reagent which was prepared fresh right before addition to the plates), 3 prepared Catch 1 filter plates and 1 prepared MyOne plate.

7. Catch 1

[0281] After a 3.5 hour equilibration time, the sample/aptamer plates were removed from the incubator, centrifuged for about 1 minute, cover removed, and placed on the deck of the Beckman Biomek FxP. The Beckman Biomek FxP program was initiated. All subsequent steps in Catch 1 were performed by the Beckman Biomek FxP robot unless otherwise noted. Within the program, the vacuum was applied to the Catch 1 filter plates to remove the bead supernatant. One hundred microlitres of each of the 5%, 0.316% and 0.01% equilibration binding reactions were added to their respective Catch 1 filtration plates, and each plate was mixed using an on-deck orbital shaker at 800 rpm for 10 minutes.

[0282] Unbound solution was removed via vacuum filtration. The Catch 1 beads were washed with 190 μL of 100 μM biotin in 1x SB17, 0.05% Tween-20 followed by 5x 190 μL of 1x SB17, 0.05% Tween-20 by dispensing the solution and immediately drawing a vacuum to filter the solution through the plate.

8. Tagging

[0283] A 100 mM NHS-PEO4-biotin aliquot in anhydrous DMSO was thawed at 37°C for 6 minutes and then diluted 1:100 with tagging buffer (SB17 at pH 7.25, 0.05% Tween-20). Upon a robot prompt, the diluted NHS-PEO4-biotin reagent was manually added to an on-deck trough and the robot program was manually re-initiated to dispense 100 μL of the NHS-PEO4-biotin into each well of each Catch 1 filter plate. This solution was allowed to incubate with Catch 1 beads shaking at 800 rpm for 5 minutes on the orbital shakers.

9. Kinetic Challenge and Photo-cleavage

[0284] The tagging reaction was removed by vacuum filtration and quenched by the addition of 150 μL of 20 mM glycine in 1x SB17, 0.05% Tween-20 to the Catch 1 plates. The NHS-tag/glycine solution was removed via vacuum

filtration. Next, 1500 μL 20 mM glycine (1x SB17, 0.05% Tween-20) was added to each plate and incubated for 1 minute on orbital shakers at 800 rpm before removal by vacuum filtration.

[0285] The wells of the Catch 1 plates were subsequently washed three times by adding 190 μL 1x SB17, 0.05% Tween-20, followed by vacuum filtration and then by adding 190 μL 1x SB17, 0.05% Tween-20 with shaking for 1 minute at 800 rpm followed by vacuum filtration. After the last wash the plates were placed on top of a 1 mL deep-well plate and removed from the deck. The Catch 1 plates were centrifuged at 1000 rpm for 1 minute to remove as much extraneous volume from the agarose beads before elution as possible.

[0286] The plates were placed back onto the Beckman Biomek FxP and 85 μL of 10 mM DxSO4 in 1x SB17, 0.05% Tween-20 was added to each well of the filter plates.

[0287] The filter plates were removed from the deck, placed onto a Variomag Thermoshaker (Thermo Fisher Scientific, Inc., Waltham, MA) under the BlackRay (Ted Pella, Inc., Red-ding, CA) light sources, and irradiated for 5 minutes while shaking at 800 rpm. After the 5 minute incubation the plates were rotated 180 degrees and irradiated with shaking for 5 minutes more.

[0288] The photocleaved solutions were sequentially eluted from each Catch 1 plate into a common deep well plate by first placing the 5% Catch 1 filter plate on top of a 1 mL deep-well plate and centrifuging at 1000 rpm for 1 minute. The 0.316% and 0.01% Catch 1 plates were then sequentially centrifuged into the same deep well plate.

10. Catch 2 bead capture

[0289] The 1 mL deep well block containing the combined eluates of Catch 1 was placed on the deck of the Beckman Biomek FxP for Catch 2.

[0290] The robot transferred all of the photo-cleaved eluate from the 1 mL deep-well plate onto the Hybaid plate containing the previously prepared Catch 2 MyOne magnetic beads (after removal of the MyOne buffer via magnetic separation).

[0291] The solution was incubated while shaking at 1350 rpm for 5 minutes at 25°C on a Variomag Thermoshaker (Thermo Fisher Scientific, Inc., Waltham, MA).

[0292] The robot transferred the plate to the on deck magnetic separator station. The plate was incubated on the magnet for 90 seconds before removal and discarding of the supernatant.

11. 37°C 30% glycerol washes

[0293] The Catch 2 plate was moved to the on-deck thermal shaker and 15 μL of 1x SB17, 0.05% Tween-20 was transferred to each well. The plate was mixed for 1 minute at 1350 rpm and 37°C to resuspend and warm the beads. To each well of the Catch 2 plate, 75 μL of 60% glycerol at 37°C was transferred and the plate continued to mix for another minute at 1350 rpm and 37°C. The robot transferred the plate to the 37°C magnetic separator where it was incubated on the magnet for 2 minutes and then the robot removed and discarded the supernatant. These washes were repeated two more times.

[0294] After removal of the third 30% glycerol wash from the Catch 2 beads, 150 μL of 1x SB17, 0.05% Tween-20 was added to each well and incubated at 37°C, shaking at 1350 rpm for 1 minute, before removal by magnetic separation on the 37°C magnet.

[0295] The Catch 2 beads were washed a final time using 150 μL 1x SB17, 0.05% Tween-20 with incubation for 1 minute while shaking at 1350 rpm at 25°C prior to magnetic separation.

12. Catch 2 Bead Elution and Neutralization

[0296] The aptamers were eluted from Catch 2 beads by adding 105 μL of 100 mM CAPSO with 1 M NaCl, 0.05% Tween-20 to each well. The beads were incubated with this solution with shaking at 1300 rpm for 5 minutes.

[0297] The Catch 2 plate was then placed onto the magnetic separator for 90 seconds prior to transferring 63 μL of the eluate to a new 96-well plate containing 7 μL of 500 mM HCl, 500 mM HEPES, 0.05% Tween-20 in each well. After transfer, the solution was mixed robotically by pipetting 60 μL up and down five times.

13. Hybridization

[0298] The Beckman Biomek FxP transferred 20 μL of the neutralized Catch 2 eluate to a fresh Hybaid plate, and 6 μL of 10x Agilent Block, containing a 10x spike of hybridization controls, was added to each well. Next, 30 μL of 2x Agilent Hybridization buffer was manually pipetted to the each well of the plate containing the neutralized samples and blocking buffer and the solution was mixed by manually pipetting 25 μL up and down 15 times slowly to avoid extensive bubble formation. The plate was spun at 1000 rpm for 1 minute.

[0299] Custom Agilent microarray slides (Agilent Technologies, Inc., Santa Clara, CA) were designed to contain probes complementary to the aptamer random region plus some primer region. For the majority of the aptamers, the optimal length of the complementary sequence was empirically determined and ranged between 40-50 nucleotides. For later aptamers a 46-mer complementary region was chosen by default. The probes were linked to the slide surface with a poly-T linker for a total probe length of 60 nucleotides.

[0300] A gasket slide was placed into an Agilent hybridization chamber and 40 μL of each of the samples containing hybridization and blocking solution was manually pipetted into each gasket. An 8-channel variable spanning pipettor was used in a manner intended to minimize bubble formation. Custom Agilent microarray slides (Agilent Technologies, Inc., Santa Clara, CA),with their Number Barcode facing up, were then slowly lowered onto the gasket slides (see Agilent manual for detailed description).

[0301] The top of the hybridization chambers were placed onto the slide/backing sandwich and clamping brackets slid over the whole assembly. These assemblies were tightly clamped by turning the screws securely.

[0302] Each slide/backing slide sandwich was visually inspected to assure the solution bubble could move freely within the sample. If the bubble did not move freely the hybridization chamber assembly was gently tapped to disengage bubbles lodged near the gasket.

[0303] The assembled hybridization chambers were incubated in an Agilent hybridization oven for 19 hours at 60°C rotating at 20 rpm.

14. Post Hybridization Washing

[0304] Approximately 400 mL Agilent Wash Buffer 1 was placed into each of two separate glass staining dishes. One of the staining dishes was placed on a magnetic stir plate and a slide rack and stir bar were placed into the buffer.

[0305] A staining dish for Agilent Wash 2 was prepared by placing a stir bar into an empty glass staining dish.

[0306] A fourth glass staining dish was set aside for the final acetonitrile wash.

[0307] Each of six hybridization chambers was disassembled. One-by-one, the slide/backing sandwich was removed from its hybridization chamber and submerged into the staining dish containing Wash 1. The slide/backing sandwich was pried apart using a pair of tweezers, while still submerging the microarray slide. The slide was quickly transferred into the slide rack in the Wash 1 staining dish on the magnetic stir plate.

[0308] The slide rack was gently raised and lowered 5 times. The magnetic stirrer was turned on at a low setting and the slides incubated for 5 minutes.

[0309] When one minute was remaining for Wash 1, Wash Buffer 2 pre-warmed to 37°C in an incubator was added to the second prepared staining dish. The slide rack was quickly transferred to Wash Buffer 2 and any excess buffer on the bottom of the rack was removed by scraping it on the top of the stain dish. The slide rack was gently raised and lowered 5 times. The magnetic stirrer was turned on at a low setting and the slides incubated for 5 minutes.

[0310] The slide rack was slowly pulled out of Wash 2, taking approximately 15 seconds to remove the slides from the solution.

[0311] With one minute remaining in Wash 2 acetonitrile (ACN) was added to the fourth staining dish. The slide rack was transferred to the acetonitrile stain dish. The slide rack was gently raised and lowered 5 times. The magnetic stirrer was turned on at a low setting and the slides incubated for 5 minutes.

[0312] The slide rack was slowly pulled out of the ACN stain dish and placed on an absorbent towel. The bottom edges of the slides were quickly dried and the slide was placed into a clean slide box.

15. Microarray Imaging

[0313] The microarray slides were placed into Agilent scanner slide holders and loaded into the Agilent Microarray scanner according to the manufacturers instructions.

[0314] The slides were imaged in the Cy3-channel at 5 μm resolution at the 100% PMT setting and the XRD option enabled at 0.05. The resulting tiff images were processed using Agilent feature extraction software version 10.5.

Example 2. Biomarker Identification

[0315] The identification of potential mesothelioma biomarkers was performed for diagnosis of mesothelioma in individuals exposed to asbestos. Enrollment criteria for this study were age 18 or older, able to give informed consent, and blood sample and documented diagnosis of mesothelioma or benign findings. For cases, blood samples collected prior to treatment or surgery and subsequently diagnosed with mesothelioma. Exclusion criteria included prior diagnosis or treatment of cancer (excluding squamous cell carcinoma of the skin) within 5 years of the blood draw. Serum samples were collected from 4 different sites and included 158 mesothelioma samples and 140 control group samples as described in Table 17. The multiplexed aptamer affinity assay as described in Example 1 was used to measure and report the RFU

value for 1045 analytes in each of these 298 samples. Since the serum samples were obtained from 4 independent studies and sites under similar protocols, an examination of site differences prior to the analysis for biomarkers discovery was performed.

[0316] Each of the case and control populations were separately compared by generating class-dependent cumulative distribution functions (cdfs) for each of the 1045 analytes. The KS-distance (Kolmogorov-Smirnov statistic) between values from two sets of samples is a non parametric measurement of the extent to which the empirical distribution of the values from one set (Set A) differs from the distribution of values from the other set (Set B). For any value of a threshold T some proportion of the values from Set A will be less than T, and some proportion of the values from Set B will be less than T. The KS-distance measures the maximum (unsigned) difference between the proportion of the values from the two sets for any choice of T.

[0317] This set of potential biomarkers can be used to build classifiers that assign samples to either a control or disease group. In fact, many such classifiers were produced from these sets of biomarkers and the frequency with which any biomarker was used in good scoring classifiers determined. Those biomarkers that occurred most frequently among the top scoring classifiers were the most useful for creating a diagnostic test. In this example, Bayesian classifiers were used to explore the classification space but many other supervised learning techniques may be employed for this purpose. The scoring fitness of any individual classifier was gauged by the area under the receiver operating characterisic curve (AUC of the ROC) of the classifier at the Bayesian surface assuming a disease prevalence of 0.5. This scoring metric varies from zero to one, with one being an error-free classifier. The details of constructing a Bayesian classifier from biomarker population measurements are described in Example 3.

[0318] Using the 66 analytes in Table 1, a total of 925 10-analyte classifiers were found with an AUC of 0.99 for diagnosing mesothelioma from the control group. From this set of classifiers, a total of 10 biomarkers were found to be present in 30% or more of the high scoring classifiers. Table 13 provides a list of these potential biomarkers and Figure 10 is a frequency plot for the identified biomarkers.

Example 3. Naive Bayesian Classification for Mesothelioma

[0319] From the list of biomarkers identified as useful for discriminating between mesothelioma and controls, a panel of ten biomarkers was selected and a naïve Bayes classifier was constructed, see Tables 16 and 18. The class-dependent probability density functions (pdfs), $p(x_i|c)$ and $p(x_i|d)$, where x_i is the log of the measured RFU value for biomarker i, and c and d refer to the control and disease populations, were modeled as log-normal distribution functions characterized by a mean μ and variance σ^2. The parameters for pdfs of the ten biomarkers are listed in Table 16 and an example of the raw data along with the model fit to a normal pdf is displayed in Figure 5. The underlying assumption appears to fit the data quite well as evidenced by Figure 5.

[0320] The naive Bayes classification for such a model is given by the following equation, where $p(d)$ is the prevalence of the disease in the population,

$$ln\left(\frac{p(d|\tilde{x})}{p(c|\tilde{x})}\right) = \sum_{i=1}^{n} ln\left(\frac{\sigma_{c,i}}{\sigma_{d,i}}\right) - \frac{1}{2}\sum_{i=1}^{n}\left[\left(\frac{x_i - \mu_{d,i}}{\sigma_{d,i}}\right)^2 - \left(\frac{x_i - \mu_{c,i}}{\sigma_{c,i}}\right)^2\right] + ln\left(\frac{p(d)}{1-p(d)}\right)$$

appropriate to the test and n = 10. Each of the terms in the summation is a log-likelihood ratio for an individual marker and the total log-likelihood ratio of a sample $\tilde{x}$ being free from the disease of interest (i.e. in this case, mesothelioma) versus having the disease is simply the sum of these individual terms plus a term that accounts for the prevalence of the disease. For simplicity, we assume $p(d)$ = 0.5 so that $ln\left(\frac{p(d)}{1-p(d)}\right) = 0.$

[0321] Given an unknown sample measurement in *log*(RFU) for each of the ten biomarkers of 9.6, 8.0, 7.4, 7.0, 7.3, 8.9, 7.3, 8.3, 10.0, 7.3, the calculation of the classification is detailed in Table 16. The individual components comprising the log likelihood ratio for disease versus control class are tabulated and can be computed from the parameters in Table 16 and the values of $\tilde{x}$. The sum of the individual log likelihood ratios is -6.364, or a likelihood of being free from the disease versus having the disease of 581, where likelihood $e^{6.364}$ = 581. The first 1 biomarker values have likelihoods more consistent with the disease group (log likelihood > 0) but the remaining 9 biomarkers are all consistently found to favor the control group. Multiplying the likelihoods together gives the same results as that shown above; a likelihood of 581 that the unknown sample is free from the disease. In fact, this sample came from the control population in the training set.

Example 4. Greedy Algorithm for Selecting Biomarker Panels for Classifiers.

[0322] This example describes the selection of biomarkers from Table 1 to form panels that can be used as classifiers in any of the methods described herein. Subsets of the biomarkers in Table 1 were selected to construct classifiers with good performance. This method was also used to determine which potential markers were included as biomarkers in Example 2.

[0323] The measure of classifier performance used here is the AUC; a performance of 0.5 is the baseline expectation for a random (coin toss) classifier, a classifier worse than random would score between 0.0 and 0.5, a classifier with better than random performance would score between 0.5 and 1.0. A perfect classifier with no errors would have a sensitivity of 1.0 and a specificity of 1.0. One can apply the methods described in Example 4 to other common measures of performance such as the F-measure, the sum of sensitivity and specificity, or the product of sensitivity and specificity. Specifically one might want to treat sensitivity and specificity with differing weight, so as to select those classifiers which perform with higher specificity at the expense of some sensitivity, or to select those classifiers which perform with higher sensitivity at the expense of some specificity. Since the method described here only involves a measure of "performance", any weighting scheme which results in a single performance measure can be used. Different applications will have different benefits for true positive and true negative findings, and also different costs associated with false positive findings from false negative findings. For example, screening asymptomatic high risk individuals and the differential diagnosis of mesothelioma from benign pleural symptoms will not in general have the same optimal trade-off between specificity and sensitivity. The different demands of the two tests will in general require setting different weighting to positive and negative misclassifications, reflected in the performance measure. Changing the performance measure will in general change the exact subset of markers selected from Table 1 for a given set of data.

[0324] For the Bayesian approach to the discrimination of mesothelioma samples from control samples described in Example 3, the classifier was completely parameterized by the distributions of biomarkers in the disease and benign training samples, and the list of biomarkers was chosen from Table 1; that is to say, the subset of markers chosen for inclusion determined a classifier in a one-to-one manner given a set of training data.

[0325] The greedy method employed here was used to search for the optimal subset of markers from Table 1. For small numbers of markers or classifiers with relatively few markers, every possible subset of markers was enumerated and evaluated in terms of the performance of the classifier constructed with that particular set of markers (see Example 4, Part 2). (This approach is well known in the field of statistics as "best subset selection"; see, e.g., Hastie et al). However, for the classifiers described herein, the number of combinations of multiple markers can be very large, and it was not feasible to evaluate every possible set of 10 markers, as there are 30,045,015 possible combinations that can be generated from a list of only 30 total analytes. Because of the impracticality of searching through every subset of markers, the single optimal subset may not be found; however, by using this approach, many excellent subsets were found, and, in many cases, any of these subsets may represent an optimal one.

[0326] Instead of evaluating every possible set of markers, a "greedy" forward stepwise approach may be followed (see, e.g., Dabney AR, Storey JD (2007) Optimality Driven Nearest Centroid Classification from Genomic Data.

[0327] PLoS ONE 2(10): e1002. doi:10.1371/journal.pone.0001002). Using this method, a classifier is started with the best single marker (based on KS-distance for the individual markers) and is grown at each step by trying, in turn, each member of a marker list that is not currently a member of the set of markers in the classifier. The one marker which scores best in combination with the existing classifier is added to the classifier. This is repeated until no further improvement in performance is achieved. Unfortunately, this approach may miss valuable combinations of markers for which some of the individual markers are not all chosen before the process stops.

[0328] The greedy procedure used here was an elaboration of the preceding forward stepwise approach, in that, to broaden the search, rather than keeping just a single candidate classifier (marker subset) at each step, a list of candidate classifiers was kept. The list was seeded with every single marker subset (using every marker in the table on its own). The list was expanded in steps by deriving new classifiers (marker subsets) from the ones currently on the list and adding them to the list. Each marker subset currently on the list was extended by adding any marker from Table 1 not already part of that classifier, and which would not, on its addition to the subset, duplicate an existing subset (these are termed "permissible markers"). Every existing marker subset was extended by every permissible marker from the list. Clearly, such a process would eventually generate every possible subset, and the list would run out of space. Therefore, all the generated classifiers were kept only while the list was less than some predetermined size (often enough to hold all three marker subsets). Once the list reached the predetermined size limit, it became elitist; that is, only those classifiers which showed a certain level of performance were kept on the list, and the others fell off the end of the list and were lost. This was achieved by keeping the list sorted in order of classifier performance; new classifiers which were at least as good as the worst classifier currently on the list were inserted, forcing the expulsion of the current bottom underachiever. One further implementation detail is that the list was completely replaced on each generational step; therefore, every classifier on the list had the same number of markers, and at each step the number of markers per classifier grew by one.

[0329] Since this method produced a list of candidate classifiers using different combinations of markers, one may

ask if the classifiers can be combined in order to avoid errors which might be made by the best single classifier, or by minority groups of the best classifiers. Such "ensemble" and "committee of experts" methods are well known in the fields of statistical and machine learning and include, for example, "Averaging", "Voting", "Stacking", "Bagging" and "Boosting" (see, e.g., Hastie et al.). These combinations of simple classifiers provide a method for reducing the variance in the classifications due to noise in any particular set of markers by including several different classifiers and therefore information from a larger set of the markers from the biomarker table, effectively averaging between the classifiers. An example of the usefulness of this approach is that it can prevent outliers in a single marker from adversely affecting the classification of a single sample. The requirement to measure a larger number of signals may be impractical in conventional "one marker at a time" antibody assays but has no downside for a fully multiplexed aptamer assay. Techniques such as these benefit from a more extensive table of biomarkers and use the multiple sources of information concerning the disease processes to provide a more robust classification.

[0330] The biomarkers selected in Table 1 gave rise to classifiers which perform better than classifiers built with "non-markers" (i.e., proteins having signals that did not meet the criteria for inclusion in Table 1 (as described in Example 2)).

[0331] For classifiers containing only one, two, and three markers, all possible classifiers obtained using the biomarkers in Table 1 were enumerated and examined for the distribution of performance compared to classifiers built from a similar table of randomly selected non-markers signals.

[0332] In Figure 11, the AUC was used as the measure of performance; a performance of 0.5 is the baseline expectation for a random (coin toss) classifier. The histogram of classifier performance was compared with the histogram of performance from a similar exhaustive enumeration of classifiers built from a "non-marker" table of 66 non-marker signals; the 66 signals were randomly chosen from aptamers that did not demonstrate differential signaling between control and disease populations.

[0333] Figure 11 shows histograms of the performance of all possible one, two, and three-marker classifiers built from the biomarker parameters in Table 14 for biomarkers that can discriminate between the control group and mesothelioma and compares these classifiers with all possible one, two, and three-marker classifiers built using the 66 "non-marker" aptamer RFU signals. Figure 11A shows the histograms of single marker classifier performance, Figure 11B shows the histogram of two marker classifier performance, and Figure 11C shows the histogram of three marker classifier performance.

[0334] In Figure 11, the solid lines represent the histograms of the classifier performance of all one, two, and three-marker classifiers using the biomarker data for asbestos exposed individuals and mesothelioma in Table 14. The dotted lines are the histograms of the classifier performance of all one, two, and three-marker classifiers using the data for controls and mesothelioma but using the set of random non-marker signals.

[0335] The classifiers built from the markers listed in Table 1 form a distinct histogram, well separated from the classifiers built with signals from the "non-markers" for all one-marker, two-marker, and three-marker comparisons. The performance and AUC score of the classifiers built from the biomarkers in Table 1 also increase faster with the number of markers than do the classifiers built from the non-markers, the separation increases between the marker and non-marker classifiers as the number of markers per classifier increases. All classifiers built using the biomarkers listed in Table 14 perform distinctly better than classifiers built using the "non-markers".

[0336] The distributions of classifier performance show that there are many possible multiple-marker classifiers that can be derived from the set of analytes in Table 1. Although some biomarkers are better than others on their own, as evidenced by the distribution of classifier scores and AUCs for single analytes, it was desirable to determine whether such biomarkers are required to construct high performing classifiers. To make this determination, the behavior of classifier performance was examined by leaving out some number of the best biomarkers. Figure 12 compares the performance of classifiers built with the full list of biomarkers in Table 1 with the performance of classifiers built with subsets of biomarkers from Table 1 that excluded top-ranked markers.

[0337] Figure 12 demonstrates that classifiers constructed without the best markers perform well, implying that the performance of the classifiers was not due to some small core group of markers and that the changes in the underlying processes associated with disease are reflected in the activities of many proteins. Many subsets of the biomarkers in Table 1 performed close to optimally, even after removing the top 15 of the 66 markers from Table 1. After dropping the 15 top-ranked markers (ranked by KS-distance) from Table 1, the classifier performance increased with the number of markers selected from the table to reach an AUC of almost 0.97, close to the performance of the optimal classifier score of 0.993 selected from the full list of biomarkers.

[0338] Finally, Figure 13 shows how the ROC performance of typical classifiers constructed from the list of parameters in Table 14 according to Example 3. A five analyte classifier was constructed with CDH1, BMPER, F9, CCL23, and CRK. Figure 13A shows the performance of the model, assuming independence of these markers, as in Example 3, and Figure 13B shows the empirical ROC curves generated from the study data set used to define the parameters in Table 14. It can be seen that the performance for a given number of selected markers was qualitatively in agreement, and that quantitative agreement was generally quite good, as evidenced by the AUCs, although the model calculation tends to overestimate classifier performance. This is consistent with the notion that the information contributed by any particular

biomarker concerning the disease processes is redundant with the information contributed by other biomarkers provided in Table 1 while the model calculation assumes complete independence. Figure 13 thus demonstrates that Table 1 in combination with the methods described in Example 3 enable the construction and evaluation of a great many classifiers useful for the discrimination of mesothelioma from the control group.

Example 5. Biomarkers for the Diagnosis of Cancer

[0339] The identification of potential biomarkers for the general diagnosis of cancer was performed. Both case and control samples were evaluated from 3 different types of cancer (mesothelioma, lung cancer, and renal cell carcinoma). Across the sites, inclusion criteria were at least 18 years old with signed informed consent. Both cases and controls were excluded for known malignancy other than the cancer in question.

[0340] Mesothelioma. Case and control samples were obtained as described in Example 2.

[0341] **Lung Cancer.** Case and control samples were obtained from three academic cancer center biorepositories and one commercial biorepository to identify potential markers for the differential diagnosis of non-small cell lung cancer (NSCLC) from a control group of high risk smokers and individuals with benign pulmomary nodules. The study was composed of 978 sampls collected from smokers and patients with benign nodules and 320 individuals diagnosed with NSCLC.

[0342] **Renal Cell Carcinoma.** Case and control samples were obtained from an academic cancer center biorepository from patients with renal cell carcioma (RCC) and benign masses (BEN). Pre-surgical samples (TP1) were obtained for all subjects. The primary analysis compared outcome data (as recorded in the SEER database field CA Status 1) for the 38 RCC patients with "Evidence of Disease" (EVD) vs 104 with "No Evidence of Disease" (NED) documented through clinical follow-up.

[0343] A final list of cancer biomarkers was identified by combining the sets of biomarkers considered for each of the 3 different cancer studies. Bayesian classifiers that used biomarker sets of increasing size were successively constructed using a greedy algorithm (as described in greater detail in Section 5.2 of this Example). The sets (or panels) of biomarkers that were useful for diagnosing cancer in general among the different sites and types of cancer were compiled as a function of set (or panel) size and analyzed for their performance. This analysis resulted in the list of 22 cancer biomarkers shown in Table 19, each of which was present in at least one of these successive marker sets, which ranged in size from three to ten markers. As an illustrative example, we describe the generation of a specific panel composed of ten cancer biomarkers, which is shown in Table 32.

5.1 Naive Bayesian Classification for Cancer

[0344] From the list of biomarkers in Table 1, a panel of ten potential cancer biomarkers was selected using a greedy algorithm for biomarker selection, as outlined in Section 5.2 of this Example. A distinct naïve Bayes classifier was constructed for each of the 3 different cancer types. The class-dependent probability density functions (pdfs), $p(x_i|c)$ and $p(X_i|d)$, where x_i is the log of the measured RFU value for biomarker *i*, and *c* and *d* refer to the control and disease populations, were modeled as log-normal distribution functions characterized by a mean μ and variance σ^2. The parameters for pdfs of the 3 models composed of the ten potential biomarkers are listed in Table 31.

[0345] The naive Bayes classification for such a model is given by the following equation, where $p((d)$ is the prevalence of the disease in the population,

$$ln\left(\frac{p(d|\tilde{x})}{p(c|\tilde{x})}\right) = \sum_{i=1}^{n} ln\left(\frac{\sigma_{c,i}}{\sigma_{d,i}}\right) - \frac{1}{2}\sum_{i=1}^{n}\left[\left(\frac{x_i - \mu_{d,i}}{\sigma_{d,i}}\right)^2 - \left(\frac{x_i - \mu_{c,i}}{\sigma_{c,i}}\right)^2\right] + ln\left(\frac{p(d)}{1-p(d)}\right)$$

appropriate to the test and n = 10. Each of the terms in the summation is a log-likelihood ratio for an individual marker and the total log-likelihood ratio of a sample $\tilde{x}$ being free from the disease interest (i.e., in this case, each particular disease from the 3 different cancer types) versus having the disease is simply the sum of these individual terms plus a term that accounts for the prevalence of the disease. For simplicity, we assume $p(d)$ = 0.5 so that

$$ln\left(\frac{p(d)}{1-p(d)}\right) = 0.$$

[0346] Given an unknown sample measurement in *log*(RFU) for each of the ten biomarkers of 9.5, 7.0, 10.5, 11.7, 8.7, 9.3, 10.9, 9.6, 7.9, 10.6, the calculation of the classification is detailed in Table 32. The individual components

comprising the log likelihood ratio for disease versus control class are tabulated and can be computed from the parameters in Table 31 and the values of $\tilde{x}$. The sum of the individual log likelihood ratios is -4.020, or a likelihood of being free from the disease versus having the disease of 56, where likelihood $e^{4.020} = 56$. Only 2 of the biomarker values have likelihoods more consistent with the disease group (log likelihood > 0) but the remaining 8 biomarkers are all consistently found to favor the control group. Multiplying the likelihoods together gives the same results as that shown above; a likelihood of 56 that the unknown sample is free from the disease. In fact, this sample came from the control population in the NSCLC training set.

5.2 Greedy Algorithm for Selecting Cancer Biomarker Panels for Classifiers

Part 1

[0347] Subsets of the biomarkers in Table 1 were selected to construct potential classifiers that could be used to determine which of the markers could be used as general cancer biomarkers to detect cancer.

[0348] Given a set of markers, a distinct model was trained for each of the 3 cancer studies, so a global measure of performance was required to select a set of biomarkers that was able to classify simultaneously many different types of cancer. The measure of classifier performance used here was the mean of the area under ROC curve across all naive Bayes classifiers. The ROC curve is a plot of a single classifier true positive rate (sensitivity) versus the false positive rate (1-specificity). The area under the ROC curve (AUC) ranges from 0 to 1.0, where an AUC of 1.0 corresponds to perfect classification and an AUC of 0.5 corresponds to random (coin toss) classifier. One can apply other common measures of performance such as the F-measure or the sum or product of sensitivity and specificity. Specifically, one might want to treat sensitivity and specificity with differing weight, in order to select those classifiers that perform with higher specificity at the expense of some sensitivity, or to select those classifiers which perform with higher sensitivity at the expense of specificity. We chose to use the AUC because it encompasses all combinations of sensitivity and specificity in a single measure. Different applications will have different benefits for true positive and true negative findings, and will have different costs associated with false positive findings from false negative findings. Changing the performance measure may change the exact subset of markers selected for a given set of data.

[0349] For the Bayesian approach to the discrimination of cancer samples from control samples described in Section 5.1 of this Example, the classifier was completely parameterized by the distributions of biomarkers in each of the 3 cancer studies, and the list of biomarkers was chosen from Table 19. That is to say, the subset of markers chosen for inclusion determined a classifier in a one-to-one manner given a set of training data.

[0350] The greedy method employed here was used to search for the optimal subset of markers from Table 1. For small numbers of markers or classifiers with relatively few markers, every possible subset of markers was enumerated and evaluated in terms of the performance of the classifier constructed with that particular set of markers (see Example 4, Part 2). (This approach is well known in the field of statistics as "best subset selection"; see, e.g., Hastie et al). However, for the classifiers described herein, the number of combinations of multiple markers can be very large, and it was not feasible to evaluate every possible set of 10 markers, as there are 30,045,015 possible combinations that can be generated from a list of only 30 total analytes. Because of the impracticality of searching through every subset of markers, the single optimal subset may not be found; however, by using this approach, many excellent subsets were found, and, in many cases, any of these subsets may represent an optimal one.

[0351] Instead of evaluating every possible set of markers, a "greedy" forward stepwise approach may be followed (see, e.g., Dabney AR, Storey JD (2007) Optimality Driven Nearest Centroid Classification from Genomic Data. PLoS ONE 2(10): e1002. doi:10.1371/journal.pone.0001002). Using this method, a classifier is started with the best single marker (based on KS-distance for the individual markers) and is grown at each step by trying, in turn, each member of a marker list that is not currently a member of the set of markers in the classifier. The one marker that scores the best in combination with the existing classifier is added to the classifier. This is repeated until no further improvement in performance is achieved. Unfortunately, this approach may miss valuable combinations of markers for which some of the individual markers are not all chosen before the process stops.

[0352] The greedy procedure used here was an elaboration of the preceding forward stepwise approach, in that, to broaden the search, rather than keeping just a single marker subset at each step, a list of candidate marker sets was kept. The list was seeded with a list of single markers. The list was expanded in steps by deriving new marker subsets from the ones currently on the list and adding them to the list. Each marker subset currently on the list was extended by adding any marker from Table 1 not already part of that classifier, and which would not, on its addition to the subset, duplicate an existing subset (these are termed "permissible markers"). Each time a new set of markers was defined, a set of classifiers composed of one for each cancer study was trained using these markers, and the global performance was measured via the mean AUC across all 3 studies. To avoid potential over fitting, the AUC for each cancer study model was calculated via a ten-fold cross validation procedure. Every existing marker subset was extended by every permissible marker from the list. Clearly, such a process would eventually generate every possible subset, and the list

would run out of space. Therefore, all the generated marker sets were kept only while the list was less than some predetermined size. Once the list reached the predetermined size limit, it became elitist; that is, only those classifier sets which showed a certain level of performance were kept on the list, and the others fell off the end of the list and were lost. This was achieved by keeping the list sorted in order of classifier set performance; new marker sets whose classifiers were globally at least as good as the worst set of classifiers currently on the list were inserted, forcing the expulsion of the current bottom underachieving classifier sets. One further implementation detail is that the list was completely replaced on each generational step; therefore, every marker set on the list had the same number of markers, and at each step the number of markers per classifier grew by one.

[0353] In one embodiment, the set (or panel) of biomarkers useful for constructing classifiers for diagnosing general cancer from non-cancer is based on the mean AUC for the particular combination of biomarkers used in the classification scheme. We identified many combinations of biomarkers derived from the markers in Table 19 that were able to effectively classify different cancer samples from controls. Representative panels are set forth in Tables 22-29, which set forth a series of 100 different panels of 3-10 biomarkers, which have the indicated mean cross validation (CV) AUC for each panel. The total number of occurrences of each marker in each of these panels is indicated at the bottom of each table.

[0354] The biomarkers selected in Table 19 gave rise to classifiers that perform better than classifiers built with "non-markers." In Figure 14, we display the performance of our ten biomarker classifiers compared to the performance of other possible classifiers.

[0355] Figure 14A shows the distribution of mean AUCs for classifiers built from randomly sampled sets of ten "non-markers" taken from the entire set of 22 present in all 3 studies, excluding the ten markers in Table 19. The performance of the ten potential cancer biomarkers is displayed as a vertical dashed line. This plot clearly shows that the performance of the ten potential biomarkers is well beyond the distribution of other marker combinations.

[0356] Figure 14B displays a similar distribution as Figure 14A, however the randomly sampled sets were restricted to the 56 biomarkers from Table 1 that were not selected by the greedy biomarker selection procedure for ten analyte classifiers. This plot demonstrates that the ten markers chosen by the greedy algorithm represent a subset of biomarkers that generalize to other types of cancer far better than classifiers built with the remaining 56 biomarkers.

[0357] Finally, Figure 15 shows the classifier ROC curve for each of the 3 cancer studies classifiers. The foregoing embodiments and examples are intended only as examples. No particular embodiment, example, or element of a particular embodiment or example is to be construed as a critical, required, or essential element or feature of any of the claims. Further, no element described herein is required for the practice of the appended claims unless expressly described as "essential" or "critical." Various alterations, modifications, substitutions, and other variations can be made to the disclosed embodiments without departing from the scope of the present application, which is defined by the appended claims. The specification, including the figures and examples, is to be regarded in an illustrative manner, rather than a restrictive one, and all such modifications and substitutions are intended to be included within the scope of the application. Accordingly, the scope of the application should be determined by the appended claims and their legal equivalents, rather than by the examples given above. For example, steps recited in any of the method or process claims may be executed in any feasible order and are not limited to an order presented in any of the embodiments, the examples, or the claims. Further, in any of the aforementioned methods, one or more biomarkers of Table 1 or Table 19 can be specifically excluded either as an individual biomarker or as a biomarker from any panel.

Table 1: Cancer Biomarkers

Column #1	Column #2	Column #3	Column #4	Column #5	Column #6
Biomarker #	Biomarker Designation Entrez Gene Symbol(s)	Entrez Gene ID	SwissProt ID	Public Name	Direction
1	ABL1	25	P00519	ABL1	Down
2	AFM	173	P43652	Afamin	Down
3	ALB	213	P02768	Albumin	Down
4	ALPL	249	P05186	Alkaline phosphatase, bone	Up
5	APOA1	335	P02647	Apo A-I	Down
6	AZU1	566	P20160	Azurocidin	Up
7	BDNF	627	P23560	BDNF	Down
8	BMP1	649	P13497	BMP-1	Down

(continued)

Column #1	Column #2	Column #3	Column #4	Column #5	Column #6
Biomarker #	Biomarker Designation Entrez Gene Symbol(s)	Entrez Gene ID	SwissProt ID	Public Name	Direction
9	BMPER	168667	Q8N8U9	BMPER	Down
10	BMX	660	P51813	BMX	Down
11	BPI	671	P17213	BPI	Up
12	C9	735	P02748	C9	Up
13	CAMK1	8536	Q14012	CAMK1	Up
14	CCDC80	151887	Q76M96	URB	Up
15	CCL23	6368	P55773	MPIF-1	Up
16	CCL23	6368	P55773	Ck-β-8-1	Up
17	CDH1	999	P12830	Cadherin-1	Down
18	CDK5-CDK5R1	1020; 1775	Q00535; Q15078	CDK5/p35	Up
19	CDK8-CCNC	1024; 892	P49336; P24863	CDK8/cyclin C	Up
20	CFHR5	81494	Q9BXR6	complement factor H-related 5	Up
21	CFL1	1072	P23528	Cofilin-1	Up
22	CFP	5199	P27918	Properdin	Down
23	CRK	1398	P46108	adaptor protein Crk	Up
24	CRP	1401	P02741	CRP	Up
25	CSN1S1	1446	P47710	Alpha-S1-casein	Down
26	CXCL13	10563	043927	BCA-1	Up
27	DDC	1644	P20711	dopa decarboxylase	Down
28	EFNA5	1946	P52803	Ephrin-A5	Up
29	EGFR	1956	P00533	ERBB1	Down
30	EIF4EBP2	1979	Q13542	eIF4E-binding protein 2	Down
31	ESM1	11082	Q9NQ30	Endocan	Up
32	F9	2158	P00740	Coagulation Factor IX	Up
33	FCN2	2220	Q15485	Ficolin-2	Up
34	FGA-FGB-FGG	2243; 2244; 2266	P02671; P02675; P02679	D-dimer	Up
35	FLT3LG	2323	P49771	Flt-3 ligand	Up
36	FN1	2335	P02751	Fibronectin FN1.4	Down
37	FN1	2335	P02751	Fibronectin	Down
38	FRZB	2487	Q92765	FRP-3, soluble	Up
39	GPC2	221914	Q8N158	Glypican 2	Down

(continued)

Column #1	Column #2	Column #3	Column #4	Column #5	Column #6
Biomarker #	Biomarker Designation Entrez Gene Symbol(s)	Entrez Gene ID	SwissProt ID	Public Name	Direction
40	GPI	2821	P06744	glucose phosphate isomerase	Up
41	H2AFZ	3015	P0C0S5	Histone H2A.z	Up
42	HINT1	3094	P49773	HINT1	Down
43	ICAM2	3384	P13598	ICAM-2, soluble	Down
44	IL31	386653	Q6EBC2	IL-31	Down
45	ITGA1-ITGB1-	3672; 3688	P56199; P05556	Integrin α1/β1	Up
46	ITIH4	3700	Q14624	Inter-α-trypsin inhibitor heavy chain H4	Up
47	KIT	3815	P10721	SCF sR	Down
48	KLK3-SERPINA3	354; 12	P07288; P01011	PSA-ACT	Up
49	LCN2	3934	P80188	Lipocalin 2	Up
50	LTF	4057	P02788	Lactoferrin	Up
51	MDK	4192	P21741	Midkine	Up
52	MMP9	4318	P14780	MMP-9	Up
53	MPO	4353	P05164	Myeloperoxidase	Up
54	MSLN	10232	Q13421	Mesothelin	Down
55	PLA2G5	5322	P39877	Group V phospholipase A2	Down
56	PRTN3	5657	P24158	Proteinase-3	Up
57	RBP4	5950	P02753	RBP	Down
58	SAA1	6288	P02735	SAA	Up
59	SERPINA4	5267	P29622	Kallistatin	Down
60	TGFB2	7042	P61812	TGF-β2	Down
61	TIMP1	7076	P01033	TIMP-1	Up
62	TNFRSF4	7293	P43489	TNR4	Down
63	TNFRSF8	943	P28908	CD30	Up
64	TPT1	7178	P13693	Fortilin	Up
65	VEGFA	7422	P15692	VEGF	Up
66	YWHAH	7533	Q04917	14-3-3 protein eta	Up

Table 2: Panels of 1 Biomarker

Markers		CV AUC
1	CDH1	0.880
2	BMPER	0.859

(continued)

Markers		CV AUC
3	KLK3-SERPINA3	0.856
4	C9	0.837
5	PLA2G5	0.826
6	CRK	0.814
7	BMX	0.807
8	VEGFA	0.806
9	F9	0.806
10	AFM	0.805
11	CCL23	0.803
12	SERPINA4	0.803
13	GPC2	0.802
14	ABL1	0.802
15	APOA1	0.796
16	IL31	0.795
17	CDK8-CCNC	0.795
18	KIT	0.789
19	FCN2	0.786
20	HINT1	0.786
21	CAMK1	0.782
22	TGFB2	0.780
23	SAA1	0.780
24	CSN1S1	0.779
25	CXCL13	0.777
26	CFL1	0.777
27	TPT1	0.776
28	CRP	0.775
29	MSLN	0.773
30	FLT3LG	0.773
31	FN1	0.773
32	ITGA1-ITGB1	0.772
33	CFP	0.772
34	TNFRSF4	0.770
35	GPI	0.768
36	BMP1	0.768
37	CCL23	0.764
38	ALB	0.762
39	DDC	0.759
40	EGFR	0.758

(continued)

Markers		CV AUC
41	BDNF	0.757
42	CFHR5	0.753
43	H2AFZ	0.747
44	ITIH4	0.747
45	EIF4EBP2	0.746
46	RBP4	0.745
47	ESM1	0.744
48	FN1	0.741
49	YWHAH	0.738
50	FRZB	0.733
51	EFNA5	0.731
52	FGA-FGB-FGG	0.729
53	CCDC80	0.727
54	TIMP1	0.722
55	CDK5-CDK5R1	0.692
56	MDK	0.680
57	BPI	0.646
58	AZU1	0.637
59	TNFRSF8	0.628
60	ICAM2	0.624
61	PRTN3	0.612
62	LTF	0.605
63	MMP9	0.593
64	ALPL	0.591
65	MPO	0.589
66	LCN2	0.573

Table 3: Panels of 2 Biomarkers

Markers			CV AUC
1	CDH1	BMPER	0.945
2	CDH1	F9	0.932
3	BMPER	CRK	0.920
4	BMPER	TPT1	0.919
5	CDH1	CCL23	0.918
6	BMPER	TGFB2	0.916
7	CDH1	FRZB	0.914
8	CDH1	ABL1	0.912

(continued)

Markers			CV AUC
9	KLK3-SERPINA3	CDH1	0.912
10	CCL23	CRK	0.911
11	CDH1	CCL23	0.911
12	CCL23	TPT1	0.911
13	CCL23	YWHAH	0.910
14	BMPER	YWHAH	0.910
15	CDH1	VEGFA	0.910
16	CDH1	AFM	0.908
17	PLA2G5	CDH1	0.908
18	CDH1	SERPINA4	0.908
19	CDH1	SAA1	0.907
20	CDH1	CCDC80	0.907
21	CDH1	CRP	0.907
22	CCDC80	BMPER	0.907
23	CDH1	CRK	0.907
24	CDH1	CSN1S1	0.905
25	CDH1	FCN2	0.904
26	CDH1	BMX	0.904
27	KIT	CDH1	0.904
28	CDH1	RBP4	0.904
29	BMPER	CFL1	0.904
30	CDH1	CXCL13	0.903
31	KLK3-SERPINA3	BMPER	0.903
32	CDH1	CFL1	0.902
33	CDH1	MDK	0.902
34	CDH1	C9	0.900
35	CDH1	APOA1	0.899
36	CDH1	ITIH4	0.899
37	ESM1	CRK	0.899
38	CDH1	YWHAH	0.898
39	CCL23	GPI	0.897
40	BDNF	CDH1	0.897
41	CDH1	FN1	0.897
42	CDH1	EFNA5	0.897
43	VEGFA	BMPER	0.896
44	CCL23	CFL1	0.896
45	KLK3-SERPINA3	GPC2	0.896
46	CDH1	BMP1	0.895

(continued)

Markers			CV AUC
47	AZU1	BMPER	0.895
48	KLK3-SERPINA3	F9	0.895
49	KLK3-SERPINA3	CCL23	0.894
50	CCL23	H2AFZ	0.894
51	CDH1	TGFB2	0.894
52	C9	FCN2	0.894
53	CDH1	CFP	0.893
54	CDH1	CFHR5	0.893
55	BMPER	HINT1	0.893
56	CDH1	TPT1	0.892
57	KLK3-SERPINA3	FCN2	0.892
58	CDH1	GPI	0.892
59	CDH1	CDK8-CCNC	0.891
60	BMPER	GPI	0.891
61	KLK3-SERPINA3	PLA2G5	0.891
62	CDH1	EGFR	0.891
63	TIMP1	CDH1	0.891
64	CCL23	BMPER	0.890
65	CDH1	GPC2	0.890
66	VEGFA	FCN2	0.890
67	C9	BMPER	0.889
68	BMPER	AFM	0.889
69	CCL23	FCN2	0.889
70	VEGFA	CRK	0.889
71	CDH1	ALB	0.889
72	KLK3-SERPINA3	ESM1	0.888
73	BMPER	BPI	0.888
74	FRZB	C9	0.888
75	ESM1	TPT1	0.887
76	C9	CRK	0.887
77	CDH1	DDC	0.887
78	PLA2G5	AFM	0.886
79	CDH1	ESM1	0.885
80	CDH1	ICAM2	0.885
81	CDH1	TNFRSF4	0.885
82	CDH1	CAMK1	0.885
83	KLK3-SERPINA3	TNFRSF4	0.885
84	PRTN3	BMPER	0.885

(continued)

Markers			CV AUC
85	ABL1	BMPER	0.885
86	KLK3-SERPINA3	EFNA5	0.885
87	F9	CRK	0.885
88	KLK3-SERPINA3	FRZB	0.884
89	CDH1	ITGA1-ITGB1	0.884
90	KLK3-SERPINA3	MDK	0.884
91	ICAM2	BMPER	0.883
92	PLA2G5	C9	0.883
93	CDH1	HINT1	0.883
94	RBP4	BMPER	0.883
95	CCL23	TGFB2	0.883
96	CDH1	FLT3LG	0.883
97	KIT	BMPER	0.883
98	PLA2G5	BMPER	0.882
99	CDH1	FN1	0.882
100	ITGA1-ITGB1	BMPER	0.882

Table 4: Panels of 3 Biomarkers

Markers				CV AUC
1	CDH1	F9	CRK	0.958
2	CDH1	FRZB	BMPER	0.955
3	CDH1	BMPER	CRK	0.954
4	CDH1	BMPER	TPT1	0.954
5	KLK3-SERPINA3	CDH1	F9	0.954
6	CDH1	TPT1	F9	0.951
7	CDH1	ICAM2	BMPER	0.950
8	CDH1	CCDC80	BMPER	0.950
9	CDH1	BMPER	YWHAH	0.949
10	CDH1	CFL1	F9	0.949
11	CDH1	CCL23	BMPER	0.947
12	CDH1	CCL23	CRK	0.947
13	CDH1	BMPER	F9	0.947
14	CDH1	BMPER	CFL1	0.946
15	CDH1	FRZB	CCL23	0.946
16	CDH1	ABL1	BMPER	0.946
17	CDH1	CAMK1	F9	0.946
18	CDH1	SERPINA4	F9	0.945

(continued)

Markers				CV AUC
19	KIT	CDH1	BMPER	0.945
20	CDH1	VEGFA	BMPER	0.945
21	CDH1	BMPER	AFM	0.945
22	CDH1	AFM	F9	0.945
23	CDH1	BMPER	TGFB2	0.945
24	CDH1	AZU1	BMPER	0.944
25	CDH1	EFNA5	F9	0.944
26	BDNF	CDH1	FRZB	0.944
27	CDH1	RBP4	BMPER	0.944
28	CDH1	BMPER	BPI	0.943
29	CDH1	MMP9	BMPER	0.943
30	CDH1	CCL23	TPT1	0.943
31	CDH1	EGFR	F9	0.942
32	CDH1	PRTN3	BMPER	0.942
33	CDH1	LTF	BMPER	0.942
34	CDH1	CCDC80	F9	0.941
35	CDH1	RBP4	F9	0.941
36	CDH1	BMPER	SAA1	0.940
37	CDH1	BMPER	CRP	0.940
38	CDH1	C9	F9	0.940
39	CDH1	CCL23	F9	0.940
40	CDH1	CCL23	F9	0.940
41	CDH1	MPO	BMPER	0.940
42	CDH1	ALPL	BMPER	0.939
43	CDH1	FRZB	SAA1	0.939
44	CDH1	CFP	BMPER	0.939
45	KLK3-SERPINA3	CDH1	BMPER	0.939
46	CDH1	BMPER	ITIH4	0.939
47	CCL23	BMPER	TPT1	0.939
48	CDH1	SERPINA4	BMPER	0.939
49	CDH1	TGFB2	F9	0.939
50	CDH1	CRP	F9	0.939
51	CDH1	CCL23	YWHAH	0.939
52	CDH1	FRZB	AFM	0.939
53	CXCL13	BMPER	CRK	0.939
54	CDH1	BMP1	F9	0.939
55	CDH1	VEGFA	FRZB	0.938
56	CDH1	CFP	F9	0.938

(continued)

Markers				CV AUC
57	CCL23	BMPER	CRK	0.938
58	CDH1	MDK	F9	0.938
59	CDH1	FRZB	CRP	0.938
60	CDH1	FN1	BMPER	0.938
61	CDH1	BMPER	CSN1S1	0.937
62	CDH1	MDK	BMPER	0.937
63	CDH1	BMPER	GPI	0.937
64	CDH1	FRZB	CCL23	0.937
65	VEGFA	BMPER	CRK	0.937
66	CDH1	CCL23	BMPER	0.937
67	CXCL13	BMPER	TPT1	0.936
68	CDH1	BMPER	TNFRSF4	0.936
69	PLA2G5	CDH1	BMPER	0.936
70	CDH1	YWHAH	F9	0.936
71	CDH1	BMP1	BMPER	0.936
72	CDH1	ABL1	F9	0.936
73	CDH1	EFNA5	BMPER	0.936
74	CDH1	C9	FCN2	0.935
75	CDH1	FRZB	CFP	0.935
76	CDH1	VEGFA	CRK	0.935
77	CDH1	LCN2	BMPER	0.935
78	PLA2G5	CDH1	F9	0.935
79	CDH1	VEGFA	F9	0.935
80	CDH1	CXCL13	BMPER	0.935
81	CDH1	FRZB	F9	0.934
82	CDH1	SAA1	F9	0.934
83	MDK	BMPER	CRK	0.934
84	CDH1	CCL23	ABL1	0.934
85	FRZB	BMPER	CRK	0.934
86	BDNF	CDH1	FCN2	0.933
87	CDH1	BMPER	EIF4EBP2	0.933
88	CDH1	VEGFA	FCN2	0.933
89	VEGFA	BMPER	TPT1	0.933
90	CDH1	BMX	BMPER	0.933
91	CDH1	CCL23	TGFB2	0.933
92	CDH1	AZU1	F9	0.932
93	KIT	CDH1	CCL23	0.932
94	KLK3-SERPINA3	CDH1	FCN2	0.932

(continued)

Markers				CV AUC
95	CDH1	SERPINA4	CRK	0.932
96	CDH1	FRZB	C9	0.932
97	CDH1	ESM1	F9	0.932
98	CDH1	FRZB	BMP1	0.932
99	CDH1	BMPER	CFHR5	0.932
100	CCL23	BMPER	CRK	0.932

Table 5: Panels of 4 Biomarkers

Markers					CV AUC
1	CDH1	EGFR	F9	CRK	0.970
2	CDH1	BMPER	F9	CRK	0.968
3	CDH1	FRZB	BMPER	CRK	0.968
4	CDH1	CCL23	F9	CRK	0.968
5	KLK3-SERPINA3	CDH1	F9	CRK	0.966
6	CDH1	BMP1	F9	CRK	0.966
7	CDH1	MDK	F9	CRK	0.965
8	CDH1	EFNA5	F9	CRK	0.965
9	CDH1	EGFR	TPT1	F9	0.965
10	CDH1	CCL23	F9	CRK	0.965
11	CDH1	SERPINA4	F9	CRK	0.964
12	CDH1	FRZB	BMPER	TPT1	0.964
13	CDH1	CDK5-CDK5R1	F9	CRK	0.963
14	BDNF	CDH1	FRZB	BMPER	0.963
15	CDH1	AFM	F9	CRK	0.963
16	CDH1	BMPER	TPT1	F9	0.963
17	CDH1	CCL23	TPT1	F9	0.962
18	CDH1	FRZB	CCL23	CRK	0.962
19	CDH1	CFP	F9	CRK	0.962
20	BDNF	CDH1	F9	CRK	0.962
21	CDH1	CCDC80	F9	CRK	0.962
22	CDH1	VEGFA	F9	CRK	0.961
23	CDH1	BMP1	TPT1	F9	0.961
24	CDH1	FRZB	CCL23	BMPER	0.961
25	CDH1	C9	F9	CRK	0.961
26	KLK3-SERPINA3	CDH1	EGFR	F9	0.961
27	CDH1	ESM1	F9	CRK	0.961
28	CDH1	FRZB	BMPER	SAA1	0.961

(continued)

Markers					CV AUC
29	BDNF	KIT	CDH1	FRZB	0.961
30	CDH1	FRZB	CFP	BMPER	0.960
31	CDH1	TGFB2	F9	CRK	0.960
32	KLK3-SERPINA3	CDH1	BMPER	F9	0.960
33	KLK3-SERPINA3	CDH1	TPT1	F9	0.960
34	CDH1	CCL23	BMPER	CRK	0.960
35	CDH1	VEGFA	BMPER	CRK	0.960
36	KLK3-SERPINA3	CDH1	MDK	F9	0.960
37	CDH1	SAA1	F9	CRK	0.960
38	CDH1	FRZB	CCDC80	BMPER	0.960
39	CDH1	CDK8-CCNC	F9	CRK	0.959
40	CDH1	FRZB	F9	CRK	0.959
41	CDH1	MMP9	F9	CRK	0.959
42	KLK3-SERPINA3	CDH1	EFNA5	F9	0.959
43	CDH1	FN1	BMPER	CRK	0.959
44	CDH1	CAMK1	F9	CRK	0.959
45	CDH1	FRZB	BMPER	AFM	0.959
46	CDH1	ICAM2	BMPER	CRK	0.959
47	KLK3-SERPINA3	CDH1	TGFB2	F9	0.959
48	CDH1	ICAM2	F9	CRK	0.959
49	CDH1	TNFRSF4	F9	CRK	0.959
50	CDH1	RBP4	F9	CRK	0.959
51	CDH1	FRZB	CCL23	TPT1	0.959
52	CDH1	MDK	TPT1	F9	0.959
53	CDH1	SERPINA4	TPT1	F9	0.958
54	BDNF	CDH1	FRZB	CRK	0.958
55	CDH1	TPT1	AFM	F9	0.958
56	CDH1	EFNA5	CFL1	F9	0.958
57	BDNF	CDH1	BMPER	CRK	0.958
58	CDH1	EFNA5	TPT1	F9	0.958
59	CDH1	ICAM2	BMPER	TPT1	0.958
60	CDH1	CCL23	BMPER	CRK	0.958
61	CDH1	FRZB	BMPER	TGFB2	0.958
62	CDH1	BMPER	CFL1	F9	0.958
63	CDH1	EGFR	TGFB2	F9	0.958
64	CDH1	FRZB	CCL23	BMPER	0.958
65	BDNF	CDH1	FRZB	ABL1	0.958
66	CDH1	CCL23	TPT1	F9	0.958

(continued)

Markers					CV AUC
67	CDH1	CFL1	F9	CRK	0.958
68	CDH1	CFHR5	F9	CRK	0.958
69	CDH1	CCL23	BMPER	TPT1	0.957
70	CDH1	MDK	BMPER	CRK	0.957
71	CDH1	FRZB	ICAM2	BMPER	0.957
72	CDH1	FRZB	BMP1	BMPER	0.957
73	KLK3-SERPINA3	CDH1	CCL23	F9	0.957
74	KIT	CDH1	BMPER	CRK	0.957
75	KIT	CDH1	FRZB	BMPER	0.957
76	CDH1	CRP	F9	CRK	0.957
77	CDH1	CCDC80	TPT1	F9	0.957
78	CDH1	FRZB	ABL1	BMPER	0.957
79	CDH1	CFP	BMPER	CRK	0.957
80	CDH1	CCL23	CFL1	F9	0.957
81	CDH1	ABL1	F9	CRK	0.957
82	CDH1	BMPER	AFM	CRK	0.957
83	CDH1	FRZB	BMPER	CRP	0.957
84	KLK3-SERPINA3	CDH1	CFP	F9	0.957
85	CDH1	FN1	F9	CRK	0.957
86	BDNF	CDH1	FRZB	TPT1	0.957
87	KIT	CDH1	F9	CRK	0.957
88	KLK3-SERPINA3	CDH1	BMP1	F9	0.957
89	CDH1	EGFR	CFL1	F9	0.957
90	CDH1	BPI	F9	CRK	0.957
91	CDH1	TNFRSF8	F9	CRK	0.956
92	PLA2G5	CDH1	F9	CRK	0.956
93	CDH1	BMPER	SAA1	CRK	0.956
94	CDH1	BMP1	BMPER	CRK	0.956
95	CDH1	VEGFA	FRZB	BMPER	0.956
96	CDH1	CCL23	BMPER	TPT1	0.956
97	CDH1	FN1	F9	CRK	0.956
98	CDH1	CCDC80	BMPER	CRK	0.956
99	CDH1	CFP	TPT1	F9	0.956
100	CDH1	BMP1	BMPER	TPT1	0.956

Table 6: Panels of 5 Biomarkers

Markers						CV AUC
1	BDNF	CDH1	FRZB	BMPER	CRK	0.980
2	BDNF	CDH1	FRZB	BMPER	TPT1	0.977
3	CDH1	CCL23	BMP1	F9	CRK	0.977
4	CDH1	EGFR	MDK	F9	CRK	0.976
5	KLK3-SERPINA3	CDH1	MDK	F9	CRK	0.976
6	CDH1	CCL23	BMPER	F9	CRK	0.976
7	CDH1	EGFR	FRZB	F9	CRK	0.975
8	CDH1	EGFR	CCL23	F9	CRK	0.974
9	KLK3-SERPINA3	CDH1	EGFR	F9	CRK	0.974
10	CDH1	FRZB	CCL23	BMPER	CRK	0.974
11	CDH1	EGFR	SERPINA4	F9	CRK	0.974
12	CDH1	BMP1	BMPER	F9	CRK	0.974
13	KLK3-SERPINA3	CDH1	CCL23	F9	CRK	0.973
14	CDH1	MDK	BMP1	F9	CRK	0.973
15	CDH1	FRZB	FN1	BMPER	CRK	0.973
16	CDH1	MDK	BMPER	F9	CRK	0.973
17	CDH1	EGFR	TGFB2	F9	CRK	0.973
18	CDH1	EFNA5	BMPER	F9	CRK	0.973
19	CDH1	EGFR	CCL23	F9	CRK	0.973
20	CDH1	CCL23	FN1	F9	CRK	0.973
21	KLK3-SERPINA3	CDH1	BMPER	F9	CRK	0.973
22	CDH1	EGFR	BMPER	F9	CRK	0.973
23	CDH1	EFNA5	EGFR	F9	CRK	0.973
24	CDH1	EGFR	CCDC80	F9	CRK	0.973
25	CDH1	EGFR	TNFRSF4	F9	CRK	0.972
26	CDH1	FRZB	BMP1	BMPER	CRK	0.972
27	CDH1	CCL23	TNFRSF4	F9	CRK	0.972
28	CDH1	ICAM2	CCL23	F9	CRK	0.972
29	CDH1	EGFR	AFM	F9	CRK	0.972
30	CDH1	FRZB	CFP	BMPER	CRK	0.972
31	CDH1	FRZB	CCL23	F9	CRK	0.972
32	CDH1	FRZB	BMPER	F9	CRK	0.972
33	CDH1	CCL23	CCL23	F9	CRK	0.972
34	CDH1	MDK	CCL23	F9	CRK	0.972
35	CDH1	CCDC80	BMPER	F9	CRK	0.971
36	CDH1	CCL23	CCDC80	F9	CRK	0.971
37	CDH1	BMPER	AFM	F9	CRK	0.971
38	CDH1	CCL23	RBP4	F9	CRK	0.971

(continued)

Markers						CV AUC
39	CDH1	EGFR	ICAM2	F9	CRK	0.971
40	CDH1	EGFR	RBP4	F9	CRK	0.971
41	CDH1	CCL23	FN1	F9	CRK	0.971
42	CDH1	FRZB	BMP1	BMPER	TPT1	0.971
43	CDH1	EGFR	FRZB	BMPER	CRK	0.971
44	CDH1	CCL23	AFM	F9	CRK	0.971
45	CDH1	EGFR	TPT1	TGFB2	F9	0.971
46	CDH1	ICAM2	BMPER	F9	CRK	0.971
47	CDH1	CCL23	BMP1	TPT1	F9	0.971
48	CDH1	CCL23	C9	F9	CRK	0.971
49	CDH1	SERPINA4	BMPER	F9	CRK	0.971
50	CDH1	CCL23	BMPER	F9	CRK	0.971
51	CDH1	EGFR	BMP1	F9	CRK	0.971
52	CDH1	MDK	SERPINA4	F9	CRK	0.971
53	BDNF	CDH1	FRZB	BMPER	TGFB2	0.970
54	CDH1	MMP9	EGFR	F9	CRK	0.970
55	CDH1	EGFR	ABL1	F9	CRK	0.970
56	CDH1	BMP1	SERPINA4	F9	CRK	0.970
57	CDH1	RBP4	BMPER	F9	CRK	0.970
58	CDH1	BMP1	BMPER	TPT1	F9	0.970
59	CDH1	EGFR	CFP	F9	CRK	0.970
60	CDH1	FRZB	CCL23	BMPER	TPT1	0.970
61	BDNF	CDH1	EGFR	F9	CRK	0.970
62	CDH1	CCL23	TPT1	F9	CRK	0.970
63	CDH1	FRZB	BMP1	F9	CRK	0.970
64	CDH1	EFNA5	MDK	F9	CRK	0.970
65	CDH1	FRZB	BMPER	SAA1	CRK	0.970
66	CDH1	EGFR	TPT1	F9	CRK	0.970
67	CDH1	CCL23	SAA1	F9	CRK	0.970
68	CDH1	VEGFA	FRZB	BMPER	CRK	0.970
69	BDNF	KIT	CDH1	FRZB	BMPER	0.970
70	CDH1	VEGFA	EGFR	F9	CRK	0.970
71	CDH1	CCL23	TGFB2	F9	CRK	0.970
72	CDH1	EFNA5	CCL23	F9	CRK	0.970
73	CDH1	BMPER	TGFB2	F9	CRK	0.970
74	CDH1	VEGFA	BMPER	F9	CRK	0.970
75	CDH1	CDK5-CDK5R1	FRZB	F9	CRK	0.970
76	CDH1	CFP	BMPER	F9	CRK	0.970

(continued)

Markers						CV AUC
77	CDH1	EGFR	SAA1	F9	CRK	0.970
78	BDNF	CDH1	BMPER	F9	CRK	0.970
79	BDNF	CDH1	VEGFA	FRZB	CRK	0.970
80	KLK3-SERPINA3	CDH1	FN1	F9	CRK	0.970
81	CDH1	BMP1	RBP4	F9	CRK	0.970
82	KLK3-SERPINA3	CDH1	EGFR	TPT1	F9	0.970
83	CDH1	EFNA5	BMP1	F9	CRK	0.970
84	CDH1	CCL23	CFL1	F9	CRK	0.970
85	CDH1	FRZB	BMPER	AFM	CRK	0.970
86	KLK3-SERPINA3	CDH1	CFP	F9	CRK	0.969
87	CDH1	ICAM2	BMP1	F9	CRK	0.969
88	CDH1	EGFR	BPI	F9	CRK	0.969
89	CDH1	C9	BMPER	F9	CRK	0.969
90	CDH1	MDK	C9	F9	CRK	0.969
91	BDNF	CDH1	MDK	F9	CRK	0.969
92	CDH1	BMP1	AFM	F9	CRK	0.969
93	BDNF	CDH1	FRZB	F9	CRK	0.969
94	BDNF	CDH1	AZU1	FRZB	BMPER	0.969
95	KLK3-SERPINA3	CDH1	BMPER	TPT1	F9	0.969
96	CDH1	CCL23	ESM1	F9	CRK	0.969
97	KLK3-SERPINA3	BDNF	CDH1	F9	CRK	0.969
98	KLK3-SERPINA3	CDH1	BMP1	F9	CRK	0.969
99	CDH1	CCDC80	BMP1	F9	CRK	0.969
100	CDH1	MDK	AFM	F9	CRK	0.969

Table 7: Panels of 6 Biomarkers

Markers						CV AUC
1	KLK3-SERPINA3 CRK	CDH1	EGFR	MDK	F9	0.982
2	CDH1 CRK	EGFR	FRZB	CCL23	F9	0.981
3	CDH1 CRK	FRZB	CCL23	BMP1	F9	0.981
4	CDH1 CRK	EGFR	MDK	SERPINA4	F9	0.981
5	CDH1 CRK	MDK	CCL23	BMP1	F9	0.981
6	CDH1 CRK	FRZB	CCL23	BMPER	F9	0.980

(continued)

Markers						CV AUC
7	CDH1 CRK	EGFR	FRZB	MDK	F9	0.980
8	BDNF CRK	KIT	CDH1	FRZB	BMPER	0.980
9	CDH1 CRK	CCL23	BMP1	BMPER	F9	0.980
10	BDNF CRK	CDH1	VEGFA	FRZB	BMPER	0.980
11	BDNF CRK	CDH1	FRZB	ICAM2	BMPER	0.980
12	KLK3-SERPINA3 CRK	CDH1	MDK	BMPER	F9	0.980
13	KLK3-SERPINA3 CRK	CDH1	MDK	FN1	F9	0.980
14	BDNF TPT1	CDH1	FRZB	CFP	BMPER	0.980
15	CDH1 CRK	EGFR	MDK	CCL23	F9	0.979
16	CDH1 CRK	MDK	BMP1	SERPINA4	F9	0.979
17	CDH1 CRK	ICAM2	CCL23	BMPER	F9	0.979
18	BDNF CRK	CDH1	MMP9	FRZB	BMPER	0.979
19	KLK3-SERPINA3 CRK	CDH1	MDK	CCL23	F9	0.979
20	CDH1 CRK	ICAM2	CCL23	BMP1	F9	0.979
21	KLK3-SERPINA3 CRK	BDNF	CDH1	MDK	F9	0.979
22	CDH1 CRK	MDK	CCL23	BMPER	F9	0.979
23	BDNF TPT1	CDH1	FRZB	ICAM2	BMPER	0.979
24	BDNF TPT1	CDH1	VEGFA	FRZB	BMPER	0.979
25	CDH1 CRK	EGFR	FRZB	SERPINA4	F9	0.979
26	CDH1 CRK	EGFR	MDK	CCDC80	F9	0.979
27	BDNF CRK	CDH1	FRZB	CFP	BMPER	0.979

(continued)

Markers						CV AUC
28	CDH1 CRK	EGFR	MDK	BMPER	F9	0.979
29	KLK3-SERPINA3 CRK	CDH1	MDK	BMP1	F9	0.979
30	CDH1 CRK	EGFR	MDK	AFM	F9	0.978
31	CDH1 CRK	EGFR	FRZB	BMPER	F9	0.978
32	CDH1 CRK	EFNA5	EGFR	MDK	F9	0.978
33	CDH1 CRK	EGFR	MDK	TPT1	F9	0.978
34	BDNF CRK	CDH1	FRZB	MDK	BMPER	0.978
35	CDH1 CRK	MDK	CCL23	FN1	F9	0.978
36	CDH1 CRK	CCL23	BMP1	TPT1	F9	0.978
37	CDH1 CRK	CCL23	FN1	BMPER	F9	0.978
38	CDH1 CRK	EGFR	MDK	CCL23	F9	0.978
39	BDNF CRK	CDH1	FRZB	BMPER	AFM	0.978
40	CDH1 CRK	MDK	BMP1	BMPER	F9	0.978
41	BDNF CRK	CDH1	EGFR	FRZB	F9	0.978
42	CDH1 CRK	CCL23	BMP1	TNFRSF4	F9	0.978
43	BDNF CRK	CDH1	FRZB	BMPER	BPI	0.978
44	CDH1 CRK	VEGFA	FRZB	FN1	BMPER	0.978
45	KLK3-SERPINA3 CRK	CDH1	EGFR	FRZB	F9	0.978
46	CDH1 CRK	EGFR	MDK	TGFB2	F9	0.978
47	CDH1 CRK	EFNA5	MDK	BMPER	F9	0.978
48	BDNF TPT1	KIT	CDH1	FRZB	BMPER	0.978

(continued)

Markers						CV AUC
49	BDNF CRK	CDH1	FRZB	BMPER	SAA1	0.978
50	CDH1 CRK	EGFR	CCL23	FN1	F9	0.978
51	CDH1 CRK	CCL23	BMPER	TNFRSF4	F9	0.978
52	CDH1 CRK	CCL23	CCDC80	BMPER	F9	0.978
53	CDH1 CRK	MDK	SERPINA4	BMPER	F9	0.978
54	CDH1 CRK	EGFR	CCL23	TNFRSF4	F9	0.977
55	CDH1 CRK	FRZB	BMP1	BMPER	F9	0.977
56	CDH1 CRK	MDK	BMP1	AFM	F9	0.977
57	CDH1 CRK	EGFR	FRZB	AFM	F9	0.977
58	BDNF CRK	CDH1	FRZB	FN1	BMPER	0.977
59	CDH1 CRK	EGFR	CCL23	CCDC80	F9	0.977
60	KIT CRK	CDH1	CCL23	BMP1	F9	0.977
61	CDH1 CRK	EGFR	FRZB	RBP4	F9	0.977
62	CDH1 CRK	EGFR	MDK	BMP1	F9	0.977
63	BDNF CRK	CDH1	FRZB	BMPER	TNFRSF4	0.977
64	CDH1 CRK	MDK	CCL23	CCDC80	F9	0.977
65	CDH1 CRK	CCDC80	BMP1	BMPER	F9	0.977
66	BDNF CRK	CDH1	FRZB	CCL23	F9	0.977
67	CDH1 CRK	EGFR	MDK	TNFRSF4	F9	0.977
68	CDH1 CRK	EGFR	MDK	RBP4	F9	0.977
69	CDH1 CRK	EGFR	FRZB	BMP1	F9	0.977

(continued)

Markers						CV AUC
70	CDH1 CRK	EGFR	FRZB	CCDC80	F9	0.977
71	CDH1 CRK	EGFR	FRZB	CCL23	F9	0.977
72	CDH1 CRK	CCL23	CCL23	BMPER	F9	0.977
73	CDH1 CRK	ICAM2	BMP1	BMPER	F9	0.977
74	CDH1 CRK	EGFR	MDK	FN1	F9	0.977
75	BDNF CRK	CDH1	LTF	FRZB	BMPER	0.977
76	KLK3-SERPINA3 F9	CDH1	EGFR	MDK	TPT1	0.977
77	CDH1 CRK	EGFR	CCL23	BMPER	F9	0.977
78	CDH1 CRK	EGFR	CCDC80	BMPER	F9	0.977
79	CDH1 CRK	MDK	CCL23	C9	F9	0.977
80	CDH1 CRK	CCL23	RBP4	BMPER	F9	0.977
81	CDH1 CRK	EGFR	ICAM2	MDK	F9	0.977
82	CDH1 CRK	EGFR	FRZB	TNFRSF4	F9	0.977
83	CDH1 CRK	CCL23	BMP1	RBP4	F9	0.977
84	CDH1 CRK	FRZB	CCL23	FN1	F9	0.977
85	BDNF TPT1	CDH1	FRZB	MDK	BMPER	0.977
86	BDNF CRK	CDH1	FRZB	BMPER	F9	0.977
87	CDH1 CRK	CCL23	CCDC80	BMP1	F9	0.977
88	CDH1 CRK	EGFR	MDK	CFL1	F9	0.977
89	KLK3-SERPINA3 CRK	CDH1	MDK	GPC2	F9	0.977
90	BDNF TPT1	CDH1	MMP9	FRZB	BMPER	0.977

(continued)

Markers						CV AUC
91	CDH1 CRK	MDK	BMPER	AFM	F9	0.977
92	BDNF CRK	CDH1	EGFR	MDK	F9	0.977
93	CDH1 CRK	EGFR	FRZB	CFP	F9	0.977
94	KLK3-SERPINA3 CRK	CDH1	MDK	TGFB2	F9	0.977
95	BDNF CRK	CDH1	FRZB	ABL1	BMPER	0.977
96	CDH1 CRK	EGFR	SERPINA4	TGFB2	F9	0.977
97	CDH1 CRK	EGFR	CCL23	TGFB2	F9	0.976
98	CDH1 CRK	EGFR	ICAM2	CCL23	F9	0.976
99	BDNF CRK	CDH1	FRZB	SERPINA4	BMPER	0.976
100	CDH1 CRK	FRZB	CFP	FN1	BMPER	0.976

Table 8: Panels of 7 Biomarkers

Markers						CV AUC
1	CDH1 F9	EGFR CRK	FRZB	MDK	SERPINA4	0.985
2	KLK3-SERPINA3 F9	CDH1 CRK	EGFR	FRZB	MDK	0.985
3	CDH1 F9	FRZB CRK	MDK	CCL23	BMP1	0.985
4	CDH1 F9	EGFR CRK	FRZB	MDK	CCL23	0.985
5	CDH1 F9	FRZB CRK	CCL23	BMP1	BMPER	0.985
6	CDH1 F9	MDK CRK	CCL23	BMP1	BMPER	0.984
7	KLK3-SERPINA3 F9	CDH1 CRK	EGFR	MDK	BMPER	0.984
8	CDH1 F9	EGFR CRK	FRZB	CCL23	TNFRSF4	0.984
9	BDNF BMPER	CDH1 TPT1	FRZB	ICAM2	CFP	0.984

(continued)

Markers						CV AUC
10	CDH1 F9	MDK CRK	CCL23	BMP1	TPT1	0.983
11	CDH1 F9	MDK CRK	CCL23	FN1	BMPER	0.983
12	CDH1 F9	EGFR CRK	MDK	CCL23	CCDC80	0.983
13	CDH1 F9	EGFR CRK	FRZB	CCL23	BMPER	0.983
14	CDH1 F9	MDK CRK	CCL23	BMP1	TNFRSF4	0.983
15	KLK3-SERPINA3 F9	CDH1 CRK	EGFR	MDK	TNFRSF4	0.983
16	CDH1 F9	EGFR CRK	MDK	CCDC80	FN1	0.983
17	CDH1 F9	EGFR CRK	MDK	CCL23	FN1	0.983
18	CDH1 F9	EGFR CRK	MDK	SERPINA4	TPT1	0.983
19	KLK3-SERPINA3 F9	CDH1 CRK	EGFR	MDK	FN1	0.983
20	CDH1 F9	FRZB CRK	MDK	CCL23	BMPER	0.983
21	KLK3-SERPINA3 F9	CDH1 CRK	EGFR	ICAM2	MDK	0.983
22	KLK3-SERPINA3 F9	CDH1 CRK	EGFR	MDK	TGFB2	0.983
23	CDH1 F9	EGFR CRK	MDK	BMP1	SERPINA4	0.983
24	CDH1 F9	EGFR CRK	FRZB	CCDC80	BMPER	0.983
25	KLK3-SERPINA3 F9	CDH1 CRK	EGFR	MDK	FCN2	0.983
26	CDH1 F9	EGFR CRK	FRZB	CCL23	FN1	0.983
27	CDH1 F9	EGFR CRK	FRZB	CCL23	TGFB2	0.983
28	CDH1 F9	EGFR CRK	MDK	SERPINA4	TGFB2	0.983
29	KLK3-SERPINA3 F9	BDNF CRK	CDH1	EGFR	MDK	0.983
30	KLK3-SERPINA3 F9	CDH1 CRK	EGFR	MDK	TPT1	0.983

(continued)

Markers						CV AUC
31	CDH1 F9	ICAM2 CRK	CCL23	BMP1	BMPER	0.983
32	CDH1 F9	EFNA5 CRK	EGFR	MDK	SERPINA4	0.983
33	CDH1 F9	MDK CRK	CCL23	BMP1	FN1	0.983
34	CDH1 F9	EGFR CRK	ICAM2	MDK	SERPINA4	0.983
35	CDH1 F9	EGFR CRK	FRZB	MDK	CCDC80	0.983
36	CDH1 F9	FRZB CRK	ICAM2	CCL23	BMP1	0.983
37	CDH1 F9	MDK CRK	BMP1	SERPINA4	BMPER	0.982
38	KLK3-SERPINA3 F9	CDH1 CRK	EGFR	MDK	GPC2	0.982
39	BDNF BMPER	CDH1 CRK	VEGFA	FRZB	ICAM2	0.982
40	CDH1 F9	FRZB CRK	CCL23	BMP1	TPT1	0.982
41	CDH1 F9	EGFR CRK	MDK	CCDC80	BMPER	0.982
42	BDNF F9	CDH1 CRK	EGFR	FRZB	MDK	0.982
43	CDH1 F9	EGFR CRK	FRZB	CCL23	CCDC80	0.982
44	CDH1 F9	EGFR CRK	MDK	SERPINA4	BMPER	0.982
45	CDH1 F9	EGFR CRK	FRZB	CCL23	AFM	0.982
46	CDH1 F9	MDK CRK	CCDC80	BMP1	BMPER	0.982
47	CDH1 F9	EGFR CRK	FRZB	RBP4	BMPER	0.982
48	CDH1 F9	FRZB CRK	CCL23	BMP1	TNFRSF4	0.982
49	CDH1 F9	FRZB CRK	ICAM2	CCL23	BMPER	0.982
50	BDNF BMPER	CDH1 CRK	FRZB	ICAM2	CFP	0.982
51	CDH1 F9	EGFR CRK	MDK	CCL23	TPT1	0.982

(continued)

Markers						CV AUC
52	CDH1 F9	EGFR CRK	FRZB	MDK	AFM	0.982
53	KLK3-SERPINA3 F9	CDH1 CRK	MMP9	EGFR	MDK	0.982
54	CDH1 F9	ICAM2 CRK	CCL23	CCDC80	BMPER	0.982
55	KLK3-SERPINA3 F9	CDH1 CRK	EGFR	MDK	ABL1	0.982
56	CDH1 F9	EGFR CRK	FRZB	BMPER	AFM	0.982
57	CDH1 F9	EGFR CRK	MDK	CCL23	AFM	0.982
58	CDH1 F9	EGFR CRK	FRZB	MDK	BMPER	0.982
59	CDH1 F9	EGFR CRK	MDK	CCDC80	SERPINA4	0.982
60	CDH1 F9	EGFR CRK	MDK	CCL23	BMPER	0.982
61	CDH1 F9	CCL23 CRK	BMP1	BMPER	TPT1	0.982
62	KLK3-SERPINA3 F9	CDH1 CRK	EGFR	MDK	CFL1	0.982
63	KLK3-SERPINA3 F9	CDH1 CRK	EGFR	MDK	CCL23	0.982
64	KLK3-SERPINA3 F9	CDH1 CRK	MDK	CCL23	BMPER	0.982
65	CDH1 F9	EGFR CRK	FRZB	CCL23	TPT1	0.982
66	CDH1 F9	EGFR CRK	MDK	FN1	SERPINA4	0.982
67	CDH1 F9	ICAM2 CRK	MDK	CCL23	BMP1	0.982
68	BDNF BMPER	KIT CRK	CDH1	FRZB	ICAM2	0.982
69	KLK3-SERPINA3 F9	CDH1 CRK	MDK	BMP1	BMPER	0.982
70	CDH1 F9	MDK CRK	BMP1	BMPER	AFM	0.982
71	BDNF BMPER	CDH1 CRK	VEGFA	FRZB	FN1	0.982
72	KIT F9	CDH1 CRK	MDK	CCL23	BMP1	0.982

(continued)

Markers						CV AUC
73	CDH1 F9	EGFR CRK	MDK	CCL23	TNFRSF4	0.982
74	CDH1 F9	EGFR CRK	FRZB	CCL23	CCL23	0.982
75	CDH1 F9	EGFR CRK	MDK	CCDC80	AFM	0.982
76	CDH1 F9	EGFR CRK	MDK	CCL23	RBP4	0.982
77	BDNF F9	CDH1 CRK	EGFR	MDK	SERPINA4	0.982
78	CDH1 F9	EGFR CRK	FRZB	CCL23	FN1	0.982
79	CDH1 F9	EGFR CRK	MDK	FN1	AFM	0.982
80	CDH1 F9	MDK CRK	CCL23	FN1	TPT1	0.982
81	KLK3-SERPINA3 F9	CDH1 CRK	MDK	CCL23	FN1	0.982
82	CDH1 F9	EGFR CRK	MDK	SERPINA4	TNFRSF4	0.982
83	CDH1 F9	EGFR CRK	MDK	CCL23	AFM	0.982
84	CDH1 F9	EGFR CRK	FRZB	CCL23	BMP1	0.982
85	CDH1 F9	EGFR CRK	MDK	CCL23	SERPINA4	0.982
86	CDH1 F9	EGFR CRK	FRZB	CFP	CCL23	0.982
87	CDH1 F9	EGFR CRK	MDK	FN1	RBP4	0.982
88	KLK3-SERPINA3 F9	CDH1 CRK	ICAM2	MDK	BMPER	0.982
89	CDH1 F9	EGFR CRK	FRZB	CCL23	RBP4	0.982
90	KLK3-SERPINA3 F9	CDH1 CRK	ICAM2	MDK	CCL23	0.982
91	KLK3-SERPINA3 F9	CDH1 CRK	MDK	FN1	BMPER	0.982
92	CDH1 F9	EGFR CRK	MDK	CCL23	CCL23	0.982
93	CDH1 F9	MDK CRK	CCL23	CCDC80	BMPER	0.982

(continued)

Markers						CV AUC
94	CDH1 F9	ICAM2 CRK	MDK	BMP1	SERPINA4	0.982
95	CDH1 F9	EGFR CRK	MDK	CCDC80	TNFRSF4	0.982
96	CDH1 F9	EGFR CRK	MDK	CCL23	FN1	0.982
97	KIT F9	CDH1 CRK	FRZB	CCL23	BMP1	0.982
98	CDH1 F9	EGFR CRK	MDK	CCDC80	RBP4	0.982
99	CDH1 F9	FRZB CRK	MDK	CCL23	FN1	0.982
100	KLK3-SERPINA3 F9	CDH1 CRK	EGFR	MDK	CCL23	0.981

Table 9: Panels of 8 Biomarkers

Markers						CV AUC
1	CDH1 BMPER	FRZB F9	MDK CRK	CCL23	BMP1	0.988
2	CDH1 BMPER	EGFR F9	FRZB CRK	MDK	CCL23	0.988
3	CDH1 FN1	EGFR F9	FRZB CRK	MDK	CCDC80	0.988
4	CDH1 FN1	EGFR F9	FRZB CRK	MDK	CCL23	0.987
5	BDNF SERPINA4	CDH1 F9	EGFR CRK	FRZB	MDK	0.987
6	CDH1 TPT1	MDK F9	CCL23 CRK	BMP1	BMPER	0.987
7	CDH1 CCDC80	EGFR F9	FRZB CRK	MDK	CCL23	0.987
8	KLK3-SERPINA3 FCN2	BDNF F9	CDH1 CRK	EGFR	MDK	0.987
9	KLK3-SERPINA3 BMPER	CDH1 F9	EGFR CRK	FRZB	MDK	0.987
10	KLK3-SERPINA3 FN1	CDH1 F9	EGFR CRK	FRZB	MDK	0.987
11	CDH1 FN1	EGFR F9	FRZB CRK	MDK	CCL23	0.986
12	CDH1 AFM	EGFR F9	FRZB CRK	MDK	FN1	0.986

(continued)

Markers						CV AUC
13	CDH1 FCN2	EGFR F9	MDK CRK	CCL23	CCDC80	0.986
14	CDH1 BMPER	EGFR F9	FRZB CRK	MDK	CCDC80	0.986
15	CDH1 TPT1	EGFR F9	FRZB CRK	MDK	CCL23	0.986
16	CDH1 TPT1	FRZB F9	MDK CRK	CCL23	BMP1	0.986
17	CDH1 AFM	EGFR F9	FRZB CRK	MDK	CCL23	0.986
18	CDH1 TNFRSF4	EGFR F9	FRZB CRK	MDK	CCL23	0.986
19	KLK3-SERPINA3 TNFRSF4	CDH1 F9	EGFR CRK	FRZB	MDK	0.986
20	KLK3-SERPINA3 MDK	BDNF F9	CDH1 CRK	EGFR	FRZB	0.986
21	CDH1 TPT1	FRZB F9	CCL23 CRK	BMP1	BMPER	0.986
22	KLK3-SERPINA3 CCL23	CDH1 F9	EGFR CRK	FRZB	MDK	0.986
23	CDH1 TGFB2	EGFR F9	FRZB CRK	MDK	SERPINA4	0.986
24	CDH1 RBP4	EGFR F9	FRZB CRK	MDK	FN1	0.986
25	CDH1 TGFB2	EGFR F9	FRZB CRK	MDK	CCL23	0.986
26	CDH1 BMX	EGFR F9	FRZB CRK	MDK	CCL23	0.986
27	CDH1 SERPINA4	EGFR F9	FRZB CRK	MDK	BMP1	0.986
28	CDH1 BMPER	FRZB F9	ICAM2 CRK	CCL23	BMP1	0.986
29	CDH1 SERPINA4	EGFR F9	FRZB CRK	MDK	FN1	0.986
30	CDH1 TPT1	EGFR F9	FRZB CRK	MDK	SERPINA4	0.986
31	CDH1 BMPER	EGFR F9	FRZB CRK	MDK	SERPINA4	0.986
32	CDH1 BMPER	EGFR F9	MDK CRK	CCDC80	FN1	0.986
33	CDH1 AFM	EGFR F9	FRZB CRK	MDK	BMPER	0.986

(continued)

Markers						CV AUC
34	CDH1 RBP4	EGFR F9	FRZB CRK	MDK	CCL23	0.986
35	CDH1 TPT1	EGFR F9	MDK CRK	CCL23	FN1	0.986
36	CDH1 TNFRSF4	EGFR F9	MDK CRK	CCDC80	FN1	0.986
37	CDH1 BMP1	EGFR F9	FRZB CRK	MDK	CCL23	0.986
38	CDH1 TPT1	MDK F9	CCL23 CRK	BMP1	FN1	0.986
39	CDH1 SAA1	EGFR F9	FRZB CRK	MDK	CCL23	0.986
40	KLK3-SERPINA3 FCN2	CDH1 F9	EGFR CRK	MDK	CCL23	0.986
41	CDH1 TPT1	ICAM2 F9	MDK CRK	CCL23	BMP1	0.986
42	CDH1 CCL23	EGFR F9	FRZB CRK	MDK	CCL23	0.986
43	CDH1 BMPER	ICAM2 F9	MDK CRK	CCL23	BMP1	0.986
44	CDH1 SERPINA4	EGFR F9	FRZB CRK	MDK	CCDC80	0.986
45	BDNF CCDC80	CDH1 F9	EGFR CRK	FRZB	MDK	0.986
46	CDH1 TNFRSF4	FRZB F9	MDK CRK	CCL23	BMP1	0.986
47	CDH1 TPT1	EGFR F9	MDK CRK	CCL23	CCDC80	0.986
48	CDH1 BMP1	EGFR F9	FRZB CRK	MDK	CCDC80	0.986
49	KIT BMPER	CDH1 F9	MDK CRK	CCL23	BMP1	0.985
50	CDH1 BMPER	ICAM2 F9	MDK CRK	BMP1	SERPINA4	0.985
51	CDH1 TNFRSF4	ICAM2 F9	MDK CRK	CCL23	BMP1	0.985
52	CDH1 TNFRSF4	EGFR F9	FRZB CRK	MDK	SERPINA4	0.985
53	CDH1 BMPER	FRZB F9	MDK CRK	CCL23	FN1	0.985
54	BDNF RBP4	CDH1 F9	EGFR CRK	FRZB	MDK	0.985

(continued)

Markers						CV AUC
55	CDH1 BMPER	EGFR F9	FRZB CRK	MDK	RBP4	0.985
56	CDH1 TPT1	MDK F9	CCL23 CRK	FN1	BMPER	0.985
57	CDH1 TPT1	EGFR F9	MDK CRK	CCDC80	FN1	0.985
58	CDH1 RBP4	EFNA5 F9	EGFR CRK	MDK	FCN2	0.985
59	CDH1 BMPER	ICAM2 F9	MDK CRK	CCL23	FN1	0.985
60	CDH1 SERPINA4	EGFR F9	FRZB CRK	ICAM2	MDK	0.985
61	KLK3-SERPINA3 TPT1	CDH1 F9	EGFR CRK	MDK	BMPER	0.985
62	CDH1 TPT1	EGFR F9	MDK CRK	BMP1	SERPINA4	0.985
63	BDNF BMPER	CDH1 F9	FRZB CRK	MDK	CCL23	0.985
64	CDH1 RBP4	EGFR F9	MDK CRK	CCL23	FCN2	0.985
65	CDH1 BMPER	MDK F9	CCL23 CRK	BMP1	FN1	0.985
66	CDH1 FCN2	EFNA5 F9	EGFR CRK	MDK	CCL23	0.985
67	CDH1 TNFRSF4	EGFR F9	MDK CRK	CCL23	FN1	0.985
68	CDH1 CCL23	EGFR F9	FRZB CRK	MDK	CFP	0.985
69	CDH1 TGFB2	EGFR F9	FRZB CRK	CCL23	BMPER	0.985
70	CDH1 TGFB2	EGFR F9	FRZB CRK	CCL23	FN1	0.985
71	KLK3-SERPINA3 TGFB2	CDH1 F9	EGFR CRK	MDK	BMPER	0.985
72	CDH1 BMPER	ICAM2 F9	MDK CRK	CCDC80	BMP1	0.985
73	BDNF CCL23	CDH1 F9	EGFR CRK	FRZB	MDK	0.985
74	BDNF TPT1	CDH1 F9	EGFR CRK	MDK	SERPINA4	0.985
75	CDH1 BMP1	FRZB F9	ICAM2 CRK	MDK	CCL23	0.985

(continued)

Markers						CV AUC
76	CDH1 TPT1	FRZB F9	MDK CRK	CCL23	BMPER	0.985
77	KLK3-SERPINA3 BMPER	CDH1 F9	FRZB CRK	MDK	CCL23	0.985
78	CDH1 TGFB2	EGFR F9	FRZB CRK	CCL23	FN1	0.985
79	CDH1 FN1	FRZB F9	MDK CRK	CCL23	BMP1	0.985
80	CDH1 RBP4	EGFR F9	FRZB CRK	MDK	BMP1	0.985
81	CDH1 TPT1	EGFR F9	MDK CRK	CCL23	BMPER	0.985
82	CDH1 TNFRSF4	MDK F9	CCL23 CRK	BMP1	BMPER	0.985
83	CDH1 CFL1	EGFR F9	FRZB CRK	MDK	CCL23	0.985
84	CDH1 BMPER	FRZB F9	MDK CRK	BMP1	SERPINA4	0.985
85	CDH1 TGFB2	EGFR F9	MDK CRK	CCL23	FN1	0.985
86	CDH1 AFM	EGFR F9	FRZB CRK	MDK	CCDC80	0.985
87	CDH1 FCN2	EGFR F9	MDK CRK	C9	CCDC80	0.985
88	CDH1 FN1	EGFR F9	ICAM2 CRK	MDK	CCDC80	0.985
89	CDH1 AFM	EGFR F9	MDK CRK	CCL23	FCN2	0.985
90	CDH1 AFM	EFNA5 F9	EGFR CRK	MDK	FCN2	0.985
91	CDH1 SAA1	FRZB F9	MDK CRK	CCL23	BMP1	0.985
92	CDH1 AFM	EGFR F9	FRZB CRK	MDK	TNFRSF4	0.985
93	CDH1 SERPINA4	MMP9 F9	EGFR CRK	FRZB	MDK	0.985
94	KLK3-SERPINA3 TPT1	CDH1 F9	EGFR CRK	FRZB	MDK	0.985
95	CDH1 CCL23	EGFR F9	FRZB CRK	ICAM2	MDK	0.985
96	KLK3-SERPINA3 FCN2	CDH1 F9	EFNA5 CRK	EGFR	MDK	0.985

(continued)

Markers						CV AUC
97	CDH1 TPT1	EGFR F9	ICAM2 CRK	MDK	SERPINA4	0.985
98	CDH1 TPT1	EGFR F9	FRZB CRK	CCL23	FN1	0.985
99	CDH1 AFM	EGFR F9	MDK CRK	FN1	TNFRSF4	0.985
100	CDH1 BMPER	EGFR F9	MDK CRK	CCL23	CCDC80	0.985

Table 10: Panels of 9 Biomarkers

Markers						CV AUC
1	CDH1 BMPER	FRZB TPT1	MDK F9	CCL23 CRK	BMP1	0.990
2	CDH1 FN1	EGFR TPT1	FRZB F9	MDK CRK	CCL23	0.990
3	CDH1 BMPER	ICAM2 TPT1	MDK F9	CCL23 CRK	BMP1	0.989
4	CDH1 FN1	EGFR TGFB2	FRZB F9	MDK CRK	CCL23	0.989
5	CDH1 BMPER	EGFR TGFB2	FRZB F9	MDK CRK	CCL23	0.989
6	BDNF SERPINA4	CDH1 TPT1	EGFR F9	FRZB CRK	MDK	0.989
7	CDH1 BMP1	FRZB BMPER	ICAM2 F9	MDK CRK	CCL23	0.989
8	CDH1 BMPER	EGFR TPT1	FRZB F9	MDK CRK	CCL23	0.989
9	BDNF BMPER	CDH1 TPT1	FRZB F9	MDK CRK	CCL23	0.989
10	CDH1 FN1	EGFR TPT1	FRZB F9	MDK CRK	CCDC80	0.989
11	CDH1 BMP1	EGFR TPT1	FRZB F9	MDK CRK	CCL23	0.989
12	CDH1 FN1	EGFR BMPER	FRZB F9	MDK CRK	CCDC80	0.989
13	CDH1 CCL23	EFNA5 BMX	EGFR F9	FRZB CRK	MDK	0.989
14	CDH1 FCN2	EGFR TPT1	MDK F9	CCL23 CRK	CCDC80	0.989
15	CDH1 BMX	FRZB BMPER	MDK F9	CCL23 CRK	BMP1	0.989

(continued)

Markers						CV AUC
16	KLK3-SERPINA3 MDK	BDNF FCN2	CDH1 F9	EGFR CRK	FRZB	0.989
17	CDH1 BMPER	FRZB TNFRSF4	MDK F9	CCL23 CRK	BMP1	0.989
18	CDH1 BMX	EGFR BMPER	FRZB F9	MDK CRK	CCL23	0.989
19	CDH1 TPT1	EGFR AFM	FRZB F9	MDK CRK	FN1	0.989
20	CDH1 FN1	EGFR TNFRSF4	FRZB F9	MDK CRK	CCL23	0.988
21	CDH1 FCN2	EFNA5 TPT1	EGFR F9	MDK CRK	CCL23	0.988
22	CDH1 CCDC80	EGFR TGFB2	FRZB F9	MDK CRK	CCL23	0.988
23	CDH1 CCDC80	EGFR BMPER	FRZB F9	MDK CRK	CCL23	0.988
24	CDH1 FCN2	EFNA5 BMX	EGFR F9	MDK CRK	CCL23	0.988
25	CDH1 CCDC80	EGFR TPT1	FRZB F9	MDK CRK	CCL23	0.988
26	CDH1 SERPINA4	EGFR TPT1	FRZB F9	MDK CRK	FN1	0.988
27	CDH1 BMPER	EGFR TGFB2	FRZB F9	MDK CRK	SERPINA4	0.988
28	CDH1 TGFB2	EGFR AFM	FRZB F9	MDK CRK	CCL23	0.988
29	BDNF CCDC80	CDH1 FN1	EGFR F9	FRZB CRK	MDK	0.988
30	CDH1 BMP1	AZU1 BMPER	FRZB F9	MDK CRK	CCL23	0.988
31	CDH1 PRTN3	FRZB BMPER	MDK F9	CCL23 CRK	BMP1	0.988
32	CDH1 TPT1	EGFR AFM	FRZB F9	MDK CRK	CCL23	0.988
33	CDH1 BMPER	FRZB TPT1	MDK F9	CCL23 CRK	FN1	0.988
34	CDH1 RBP4	EGFR TGFB2	FRZB F9	MDK CRK	FN1	0.988
35	KLK3-SERPINA3 MDK	BDNF TPT1	CDH1 F9	EGFR CRK	FRZB	0.988
36	CDH1 FN1	EGFR TPT1	FRZB F9	MDK CRK	CCL23	0.988

(continued)

	Markers					CV AUC
37	KIT BMP1	CDH1 BMPER	FRZB F9	MDK CRK	CCL23	0.988
38	CDH1 BMPER	FRZB BPI	MDK F9	CCL23 CRK	BMP1	0.988
39	CDH1 SERPINA4	EGFR TGFB2	FRZB F9	MDK CRK	FN1	0.988
40	CDH1 FN1	EGFR TNFRSF4	FRZB F9	MDK CRK	CCDC80	0.988
41	BDNF BMP1	CDH1 TPT1	FRZB F9	MDK CRK	CCL23	0.988
42	CDH1 BMP1	EGFR TGFB2	FRZB F9	MDK CRK	CCL23	0.988
43	CDH1 TNFRSF4	EGFR TPT1	FRZB F9	MDK CRK	CCL23	0.988
44	CDH1 BMP1	EGFR BMX	FRZB F9	MDK CRK	CCL23	0.988
45	CDH1 TGFB2	EGFR AFM	FRZB F9	MDK CRK	FN1	0.988
46	BDNF CCL23	CDH1 TPT1	EGFR F9	FRZB CRK	MDK	0.988
47	CDH1 CCL23	EGFR TPT1	FRZB F9	MDK CRK	CCL23	0.988
48	CDH1 RBP4	EGFR TPT1	FRZB F9	MDK CRK	FN1	0.988
49	CDH1 TGFB2	EGFR SAA1	FRZB F9	MDK CRK	CCL23	0.988
50	CDH1 FN1	EGFR BMPER	FRZB F9	MDK CRK	CCL23	0.988
51	CDH1 FN1	FRZB TPT1	MDK F9	CCL23 CRK	BMP1	0.988
52	CDH1 BMP1	LTF BMPER	FRZB F9	MDK CRK	CCL23	0.988
53	CDH1 FN1	FRZB BMPER	MDK F9	CCL23 CRK	BMP1	0.988
54	CDH1 CCDC80	EGFR BMX	FRZB F9	MDK CRK	CCL23	0.988
55	CDH1 BMP1	ALPL BMPER	FRZB F9	MDK CRK	CCL23	0.988
56	CDH1 BMP1	VEGFA BMPER	FRZB F9	MDK CRK	CCL23	0.988
57	CDH1 SERPINA4	EGFR TPT1	FRZB F9	MDK CRK	BMP1	0.988

(continued)

Markers						CV AUC
58	CDH1 CCL23	EGFR FN1	FRZB F9	ICAM2 CRK	MDK	0.988
59	CDH1 CCDC80	EGFR FCN2	FRZB F9	MDK CRK	CCL23	0.988
60	CDH1 BMP1	MMP9 BMPER	FRZB F9	MDK CRK	CCL23	0.988
61	CDH1 TPT1	EGFR SAA1	FRZB F9	MDK CRK	CCL23	0.988
62	BDNF SERPINA4	CDH1 TGFB2	EGFR F9	FRZB CRK	MDK	0.988
63	CDH1 BMPER	MDK TPT1	CCL23 F9	BMP1 CRK	FN1	0.988
64	CDH1 BMP1	EGFR BMPER	FRZB F9	MDK CRK	CCDC80	0.988
65	CDH1 FN1	EGFR TGFB2	FRZB F9	MDK CRK	CCDC80	0.988
66	CDH1 FN1	EGFR CFL1	FRZB F9	MDK CRK	CCL23	0.988
67	CDH1 CCL23	EGFR BMPER	FRZB F9	ICAM2 CRK	MDK	0.988
68	BDNF FN1	CDH1 RBP4	EGFR F9	FRZB CRK	MDK	0.988
69	CDH1 CCL23	EGFR CCDC80	FRZB F9	ICAM2 CRK	MDK	0.988
70	CDH1 TPT1	FRZB SAA1	MDK F9	CCL23 CRK	BMP1	0.988
71	CDH1 SERPINA4	EGFR TGFB2	FRZB F9	MDK CRK	BMP1	0.988
72	CDH1 RBP4	EGFR TGFB2	FRZB F9	MDK CRK	CCL23	0.988
73	BDNF FN1	CDH1 AFM	EGFR F9	FRZB CRK	MDK	0.988
74	BDNF CCL23	CDH1 FN1	EGFR F9	FRZB CRK	MDK	0.988
75	BDNF MDK	CDH1 FCN2	EFNA5 F9	EGFR CRK	FRZB	0.988
76	CDH1 BMPER	EGFR TPT1	FRZB F9	MDK CRK	SERPINA4	0.988
77	CDH1 TNFRSF4	MDK TPT1	CCL23 F9	BMP1 CRK	BMPER	0.988
78	CDH1 BMPER	FRZB SAA1	MDK F9	CCL23 CRK	BMP1	0.988

(continued)

Markers						CV AUC
79	CDH1 CCDC80	EGFR TPT1	ICAM2 F9	MDK CRK	CCL23	0.988
80	KLK3-SERPINA3 MDK	BDNF TGFB2	CDH1 F9	EGFR CRK	FRZB	0.988
81	CDH1 BMP1	FRZB BMPER	MDK F9	CCL23 CRK	CCDC80	0.988
82	CDH1 BMP1	LCN2 BMPER	FRZB F9	MDK CRK	CCL23	0.988
83	CDH1 TNFRSF4	EGFR AFM	FRZB F9	MDK CRK	FN1	0.988
84	CDH1 BMPER	EGFR TPT1	FRZB F9	MDK CRK	CCDC80	0.988
85	CDH1 RBP4	EGFR BMPER	FRZB F9	MDK CRK	CCL23	0.988
86	CDH1 FN1	EGFR CFL1	FRZB F9	MDK CRK	CCDC80	0.988
87	CDH1 BMX	EGFR AFM	FRZB F9	MDK CRK	CCL23	0.988
88	CDH1 FN1	ICAM2 TPT1	MDK F9	CCL23 CRK	BMP1	0.988
89	CDH1 FCN2	EGFR BMP1	MDK F9	CCL23 CRK	CCDC80	0.988
90	CDH1 CCDC80	EGFR FCN2	ICAM2 F9	MDK CRK	CCL23	0.988
91	CDH1 CCL23	EGFR FN1	FRZB F9	MDK CRK	CFP	0.988
92	BDNF CCL23	CDH1 FN1	EGFR F9	FRZB CRK	MDK	0.988
93	BDNF CCL23	CDH1 TGFB2	EGFR F9	FRZB CRK	MDK	0.988
94	CDH1 BMP1	MPO BMPER	FRZB F9	MDK CRK	CCL23	0.988
95	CDH1 RBP4	EGFR TPT1	FRZB F9	MDK CRK	CCL23	0.988
96	BDNF TPT1	CDH1 AFM	EGFR F9	FRZB CRK	MDK	0.988
97	CDH1 ABL1	EGFR FN1	FRZB F9	MDK CRK	CCL23	0.988
98	CDH1 BMP1	FRZB TPT1	ICAM2 F9	MDK CRK	CCL23	0.988
99	KLK3-SERPINA3 FN1	CDH1 TGFB2	EGFR F9	FRZB CRK	MDK	0.988

(continued)

Markers						CV AUC
100	BDNF RBP4	CDH1 BMPER	EGFR F9	FRZB CRK	MDK	0.988

Table 11: Panels of 10 Biomarkers

Markers						CV AUC
1	CDH1 BMP1	FRZB BMPER	ICAM2 TPT1	MDK F9	CCL23 CRK	0.992
2	CDH1 FN1	EGFR BMPER	FRZB TPT1	MDK F9	CCDC80 CRK	0.991
3	CDH1 FN1	EGFR BMPER	FRZB TPT1	MDK F9	CCL23 CRK	0.991
4	CDH1 FN1	EGFR TPT1	FRZB TGFB2	MDK F9	CCL23 CRK	0.991
5	CDH1 FN1	EGFR BMPER	FRZB TGFB2	MDK F9	CCL23 CRK	0.991
6	CDH1 BMP1	VEGFA BMPER	FRZB TPT1	MDK F9	CCL23 CRK	0.991
7	CDH1 CCL23	EGFR BMPER	FRZB TPT1	ICAM2 F9	MDK CRK	0.991
8	BDNF BMP1	CDH1 BMPER	FRZB TPT1	MDK F9	CCL23 CRK	0.991
9	CDH1 CCL23	EGFR FN1	FRZB TPT1	ICAM2 F9	MDK CRK	0.991
10	CDH1 BMPER	ICAM2 TNFRSF4	MDK TPT1	CCL23 F9	BMP1 CRK	0.991
11	CDH1 BMP1	EGFR BMX	FRZB TPT1	MDK F9	CCL23 CRK	0.990
12	CDH1 BMPER	FRZB TNFRSF4	MDK TPT1	CCL23 F9	BMP1 CRK	0.990
13	CDH1 FN1	EGFR TNFRSF4	FRZB TPT1	MDK F9	CCL23 CRK	0.990
14	BDNF CCL23	CDH1 FN1	EGFR TPT1	FRZB F9	MDK CRK	0.990
15	BDNF CCL23	CDH1 BMPER	FRZB TPT1	ICAM2 F9	MDK CRK	0.990
16	CDH1 BMP1	MMP9 BMPER	FRZB TPT1	MDK F9	CCL23 CRK	0.990
17	CDH1 CCDC80	EGFR BMPER	FRZB TPT1	MDK F9	CCL23 CRK	0.990
18	CDH1 BMX	EGFR BMPER	FRZB TPT1	MDK F9	CCL23 CRK	0.990

(continued)

Markers						CV AUC
19	CDH1 BMP1	EGFR BMX	FRZB BMPER	MDK F9	CCL23 CRK	0.990
20	BDNF SERPINA4	CDH1 BMPER	EGFR TPT1	FRZB F9	MDK CRK	0.990
21	CDH1 BMP1	ICAM2 BMPER	MDK TPT1	CCL23 F9	CCDC80 CRK	0.990
22	CDH1 FN1	EGFR TNFRSF4	FRZB TGFB2	MDK F9	CCL23 CRK	0.990
23	CDH1 BMP1	EGFR BMPER	FRZB TPT1	MDK F9	CCL23 CRK	0.990
24	CDH1 FN1	FRZB BMPER	MDK TPT1	CCL23 F9	BMP1 CRK	0.990
25	BDNF CCDC80	CDH1 FN1	EGFR TPT1	FRZB F9	MDK CRK	0.990
26	CDH1 CCL23	EGFR BMP1	FRZB TPT1	ICAM2 F9	MDK CRK	0.990
27	KLK3-SERPINA3 MDK	BDNF FCN2	CDH1 TPT1	EGFR F9	FRZB CRK	0.990
28	CDH1 BMPER	EGFR TPT1	FRZB TGFB2	MDK F9	CCL23 CRK	0.990
29	BDNF FN1	CDH1 TPT1	EGFR AFM	FRZB F9	MDK CRK	0.990
30	BDNF CCL23	CDH1 BMPER	EGFR TPT1	FRZB F9	MDK CRK	0.990
31	CDH1 FN1	EGFR BMPER	FRZB TGFB2	MDK F9	CCDC80 CRK	0.990
32	BDNF CCL23	CDH1 CCDC80	EGFR FCN2	FRZB F9	MDK CRK	0.990
33	BDNF CCL23	CDH1 TNFRSF4	EGFR TPT1	FRZB F9	MDK CRK	0.990
34	CDH1 BMP1	EGFR TPT1	FRZB TGFB2	MDK F9	CCL23 CRK	0.990
35	BDNF BMPER	CDH1 TPT1	EGFR AFM	FRZB F9	MDK CRK	0.990
36	BDNF FN1	CDH1 RBP4	EGFR TGFB2	FRZB F9	MDK CRK	0.990
37	BDNF MDK	CDH1 SERPINA4	EGFR TPT1	FRZB F9	ICAM2 CRK	0.990
38	CDH1 CCDC80	EGFR FCN2	ICAM2 TPT1	MDK F9	CCL23 CRK	0.990
39	BDNF FN1	CDH1 RBP4	EGFR TPT1	FRZB F9	MDK CRK	0.990

(continued)

Markers						CV AUC
40	BDNF SERPINA4	CDH1 TPT1	EGFR TGFB2	FRZB F9	MDK CRK	0.990
41	BDNF FN1	CDH1 BMPER	FRZB TPT1	MDK F9	CCL23 CRK	0.990
42	CDH1 BMPER	FRZB TPT1	MDK SAA1	CCL23 F9	BMP1 CRK	0.990
43	BDNF CCL23	CDH1 FN1	EGFR TGFB2	FRZB F9	MDK CRK	0.990
44	CDH1 BMP1	EGFR FN1	FRZB TPT1	MDK F9	CCL23 CRK	0.990
45	BDNF FN1	CDH1 SERPINA4	EGFR TPT1	FRZB F9	MDK CRK	0.990
46	CDH1 FCN2	EGFR BMP1	MDK TPT1	CCL23 F9	CCDC80 CRK	0.990
47	CDH1 CCDC80	EGFR BMPER	FRZB TGFB2	MDK F9	CCL23 CRK	0.990
48	CDH1 BMP1	EGFR TNFRSF4	FRZB TPT1	MDK F9	CCL23 CRK	0.990
49	KLK3-SERPINA3 MDK	BDNF BMPER	CDH1 TPT1	EGFR F9	FRZB CRK	0.990
50	CDH1 BMPER	EGFR TPT1	FRZB AFM	MDK F9	FN1 CRK	0.990
51	CDH1 CCL23	EGFR FN1	FRZB TPT1	MDK F9	CFP CRK	0.990
52	CDH1 CCL23	MMP9 FN1	EGFR TPT1	FRZB F9	MDK CRK	0.990
53	CDH1 BMPER	EGFR TGFB2	FRZB AFM	MDK F9	FN1 CRK	0.990
54	CDH1 CCDC80	EGFR FCN2	FRZB TPT1	MDK F9	CCL23 CRK	0.990
55	BDNF CCDC80	CDH1 BMPER	EGFR TPT1	FRZB F9	MDK CRK	0.990
56	CDH1 CCDC80	EGFR FN1	FRZB TPT1	MDK F9	CCL23 CRK	0.990
57	CDH1 CCDC80	EGFR BMX	FRZB BMPER	MDK F9	CCL23 CRK	0.990
58	CDH1 BMP1	FRZB FN1	ICAM2 TPT1	MDK F9	CCL23 CRK	0.990
59	CDH1 BMPER	EGFR TNFRSF4	FRZB TPT1	MDK F9	CCL23 CRK	0.990
60	BDNF RBP4	CDH1 BMPER	EGFR TPT1	FRZB F9	MDK CRK	0.990

(continued)

Markers						CV AUC
61	BDNF FN1	CDH1 TGFB2	EGFR AFM	FRZB F9	MDK CRK	0.990
62	CDH1 FCN2	EGFR BMP1	FRZB TPT1	MDK F9	CCL23 CRK	0.990
63	BDNF CCDC80	CDH1 FN1	EGFR TGFB2	FRZB F9	MDK CRK	0.990
64	BDNF CCL23	CDH1 FN1	EGFR TPT1	FRZB F9	MDK CRK	0.990
65	CDH1 BMP1	EGFR BMPER	FRZB TPT1	MDK F9	CCDC80 CRK	0.990
66	CDH1 FN1	FRZB BMPER	ICAM2 TPT1	MDK F9	CCL23 CRK	0.990
67	KIT CCL23	CDH1 FN1	EGFR TGFB2	FRZB F9	MDK CRK	0.990
68	BDNF MDK	CDH1 FCN2	EFNA5 RBP4	EGFR F9	FRZB CRK	0.990
69	BDNF CCDC80	CDH1 FCN2	EGFR RBP4	FRZB F9	MDK CRK	0.990
70	KLK3-SERPINA3 MDK	BDNF FCN2	CDH1 TGFB2	EGFR F9	FRZB CRK	0.990
71	CDH1 CCL23	AZU1 BMP1	FRZB BMPER	ICAM2 F9	MDK CRK	0.990
72	CDH1 CCL23	EGFR FN1	FRZB TGFB2	ICAM2 F9	MDK CRK	0.990
73	CDH1 CCDC80	EGFR FCN2	FRZB BMP1	MDK F9	CCL23 CRK	0.990
74	BDNF MDK	CDH1 FCN2	EFNA5 TPT1	EGFR F9	FRZB CRK	0.990
75	CDH1 CCL23	EFNA5 BMX	EGFR BMPER	FRZB F9	MDK CRK	0.990
76	BDNF CCDC80	CDH1 FCN2	EGFR BMP1	FRZB F9	MDK CRK	0.990
77	BDNF BMP1	CDH1 BMX	FRZB BMPER	MDK F9	CCL23 CRK	0.990
78	CDH1 FN1	ICAM2 BMPER	MDK TPT1	CCL23 F9	BMP1 CRK	0.990
79	CDH1 CCDC80	EGFR FN1	FRZB BMPER	ICAM2 F9	MDK CRK	0.990
80	BDNF CCDC80	CDH1 FCN2	EGFR TPT1	MDK F9	CCL23 CRK	0.990
81	CDH1 FN1	EGFR TNFRSF4	FRZB TPT1	MDK F9	CCDC80 CRK	0.990

(continued)

Markers						CV AUC
82	CDH1 CCL23	EGFR BMPER	FRZB TPT1	MDK F9	CCL23 CRK	0.990
83	KLK3-SERPINA3 ICAM2	BDNF MDK	CDH1 FCN2	EGFR F9	FRZB CRK	0.990
84	CDH1 BMP1	MMP9 BMPER	ICAM2 TPT1	MDK F9	CCL23 CRK	0.990
85	CDH1 BMP1	EFNA5 BMX	FRZB BMPER	MDK F9	CCL23 CRK	0.990
86	BDNF CCDC80	CDH1 FN1	EGFR TNFRSF4	FRZB F9	MDK CRK	0.990
87	CDH1 RBP4	EGFR BMPER	FRZB TPT1	MDK F9	FN1 CRK	0.990
88	CDH1 FN1	EGFR TPT1	FRZB AFM	MDK F9	CCL23 CRK	0.990
89	BDNF BMX	CDH1 BMPER	FRZB TPT1	MDK F9	CCL23 CRK	0.990
90	KLK3-SERPINA3 MDK	BDNF CCL23	CDH1 FCN2	EGFR F9	FRZB CRK	0.990
91	CDH1 BMP1	VEGFA FN1	FRZB TPT1	MDK F9	CCL23 CRK	0.990
92	CDH1 CCL23	EGFR CCDC80	FRZB TPT1	ICAM2 F9	MDK CRK	0.990
93	CDH1 CCL23	EFNA5 BMX	EGFR TPT1	FRZB F9	MDK CRK	0.990
94	BDNF CCL23	CDH1 CCDC80	EGFR FCN2	FRZB F9	MDK CRK	0.990
95	CDH1 BMX	FRZB BMPER	MDK TPT1	CCL23 F9	BMP1 CRK	0.989
96	CDH1 FN1	EGFR BMPER	FRZB TPT1	MDK F9	CCL23 CRK	0.989
97	CDH1 BMP1	VEGFA BMPER	ICAM2 TPT1	MDK F9	CCL23 CRK	0.989
98	KIT BMP1	CDH1 FN1	FRZB BMPER	MDK F9	CCL23 CRK	0.989
99	CDH1 FN1	EGFR PRTN3	FRZB TPT1	MDK F9	CCL23 CRK	0.989
100	BDNF BMPER	CDH1 TNFRSF4	FRZB TPT1	MDK F9	CCL23 CRK	0.989

Table 12: Counts of markers in biomarker panels

	Panel Size							
Biomarker	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
ABL1	47	48	32	27	25	24	15	16
AFM	52	65	55	69	83	94	111	111
ALB	3	1	0	0	0	0	0	0
ALPL	5	11	10	10	8	7	10	11
APOA1	9	0	0	0	0	0	0	0
AZU1	5	23	14	15	16	18	27	30
BDNF	43	73	134	131	139	103	181	321
BMP1	35	44	86	158	223	261	255	290
BMPER	228	341	255	351	321	324	336	396
BMX	25	16	12	11	10	22	51	107
BPI	7	19	18	16	19	15	23	26
C9	48	37	32	27	22	22	19	12
CAMK1	3	31	13	5	1	1	0	0
CCDC80	54	63	56	78	88	136	188	202
CCL23	44	43	55	60	65	77	70	58
CCL23	125	116	137	301	407	453	502	605
CDH1	653	971	996	1000	1000	1000	1000	1000
CDK5-CDK5R1	15	9	27	15	5	0	0	0
CDK8-CCNC	7	6	17	9	4	5	2	0
CFHR5	11	13	14	10	8	4	4	6
CFL1	55	66	53	29	33	31	38	32
CFP	29	37	58	43	29	24	24	33
CRK	149	205	643	888	963	994	999	1000
CRP	35	30	19	14	9	11	6	4
CSN1S1	30	15	9	9	6	5	4	2
CXCL13	35	10	7	1	0	0	0	0
DDC	4	4	2	0	0	0	0	0
EFNA5	39	52	74	69	55	55	67	67
EGFR	30	64	137	337	468	662	778	789
EIF4EBP2	15	12	10	9	7	8	6	4
ESM1	55	14	29	13	0	0	0	0
F9	82	486	768	862	936	987	1000	1000
FCN2	86	40	14	18	39	79	152	191
FGA-FGB-FGG	1	1	0	0	0	0	0	0
FLT3LG	3	4	1	0	0	0	0	0
FN1	48	52	55	76	136	229	290	340
FN1	11	10	25	39	30	17	8	2

(continued)

	Panel Size							
Biomarker	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
FRZB	91	201	224	261	357	450	669	830
GPC2	18	3	3	5	13	12	7	3
GPI	25	16	9	10	14	16	19	12
H2AFZ	8	0	0	0	0	0	0	0
HINT1	6	6	1	0	0	0	0	0
ICAM2	26	49	52	76	128	143	178	222
ITGA1-ITGB1	3	1	0	0	0	0	0	0
ITIH4	13	3	0	0	0	0	0	0
KIT	35	46	35	27	33	23	30	47
KLK3-SERPINA3	96	72	125	131	169	171	128	76
LCN2	2	7	10	8	7	7	8	7
LTF	5	14	11	13	10	9	16	18
MDK	49	47	94	244	559	838	975	998
MMP9	11	20	21	29	34	36	36	58
MPO	3	10	11	7	11	6	7	9
MSLN	2	2	0	0	0	0	0	0
PLA2G5	46	30	23	15	11	9	3	0
PRTN3	5	16	12	10	13	11	22	27
RBP4	27	43	37	59	62	76	79	86
SAA1	43	42	43	36	32	29	27	26
SERPINA4	43	53	52	64	95	117	116	80
TGFB2	60	50	53	61	62	82	107	182
TIMP1	10	6	9	4	0	0	0	0
TNFRSF4	27	24	31	56	78	119	123	142
TNFRSF8	2	4	6	1	0	0	0	0
TPT1	82	104	197	132	109	153	259	479
VEGFA	74	57	62	50	47	25	25	43
YWHAH	62	42	12	1	1	0	0	0

Table 13: Analytes in ten marker classifiers

CDH1	CRK
F9	MDK
FRZB	EGFR
CCL23	TPT1
BMPER	FN1

Table 14: Parameters derived from training set for naïve Bayes classifier.

Biomarker	μ_c	μ_d	σ_c	σ_d
CSN1S1	8.744	8.621	0.087	0.132
BMPER	7.309	7.061	0.206	0.247
CFHR5	8.943	9.232	0.239	0.344
CCL23	8.276	8.608	0.235	0.461
CDH1	9.132	8.827	0.161	0.267
CCDC80	8.588	8.846	0.218	0.365
TGFB2	6.882	6.833	0.044	0.049
FCN2	7.792	8.187	0.175	0.283
SERPINA4	10.713	10.398	0.130	0.433
MPO	9.440	9.975	0.808	0.968
CRP	7.836	9.788	1.059	1.962
FRZB	8.136	8.466	0.315	0.297
BDNF	6.828	6.709	0.103	0.094
FGA-FGB-FGG	9.639	10.247	0.514	0.620
H2AFZ	6.664	6.894	0.119	0.311
AFM	10.236	9.850	0.199	0.465
CRK	7.196	7.686	0.252	0.414
CFL1	7.949	8.169	0.143	0.288
BMX	7.153	7.066	0.083	0.061
RBP4	8.856	8.603	0.171	0.333
C9	11.525	11.955	0.199	0.291
MDK	7.033	7.244	0.179	0.470
ESM1	7.562	7.751	0.139	0.332
TNFRSF8	7.184	7.219	0.053	0.081
CFP	9.650	9.449	0.160	0.214
FLT3LG	6.636	6.797	0.125	0.133
ITIH4	10.180	10.461	0.318	0.337
MMP9	10.371	10.311	0.598	0.632
LTF	11.096	11.579	0.686	0.881
KIT	9.389	9.181	0.156	0.196
CDK5-CDK5R1	6.745	6.870	0.105	0.153
VEGFA	7.521	7.711	0.098	0.249
CDK8-CCNC	6.724	6.854	0.097	0.107
MSLN	8.101	8.001	0.062	0.102
ABL1	8.093	7.967	0.096	0.119
LCN2	9.887	10.049	0.384	0.546
GPC2	6.357	6.292	0.045	0.042

(continued)

Biomarker	μ_c	μ_d	σ_c	σ_d
TIMP1	8.763	8.927	0.118	0.310
FN1	10.929	10.599	0.333	0.422
ICAM2	7.392	7.369	0.041	0.063
ALB	9.491	9.281	0.124	0.299
CAMK1	8.318	8.527	0.157	0.208
PRTN3	8.779	9.243	0.860	0.894
YWHAH	7.820	8.309	0.310	0.645
HINT1	6.639	6.583	0.051	0.045
EGFR	10.463	10.264	0.111	0.209
EFNA5	6.697	6.833	0.113	0.253
IL31	6.478	6.407	0.045	0.046
BPI	10.379	11.043	1.014	1.250
BMP1	8.616	8.303	0.271	0.350
CCL23	7.224	7.528	0.152	0.259
GPI	7.834	8.422	0.454	0.701
EIF4EBP2	6.532	6.470	0.048	0.053
PLA2G5	7.021	6.926	0.063	0.068
ITGA1-ITGB1	7.345	7.977	0.377	0.671
TPT1	9.224	10.393	0.805	1.202
DDC	6.553	6.499	0.043	0.049
TNFRSF4	7.171	7.094	0.064	0.078
ALPL	7.799	8.245	0.662	0.937
SAA1	6.891	8.598	1.076	2.033
APOA1	8.557	8.281	0.164	0.258
CXCL13	6.890	7.020	0.084	0.145
KLK3-SERPINA3	7.997	8.511	0.161	0.530
FN1	8.923	8.533	0.362	0.378
AZU1	7.053	7.556	0.720	0.764
F9	8.870	9.498	0.627	0.345

Table 15: AUC for exemplary combinations of biomarkers

#											AUC
1	CDH1										0.884
2	CDH1	BMPER									0.947
3	CDH1	BMPER	F9								0.951
4	CDH1	BMPER	F9	CCL23							0.954
5	CDH1	BMPER	F9	CCL23	CRK						0.980

(continued)

#											AUC
6	CDH1	BMPER	F9	CCL23	CRK	BMP1					0.983
7	CDH1	BMPER	F9	CCL23	CRK	BMP1	TPT1				0.983
8	CDH1	BMPER	F9	CCL23	CRK	BMP1	TPT1	FRZB			0.987
9	CDH1	BMPER	F9	CCL23	CRK	BMP1	TPT1	FRZB	MDK		0.992
10	CDH1	BMPER	F9	CCL23	CRK	BMP1	TPT1	FRZB	MDK	ICAM2	0.993

Table 16: Calculations derived from training set for naïve Bayes classifier.

Biomarker	μ_c	μ_d	σ_c	σ_d	$\tilde{x}$	$p(c\|\tilde{x})$	$p(d\|\tilde{x})$	$ln(c\|\tilde{x})/p(c\|\tilde{x})$
BMPER	7.309	7.061	0.206	0.247	7.290	1.933	1.049	-0.611
CRK	7.196	7.686	0.252	0.414	7.323	1.396	0.656	-0.756
BMP1	8.616	8.303	0.271	0.350	8.878	0.921	0.295	-1.138
CCL23	7.224	7.528	0.152	0.259	7.283	2.434	0.986	-0.904
CDH1	9.132	8.827	0.161	0.267	9.594	0.040	0.024	-0.524
TPT1	9.224	10.393	0.805	1.202	8.304	0.258	0.073	-1.257
MDK	7.033	7.244	0.179	0.470	7.047	2.220	0.777	-1.050
ICAM2	7.392	7.369	0.041	0.063	7.447	3.931	2.928	-0.294
FRZB	8.136	8.466	0.315	0.297	8.025	1.190	0.448	-0.978
F9	8.870	9.498	0.627	0.345	10.009	0.122	0.385	1.147

Table 17: Clinical characteristics of the training set

Meta Data	Levels	Control	Meso	p-value
Samples		140	158	
GENDER	F M	41 99	28 130	2.61e-02
AGE	Mean SD	61.4 10.6	64.6 9.8	6.80e-03

Table 18: Ten biomarker classifier proteins

Biomarker	UniProt ID	Direction*	Biological Process (GO)
CDH1	P12830	Down	regulation of cell death
FRZB	Q92765	Up	regulation of signaling pathway
ICAM2	P13598	Down	
MDK	P21741	Up	response to stress signaling process

(continued)

Biomarker	UniProt ID	Direction*	Biological Process (GO)
CCL23	P55773	Up	immune system process response to stress cell communication signaling process signaling
BMP1	P13497	Down	proteolysis
BMPER	Q8N8U9	Down	regulation of signaling pathway
TPT1	P13693	Up	regulation of cell death
F9	P00740	Up	proteolysis
CRK	P46108	Up	signaling process signaling regulation of signaling pathway

Table 19: Biomarkers of general cancer

KIT	KLK3-SERPINA3
C9	BMPER
AFM	VEGFA
CCL23	CDK8-CCNC
SERPINA4	DDC
CRP	APOA1
BMP1	ALB
EGFR	FGA-FGB-FGG
BDNF	FN1
ITIH4	CFHR5
CDK5-CDK5R1	EFNA5

Table 20: Panels of 1 Biomarker

Markers		Mean CV AUC
1	C9	0.792
2	KLK3-SERPINA3	0.782
3	CRP	0.763
4	SERPINA4	0.762
5	AFM	0.750
6	BMPER	0.745
7	ALB	0.737
8	APOA1	0.733
9	BMP1	0.732
10	KIT	0.729

(continued)

Markers		Mean CV AUC
11	EGFR	0.726
12	ITIH4	0.721
13	VEGFA	0.720
14	BDNF	0.720
15	FGA-FGB-FGG	0.712
16	EFNA5	0.697
17	DDC	0.696
18	FN1	0.694
19	CDK8-CCNC	0.692
20	CCL23	0.692
21	CFHR5	0.674
22	CDK5-CDK5R1	0.666

Table 21: Panels of 2 Biomarkers

Markers			Mean CV AUC
1	KLK3-SERPINA3	EGFR	0.826
2	KLK3-SERPINA3	BDNF	0.823
3	KLK3-SERPINA3	EFNA5	0.820
4	KIT	C9	0.819
5	BDNF	C9	0.818
6	KLK3-SERPINA3	BMP1	0.816
7	KLK3-SERPINA3	BMPER	0.816
8	KLK3-SERPINA3	KIT	0.815
9	C9	BMPER	0.814
10	EFNA5	C9	0.812
11	KLK3-SERPINA3	C9	0.811
12	KLK3-SERPINA3	CRP	0.811
13	EGFR	C9	0.811
14	BMPER	CRP	0.810
15	BDNF	CRP	0.810
16	C9	DDC	0.809
17	KLK3-SERPINA3	DDC	0.807
18	KLK3-SERPINA3	ALB	0.806
19	BDNF	SERPINA4	0.805
20	BMP1	CRP	0.805
21	C9	CRP	0.802
22	C9	ALB	0.802

(continued)

Markers			Mean CV AUC
23	KLK3-SERPINA3	CCL23	0.802
24	KLK3-SERPINA3	FN1	0.801
25	BDNF	KIT	0.801
26	EGFR	SERPINA4	0.801
27	KLK3-SERPINA3	CDK5-CDK5R1	0.800
28	EFNA5	CRP	0.799
29	EGFR	ITIH4	0.799
30	BMPER	AFM	0.798
31	C9	BMP1	0.798
32	KIT	CRP	0.798
33	C9	SERPINA4	0.798
34	C9	ITIH4	0.797
35	SERPINA4	BMPER	0.796
36	EFNA5	SERPINA4	0.796
37	KLK3-SERPINA3	APOA1	0.795
38	EGFR	CRP	0.795
39	KIT	SERPINA4	0.795
40	EGFR	AFM	0.795
41	VEGFA	C9	0.795
42	C9	FN1	0.794
43	C9	AFM	0.793
44	KLK3-SERPINA3	AFM	0.793
45	KLK3-SERPINA3	SERPINA4	0.792
46	BMP1	SERPINA4	0.792
47	KIT	BMP1	0.791
48	BDNF	AFM	0.791
49	CCL23	C9	0.791
50	KIT	BMPER	0.790
51	KLK3-SERPINA3	ITIH4	0.790
52	DDC	CRP	0.789
53	CCL23	CRP	0.789
54	C9	CDK5-CDK5R1	0.788
55	BDNF	VEGFA	0.788
56	EGFR	ALB	0.788
57	KIT	AFM	0.787
58	BMPER	ITIH4	0.786
59	BDNF	ALB	0.785
60	KLK3-SERPINA3	CDK8-CCNC	0.785

(continued)

Markers			Mean CV AUC
61	FN1	CRP	0.784
62	BDNF	BMPER	0.784
63	APOA1	C9	0.784
64	C9	CDK8-CCNC	0.784
65	EGFR	BMPER	0.783
66	EFNA5	AFM	0.783
67	VEGFA	CRP	0.783
68	SERPINA4	DDC	0.783
69	CRP	AFM	0.783
70	BMP1	BMPER	0.783
71	DDC	ITIH4	0.783
72	KLK3-SERPINA3	VEGFA	0.782
73	CRP	CDK5-CDK5R1	0.782
74	DDC	AFM	0.782
75	BMP1	AFM	0.782
76	EFNA5	BMPER	0.781
77	CRP	ITIH4	0.781
78	FN1	SERPINA4	0.780
79	BDNF	ITIH4	0.780
80	ALB	CRP	0.779
81	VEGFA	EGFR	0.779
82	EFNA5	BMP1	0.778
83	C9	CFHR5	0.777
84	BDNF	EGFR	0.776
85	SERPINA4	CRP	0.776
86	BDNF	DDC	0.776
87	SERPINA4	AFM	0.775
88	KIT	EGFR	0.775
89	EFNA5	ALB	0.775
90	KLK3-SERPINA3	FGA-FGB-FGG	0.775
91	APOA1	CRP	0.774
92	CDK8-CCNC	CRP	0.774
93	BMP1	ALB	0.774
94	BMP1	DDC	0.774
95	DDC	BMPER	0.774
96	BMP1	ITIH4	0.774
97	EFNA5	EGFR	0.773
98	KIT	ITIH4	0.773

(continued)

Markers			Mean CV AUC
99	EFNA5	APOA1	0.772
100	FN1	AFM	0.772

Table 22: Panels of 3 Biomarkers

Markers				Mean CV AUC
1	BDNF	KIT	C9	0.846
2	KLK3-SERPINA3	BDNF	KIT	0.842
3	KLK3-SERPINA3	KIT	EFNA5	0.838
4	BDNF	KIT	CRP	0.837
5	KLK3-SERPINA3	EFNA5	EGFR	0.836
6	KLK3-SERPINA3	BDNF	C9	0.836
7	KLK3-SERPINA3	EFNA5	C9	0.835
8	KLK3-SERPINA3	EFNA5	BMP1	0.835
9	KLK3-SERPINA3	BDNF	CRP	0.834
10	KLK3-SERPINA3	EFNA5	CRP	0.833
11	KLK3-SERPINA3	KIT	BMP1	0.833
12	BDNF	KIT	SERPINA4	0.833
13	KLK3-SERPINA3	BDNF	EGFR	0.833
14	KLK3-SERPINA3	KIT	EGFR	0.833
15	KLK3-SERPINA3	BDNF	EFNA5	0.833
16	KLK3-SERPINA3	EGFR	ITIH4	0.832
17	KLK3-SERPINA3	EFNA5	BMPER	0.831
18	KLK3-SERPINA3	EGFR	C9	0.831
19	KLK3-SERPINA3	EGFR	BMPER	0.830
20	BDNF	C9	CRP	0.830
21	KIT	C9	BMPER	0.830
22	KIT	EFNA5	C9	0.830
23	KLK3-SERPINA3	EFNA5	ALB	0.829
24	KLK3-SERPINA3	KIT	BMPER	0.829
25	KLK3-SERPINA3	BDNF	DDC	0.829
26	BDNF	EGFR	C9	0.829
27	KLK3-SERPINA3	EGFR	ALB	0.829
28	KLK3-SERPINA3	KIT	C9	0.828
29	KLK3-SERPINA3	BDNF	SERPINA4	0.828
30	KIT	C9	DDC	0.828
31	BDNF	C9	SERPINA4	0.828
32	KLK3-SERPINA3	EGFR	CRP	0.828

(continued)

Markers				Mean CV AUC
33	KLK3-SERPINA3	EGFR	BMP1	0.827
34	BDNF	C9	DDC	0.827
35	KIT	EGFR	C9	0.827
36	KLK3-SERPINA3	BMP1	CRP	0.826
37	KLK3-SERPINA3	C9	BMPER	0.826
38	KLK3-SERPINA3	C9	DDC	0.825
39	KLK3-SERPINA3	BDNF	ALB	0.825
40	KLK3-SERPINA3	EFNA5	SERPINA4	0.825
41	KLK3-SERPINA3	EGFR	DDC	0.825
42	EGFR	C9	ITIH4	0.825
43	KLK3-SERPINA3	EFNA5	APOA1	0.825
44	KLK3-SERPINA3	EGFR	FN1	0.825
45	KIT	C9	BMP1	0.825
46	KLK3-SERPINA3	BMPER	CRP	0.825
47	KLK3-SERPINA3	EFNA5	FN1	0.824
48	KLK3-SERPINA3	BDNF	BMPER	0.824
49	EGFR	C9	BMPER	0.824
50	KLK3-SERPINA3	EGFR	SERPINA4	0.824
51	BDNF	EFNA5	C9	0.824
52	KLK3-SERPINA3	BDNF	BMP1	0.824
53	KLK3-SERPINA3	BDNF	VEGFA	0.824
54	BDNF	VEGFA	C9	0.824
55	KLK3-SERPINA3	EFNA5	CDK8-CCNC	0.824
56	BDNF	C9	BMPER	0.824
57	KIT	BMP1	CRP	0.824
58	BDNF	C9	ALB	0.824
59	KLK3-SERPINA3	EFNA5	ITIH4	0.823
60	KLK3-SERPINA3	CDK5-CDK5R1	KIT	0.823
61	KLK3-SERPINA3	VEGFA	EGFR	0.823
62	KLK3-SERPINA3	BDNF	FN1	0.823
63	EGFR	C9	ALB	0.823
64	KLK3-SERPINA3	KIT	CRP	0.823
65	C9	DDC	BMPER	0.823
66	KLK3-SERPINA3	EGFR	AFM	0.823
67	KLK3-SERPINA3	CDK5-CDK5R1	EFNA5	0.823
68	EFNA5	EGFR	C9	0.823
69	KLK3-SERPINA3	DDC	BMPER	0.823
70	EFNA5	C9	BMPER	0.822

(continued)

Markers				Mean CV AUC
71	KLK3-SERPINA3	BMP1	DDC	0.822
72	EFNA5	BMP1	CRP	0.822
73	BDNF	C9	AFM	0.822
74	C9	BMPER	CRP	0.822
75	KLK3-SERPINA3	BDNF	ITIH4	0.822
76	KLK3-SERPINA3	KIT	FN1	0.822
77	BDNF	EGFR	CRP	0.821
78	KIT	C9	CRP	0.821
79	EGFR	C9	SERPINA4	0.821
80	BDNF	EGFR	SERPINA4	0.821
81	KLK3-SERPINA3	BMP1	BMPER	0.821
82	KIT	EFNA5	SERPINA4	0.821
83	EFNA5	C9	CRP	0.821
84	BDNF	BMPER	CRP	0.821
85	KLK3-SERPINA3	KIT	CCL23	0.821
86	KLK3-SERPINA3	CDK5-CDK5R1	EGFR	0.821
87	KLK3-SERPINA3	EFNA5	DDC	0.821
88	C9	BMPER	ITIH4	0.821
89	KIT	BMPER	CRP	0.821
90	EGFR	C9	AFM	0.820
91	KLK3-SERPINA3	BDNF	AFM	0.820
92	BDNF	BMP1	CRP	0.820
93	KLK3-SERPINA3	EGFR	CCL23	0.820
94	KIT	C9	SERPINA4	0.820
95	BDNF	EFNA5	CRP	0.820
96	EFNA5	C9	ALB	0.820
97	KLK3-SERPINA3	BMPER	ITIH4	0.819
98	KLK3-SERPINA3	CCL23	CRP	0.819
99	EGFR	C9	DDC	0.819
100	KIT	EFNA5	CRP	0.819

Table 23: Panels of 4 Biomarkers

Markers					Mean CV AUC
1	KLK3-SERPINA3	BDNF	KIT	C9	0.849
2	KLK3-SERPINA3	KIT	EFNA5	BMP1	0.848
3	KLK3-SERPINA3	BDNF	KIT	CRP	0.848
4	KLK3-SERPINA3	KIT	EFNA5	C9	0.847

(continued)

Markers					Mean CV AUC
5	KLK3-SERPINA3	BDNF	KIT	EGFR	0.846
6	KLK3-SERPINA3	BDNF	KIT	EFNA5	0.846
7	KLK3-SERPINA3	KIT	EFNA5	CRP	0.845
8	BDNF	KIT	C9	DDC	0.845
9	BDNF	KIT	C9	CRP	0.844
10	BDNF	KIT	EGFR	C9	0.844
11	KLK3-SERPINA3	EFNA5	EGFR	C9	0.844
12	BDNF	KIT	C9	SERPINA4	0.844
13	KLK3-SERPINA3	BDNF	KIT	DDC	0.844
14	KLK3-SERPINA3	EFNA5	EGFR	ITIH4	0.843
15	BDNF	KIT	EFNA5	C9	0.843
16	BDNF	KIT	C9	BMPER	0.843
17	KLK3-SERPINA3	EFNA5	BMP1	CRP	0.843
18	KLK3-SERPINA3	BDNF	EFNA5	C9	0.843
19	KLK3-SERPINA3	KIT	EFNA5	BMPER	0.843
20	KLK3-SERPINA3	BDNF	EFNA5	CRP	0.843
21	KLK3-SERPINA3	BDNF	KIT	SERPINA4	0.842
22	KLK3-SERPINA3	BDNF	EGFR	C9	0.842
23	BDNF	KIT	C9	CDK8-CCNC	0.842
24	KLK3-SERPINA3	KIT	EFNA5	EGFR	0.842
25	BDNF	KIT	BMP1	CRP	0.842
26	KLK3-SERPINA3	EFNA5	EGFR	CRP	0.841
27	BDNF	KIT	VEGFA	C9	0.841
28	KLK3-SERPINA3	BDNF	KIT	BMP1	0.841
29	KLK3-SERPINA3	KIT	EGFR	ITIH4	0.841
30	KLK3-SERPINA3	CDK5-CDK5R1	KIT	EFNA5	0.841
31	BDNF	KIT	C9	ALB	0.841
32	KLK3-SERPINA3	KIT	C9	BMPER	0.841
33	KLK3-SERPINA3	BDNF	KIT	BMPER	0.841
34	KLK3-SERPINA3	BDNF	EGFR	CRP	0.840
35	KLK3-SERPINA3	KIT	EGFR	BMPER	0.840
36	KLK3-SERPINA3	BDNF	C9	CRP	0.840
37	KLK3-SERPINA3	KIT	EFNA5	ITIH4	0.840
38	KLK3-SERPINA3	EFNA5	EGFR	ALB	0.840
39	KLK3-SERPINA3	EFNA5	EGFR	BMP1	0.840
40	KLK3-SERPINA3	BDNF	KIT	FN1	0.840
41	BDNF	KIT	C9	AFM	0.840
42	KLK3-SERPINA3	EFNA5	C9	BMPER	0.840

(continued)

Markers					Mean CV AUC
43	KLK3-SERPINA3	BDNF	KIT	VEGFA	0.840
44	KLK3-SERPINA3	KIT	EFNA5	ALB	0.839
45	KLK3-SERPINA3	KIT	EGFR	BMP1	0.839
46	BDNF	KIT	EFNA5	CRP	0.839
47	BDNF	KIT	C9	FN1	0.839
48	KLK3-SERPINA3	KIT	EFNA5	SERPINA4	0.839
49	KLK3-SERPINA3	BDNF	KIT	ALB	0.839
50	BDNF	KIT	C9	BMP1	0.839
51	KLK3-SERPINA3	KIT	EGFR	C9	0.839
52	KLK3-SERPINA3	BDNF	C9	DDC	0.839
53	KLK3-SERPINA3	EGFR	BMPER	ITIH4	0.839
54	KLK3-SERPINA3	EFNA5	BMPER	CRP	0.838
55	KLK3-SERPINA3	EGFR	DDC	ITIH4	0.838
56	KLK3-SERPINA3	EFNA5	C9	CRP	0.838
57	KLK3-SERPINA3	KIT	EFNA5	FN1	0.838
58	BDNF	KIT	SERPINA4	BMPER	0.838
59	KLK3-SERPINA3	EFNA5	C9	ALB	0.838
60	KLK3-SERPINA3	BDNF	BMP1	CRP	0.838
61	KLK3-SERPINA3	EFNA5	EGFR	BMPER	0.838
62	KLK3-SERPINA3	BDNF	EFNA5	SERPINA4	0.837
63	KLK3-SERPINA3	BDNF	KIT	AFM	0.837
64	KLK3-SERPINA3	BDNF	EFNA5	EGFR	0.837
65	KLK3-SERPINA3	KIT	EGFR	ALB	0.837
66	KLK3-SERPINA3	EFNA5	FN1	CRP	0.837
67	KLK3-SERPINA3	KIT	C9	DDC	0.837
68	BDNF	KIT	EFNA5	SERPINA4	0.837
69	KLK3-SERPINA3	KIT	BMP1	BMPER	0.837
70	KLK3-SERPINA3	BDNF	KIT	ITIH4	0.837
71	KLK3-SERPINA3	EFNA5	C9	DDC	0.837
72	KIT	EFNA5	C9	BMPER	0.837
73	BDNF	KIT	EGFR	SERPINA4	0.837
74	BDNF	KIT	SERPINA4	DDC	0.837
75	KLK3-SERPINA3	BDNF	EGFR	ITIH4	0.837
76	KLK3-SERPINA3	EGFR	C9	BMPER	0.837
77	BDNF	KIT	EGFR	CRP	0.837
78	KLK3-SERPINA3	EGFR	CRP	ITIH4	0.837
79	KLK3-SERPINA3	EFNA5	EGFR	FN1	0.836
80	KLK3-SERPINA3	EFNA5	EGFR	SERPINA4	0.836

(continued)

Markers					Mean CV AUC
81	KLK3-SERPINA3	BDNF	C9	SERPINA4	0.836
82	KLK3-SERPINA3	KIT	BMP1	ALB	0.836
83	BDNF	EGFR	C9	CRP	0.836
84	KLK3-SERPINA3	BDNF	C9	ALB	0.836
85	KLK3-SERPINA3	EFNA5	CRP	ITIH4	0.836
86	KLK3-SERPINA3	EGFR	C9	DDC	0.836
87	KLK3-SERPINA3	BDNF	DDC	CRP	0.836
88	KLK3-SERPINA3	BDNF	EGFR	SERPINA4	0.836
89	KLK3-SERPINA3	BDNF	KIT	CCL23	0.836
90	KLK3-SERPINA3	BDNF	C9	FN1	0.836
91	KIT	EFNA5	C9	SERPINA4	0.836
92	BDNF	KIT	VEGFA	CRP	0.836
93	KLK3-SERPINA3	EFNA5	BMP1	BMPER	0.836
94	BDNF	KIT	BMPER	CRP	0.836
95	KLK3-SERPINA3	KIT	BMP1	CRP	0.836
96	KIT	EFNA5	BMP1	CRP	0.836
97	KLK3-SERPINA3	EGFR	FN1	ITIH4	0.836
98	KLK3-SERPINA3	KIT	BMPER	CRP	0.835
99	KLK3-SERPINA3	KIT	BMP1	DDC	0.835
100	KLK3-SERPINA3	EGFR	C9	ALB	0.835

Table 24: Panels of 5 Biomarkers

Markers						Mean CV AUC
1	KLK3-SERPINA3	BDNF	KIT	EFNA5	C9	0.854
2	KLK3-SERPINA3	BDNF	KIT	EFNA5	CRP	0.853
3	KLK3-SERPINA3	KIT	EFNA5	BMP1	CRP	0.852
4	KLK3-SERPINA3	BDNF	KIT	EGFR	C9	0.851
5	KLK3-SERPINA3	KIT	EFNA5	C9	BMPER	0.851
6	KLK3-SERPINA3	KIT	EFNA5	EGFR	C9	0.850
7	KLK3-SERPINA3	KIT	EFNA5	EGFR	ITIH4	0.849
8	KLK3-SERPINA3	EFNA5	EGFR	CRP	ITIH4	0.849
9	KLK3-SERPINA3	KIT	EFNA5	BMP1	BMPER	0.849
10	KLK3-SERPINA3	KIT	EFNA5	C9	ALB	0.849
11	KLK3-SERPINA3	CDK5-CDK5R1	KIT	EFNA5	C9	0.849
12	KLK3-SERPINA3	BDNF	KIT	C9	ALB	0.849
13	KLK3-SERPINA3	KIT	EFNA5	C9	CRP	0.849
14	KLK3-SERPINA3	BDNF	KIT	C9	DDC	0.849

(continued)

Markers						Mean CV AUC
15	KLK3-SERPINA3	BDNF	KIT	C9	CRP	0.849
16	BDNF	KIT	EFNA5	C9	CRP	0.849
17	KLK3-SERPINA3	BDNF	KIT	BMP1	CRP	0.848
18	KLK3-SERPINA3	KIT	EFNA5	C9	BMP1	0.848
19	KLK3-SERPINA3	BDNF	KIT	C9	BMPER	0.848
20	KLK3-SERPINA3	CDK5-CDK5R1	KIT	EFNA5	CRP	0.848
21	KLK3-SERPINA3	BDNF	EFNA5	C9	CRP	0.848
22	KLK3-SERPINA3	BDNF	KIT	EGFR	ITIH4	0.848
23	KLK3-SERPINA3	BDNF	KIT	EFNA5	SERPINA4	0.848
24	KLK3-SERPINA3	BDNF	KIT	EFNA5	EGFR	0.848
25	KLK3-SERPINA3	KIT	EFNA5	C9	FN1	0.848
26	KLK3-SERPINA3	KIT	EFNA5	C9	DDC	0.848
27	KLK3-SERPINA3	KIT	EFNA5	EGFR	CRP	0.848
28	KLK3-SERPINA3	KIT	EFNA5	EGFR	BMP1	0.848
29	KLK3-SERPINA3	KIT	EGFR	C9	BMPER	0.847
30	KLK3-SERPINA3	KIT	EFNA5	BMPER	CRP	0.847
31	KLK3-SERPINA3	KIT	EFNA5	BMP1	ALB	0.847
32	KLK3-SERPINA3	KIT	EFNA5	FN1	CRP	0.847
33	KLK3-SERPINA3	BDNF	KIT	C9	SERPINA4	0.847
34	BDNF	KIT	EGFR	C9	CRP	0.847
35	KLK3-SERPINA3	BDNF	KIT	EGFR	FN1	0.846
36	KLK3-SERPINA3	KIT	EFNA5	BMP1	FN1	0.846
37	KLK3-SERPINA3	BDNF	KIT	EFNA5	BMP1	0.846
38	KLK3-SERPINA3	KIT	EFNA5	C9	ITIH4	0.846
39	KLK3-SERPINA3	CDK5-CDK5R1	BDNF	KIT	C9	0.846
40	BDNF	KIT	EGFR	C9	SERPINA4	0.846
41	KLK3-SERPINA3	KIT	EFNA5	SERPINA4	BMPER	0.846
42	KLK3-SERPINA3	BDNF	KIT	EGFR	CRP	0.846
43	BDNF	KIT	C9	SERPINA4	BMPER	0.846
44	BDNF	KIT	EFNA5	C9	SERPINA4	0.846
45	KLK3-SERPINA3	BDNF	KIT	EGFR	SERPINA4	0.846
46	KLK3-SERPINA3	EFNA5	EGFR	FN1	ITIH4	0.846
47	KLK3-SERPINA3	BDNF	KIT	EFNA5	ITIH4	0.846
48	KLK3-SERPINA3	CDK5-CDK5R1	KIT	EFNA5	ITIH4	0.846
49	KLK3-SERPINA3	BDNF	KIT	FN1	CRP	0.846
50	KLK3-SERPINA3	KIT	EFNA5	BMP1	SERPINA4	0.846
51	KLK3-SERPINA3	BDNF	KIT	EGFR	ALB	0.846
52	KLK3-SERPINA3	BDNF	KIT	C9	FN1	0.846

(continued)

Markers						Mean CV AUC
53	BDNF	KIT	EGFR	C9	AFM	0.846
54	KLK3-SERPINA3	KIT	EFNA5	CRP	ITIH4	0.846
55	KLK3-SERPINA3	BDNF	KIT	EGFR	BMPER	0.845
56	KLK3-SERPINA3	EFNA5	EGFR	C9	ALB	0.845
57	KLK3-SERPINA3	BDNF	EFNA5	BMP1	CRP	0.845
58	KLK3-SERPINA3	KIT	EFNA5	C9	SERPINA4	0.845
59	KLK3-SERPINA3	KIT	EFNA5	EGFR	ALB	0.845
60	KLK3-SERPINA3	BDNF	KIT	CCL23	C9	0.845
61	KLK3-SERPINA3	BDNF	KIT	SERPINA4	BMPER	0.845
62	KLK3-SERPINA3	CDK5-CDK5R1	BDNF	KIT	CRP	0.845
63	KLK3-SERPINA3	BDNF	EFNA5	EGFR	C9	0.845
64	KLK3-SERPINA3	EFNA5	EGFR	C9	ITIH4	0.845
65	KLK3-SERPINA3	EFNA5	EGFR	DDC	ITIH4	0.845
66	KLK3-SERPINA3	EFNA5	EGFR	C9	CRP	0.845
67	KLK3-SERPINA3	BDNF	KIT	EFNA5	ALB	0.845
68	KLK3-SERPINA3	BDNF	KIT	C9	BMP1	0.845
69	KLK3-SERPINA3	KIT	EFNA5	BMP1	ITIH4	0.845
70	BDNF	KIT	C9	BMPER	CRP	0.845
71	KLK3-SERPINA3	BDNF	EGFR	C9	CRP	0.845
72	BDNF	KIT	EGFR	C9	ALB	0.845
73	KLK3-SERPINA3	BDNF	KIT	VEGFA	EGFR	0.845
74	KLK3-SERPINA3	KIT	EGFR	C9	ALB	0.845
75	KLK3-SERPINA3	KIT	EGFR	DDC	ITIH4	0.845
76	KLK3-SERPINA3	KIT	EFNA5	EGFR	SERPINA4	0.845
77	KLK3-SERPINA3	KIT	EFNA5	C9	CDK8-CCNC	0.844
78	KLK3-SERPINA3	BDNF	KIT	EGFR	BMP1	0.844
79	KLK3-SERPINA3	BDNF	KIT	FN1	SERPINA4	0.844
80	KLK3-SERPINA3	BDNF	KIT	EFNA5	DDC	0.844
81	KLK3-SERPINA3	BDNF	KIT	SERPINA4	DDC	0.844
82	BDNF	KIT	EGFR	C9	ITIH4	0.844
83	BDNF	KIT	EGFR	C9	FN1	0.844
84	BDNF	KIT	C9	BMP1	CRP	0.844
85	KLK3-SERPINA3	KIT	EFNA5	BMP1	AFM	0.844
86	BDNF	KIT	C9	SERPINA4	DDC	0.844
87	KLK3-SERPINA3	KIT	EFNA5	EGFR	FN1	0.844
88	KLK3-SERPINA3	BDNF	KIT	C9	ITIH4	0.844
89	KLK3-SERPINA3	BDNF	KIT	VEGFA	CRP	0.844
90	KLK3-SERPINA3	KIT	EFNA5	EGFR	BMPER	0.844

(continued)

Markers						Mean CV AUC
91	KLK3-SERPINA3	EFNA5	EGFR	C9	FN1	0.844
92	KLK3-SERPINA3	BDNF	KIT	C9	AFM	0.844
93	KLK3-SERPINA3	EFNA5	EGFR	BMP1	CRP	0.844
94	KLK3-SERPINA3	EFNA5	EGFR	C9	BMP1	0.844
95	KLK3-SERPINA3	EFNA5	EGFR	C9	BMPER	0.844
96	KLK3-SERPINA3	BDNF	KIT	VEGFA	C9	0.844
97	KLK3-SERPINA3	BDNF	KIT	C9	CDK8-CCNC	0.844
98	KIT	EFNA5	EGFR	C9	SERPINA4	0.844
99	BDNF	KIT	C9	FN1	SERPINA4	0.844
100	KLK3-SERPINA3	EFNA5	EGFR	C9	DDC	0.844

Table 25: Panels of 6 Biomarkers

Markers						Mean CV AUC
1	KLK3-SERPINA3 CRP	BDNF	KIT	EFNA5	C9	0.856
2	KLK3-SERPINA3 ITIH4	KIT	EFNA5	EGFR	CRP	0.855
3	KLK3-SERPINA3 CRP	BDNF	KIT	EFNA5	BMP1	0.854
4	KLK3-SERPINA3 C9	BDNF	KIT	EFNA5	EGFR	0.854
5	KLK3-SERPINA3 ITIH4	BDNF	KIT	EFNA5	CRP	0.853
6	KLK3-SERPINA3 ALB	BDNF	KIT	EFNA5	C9	0.853
7	KLK3-SERPINA3 BMPER	KIT	EFNA5	EGFR	C9	0.853
8	KLK3-SERPINA3 CDK8-CCNC	BDNF	KIT	EFNA5	C9	0.852
9	KLK3-SERPINA3 SERPINA4	BDNF	KIT	EFNA5	C9	0.852
10	KLK3-SERPINA3 ITIH4	BDNF	KIT	EFNA5	EGFR	0.852
11	KLK3-SERPINA3 ITIH4	KIT	EFNA5	EGFR	C9	0.852
12	KLK3-SERPINA3 CDK5-CDK5R1	BDNF	KIT	EFNA5	CRP	0.852
13	KLK3-SERPINA3 CRP	KIT	EFNA5	BMP1	BMPER	0.852
14	KLK3-SERPINA3 CRP	KIT	EFNA5	BMP1	FN1	0.852

(continued)

Markers						Mean CV AUC
15	KLK3-SERPINA3 ALB	KIT	EFNA5	EGFR	C9	0.852
16	KLK3-SERPINA3 ALB	BDNF	KIT	EGFR	C9	0.852
17	KLK3-SERPINA3 ITIH4	BDNF	KIT	EFNA5	C9	0.852
18	KLK3-SERPINA3 BMPER	KIT	EFNA5	C9	DDC	0.852
19	KLK3-SERPINA3 CRP	BDNF	KIT	EFNA5	EGFR	0.851
20	KLK3-SERPINA3 FN1	KIT	EFNA5	EGFR	C9	0.851
21	KLK3-SERPINA3 ITIH4	KIT	EFNA5	EGFR	FN1	0.851
22	KLK3-SERPINA3 CRP	KIT	EFNA5	EGFR	BMP1	0.851
23	KLK3-SERPINA3 BMP1	KIT	EFNA5	EGFR	C9	0.851
24	KLK3-SERPINA3 ITIH4	KIT	EFNA5	BMP1	CRP	0.851
25	KLK3-SERPINA3 DDC	BDNF	KIT	EFNA5	C9	0.851
26	KLK3-SERPINA3 BMP1	BDNF	KIT	EFNA5	C9	0.851
27	KLK3-SERPINA3 ALB	KIT	EFNA5	C9	FN1	0.851
28	KLK3-SERPINA3 ITIH4	KIT	EFNA5	EGFR	BMP1	0.851
29	KLK3-SERPINA3 CDK5-CDK5R1	BDNF	KIT	EFNA5	C9	0.851
30	KLK3-SERPINA3 FN1	BDNF	KIT	EFNA5	C9	0.851
31	KLK3-SERPINA3 CRP	KIT	EFNA5	EGFR	C9	0.851
32	KLK3-SERPINA3 CDK5-CDK5R1	KIT	EFNA5	EGFR	C9	0.851
33	KLK3-SERPINA3 ALB	KIT	EFNA5	EGFR	BMP1	0.851
34	KLK3-SERPINA3 BMPER	BDNF	KIT	EGFR	C9	0.850
35	KLK3-SERPINA3 CDK5-CDK5R1	KIT	EFNA5	CRP	ITIH4	0.850

(continued)

Markers						Mean CV AUC
36	KLK3-SERPINA3 ITIH4	BDNF	KIT	EGFR	C9	0.850
37	KLK3-SERPINA3 CDK5-CDK5R1	KIT	EFNA5	C9	ITIH4	0.850
38	KLK3-SERPINA3 SERPINA4	KIT	EFNA5	EGFR	C9	0.850
39	KLK3-SERPINA3 FN1	BDNF	KIT	EGFR	C9	0.850
40	KLK3-SERPINA3 ITIH4	KIT	EFNA5	EGFR	ALB	0.850
41	KLK3-SERPINA3 ITIH4	KIT	EFNA5	EGFR	DDC	0.850
42	KLK3-SERPINA3 CRP	BDNF	KIT	EFNA5	SERPINA4	0.850
43	KLK3-SERPINA3 CRP	KIT	EFNA5	C9	BMP1	0.850
44	KLK3-SERPINA3 CRP	BDNF	KIT	EFNA5	FN1	0.850
45	KLK3-SERPINA3 CDK5-CDK5R1	KIT	EFNA5	C9	CRP	0.850
46	KLK3-SERPINA3 ALB	KIT	EFNA5	C9	BMP1	0.850
47	KLK3-SERPINA3 DDC	KIT	EFNA5	EGFR	C9	0.850
48	KLK3-SERPINA3 CRP	BDNF	KIT	EFNA5	CDK8-CCNC	0.850
49	KLK3-SERPINA3 BMPER	KIT	EFNA5	C9	SERPINA4	0.850
50	KLK3-SERPINA3 ITIH4	BDNF	KIT	EGFR	CRP	0.850
51	KLK3-SERPINA3 CRP	KIT	EFNA5	EGFR	BMPER	0.850
52	KLK3-SERPINA3 SERPINA4	BDNF	KIT	EFNA5	EGFR	0.850
53	KLK3-SERPINA3 CRP	BDNF	KIT	EGFR	C9	0.850
54	KLK3-SERPINA3 BMPER	KIT	EFNA5	C9	FN1	0.850
55	KLK3-SERPINA3 CRP	KIT	EFNA5	BMP1	ALB	0.850
56	KLK3-SERPINA3 SERPINA4	BDNF	KIT	EGFR	C9	0.850

(continued)

Markers						Mean CV AUC
57	KLK3-SERPINA3 ITIH4	KIT	EFNA5	C9	BMPER	0.850
58	KLK3-SERPINA3 CRP	BDNF	KIT	BMP1	FN1	0.850
59	KLK3-SERPINA3 ITIH4	KIT	EFNA5	EGFR	BMPER	0.850
60	KLK3-SERPINA3 CRP	BDNF	EFNA5	EGFR	C9	0.850
61	KLK3-SERPINA3 ITIH4	KIT	EFNA5	C9	DDC	0.850
62	KLK3-SERPINA3 BMPER	BDNF	KIT	EFNA5	C9	0.850
63	KLK3-SERPINA3 CRP	KIT	EFNA5	C9	BMPER	0.850
64	KLK3-SERPINA3 C9	BDNF	KIT	VEGFA	EFNA5	0.849
65	KLK3-SERPINA3 CRP	BDNF	KIT	EFNA5	DDC	0.849
66	KLK3-SERPINA3 CRP	KIT	EFNA5	EGFR	FN1	0.849
67	KLK3-SERPINA3 BMPER	KIT	EFNA5	BMP1	SERPINA4	0.849
68	KLK3-SERPINA3 DDC	KIT	EFNA5	C9	FN1	0.849
69	KLK3-SERPINA3 CDK5-CDK5R1	KIT	EFNA5	DDC	ITIH4	0.849
70	KLK3-SERPINA3 CRP	BDNF	KIT	EFNA5	BMPER	0.849
71	KLK3-SERPINA3 BMPER	KIT	EFNA5	C9	BMP1	0.849
72	KLK3-SERPINA3 SERPINA4	KIT	EFNA5	BMP1	FN1	0.849
73	KLK3-SERPINA3 CDK5-CDK5R1	KIT	EFNA5	FN1	CRP	0.849
74	KLK3-SERPINA3 CDK5-CDK5R1	KIT	EFNA5	EGFR	ITIH4	0.849
75	KLK3-SERPINA3 SERPINA4	BDNF	KIT	EGFR	FN1	0.849
76	KLK3-SERPINA3 ITIH4	KIT	EFNA5	FN1	CRP	0.849
77	KLK3-SERPINA3 ITIH4	EFNA5	EGFR	FN1	CRP	0.849

(continued)

Markers						Mean CV AUC
78	KLK3-SERPINA3 SERPINA4	BDNF	KIT	EFNA5	FN1	0.849
79	KLK3-SERPINA3 CRP	BDNF	KIT	C9	BMPER	0.849
80	KLK3-SERPINA3 BMP1	BDNF	KIT	EGFR	C9	0.849
81	KLK3-SERPINA3 CRP	BDNF	KIT	VEGFA	EFNA5	0.849
82	KLK3-SERPINA3 FN1	KIT	EFNA5	EGFR	BMP1	0.849
83	KLK3-SERPINA3 AFM	BDNF	KIT	EGFR	C9	0.849
84	KLK3-SERPINA3 CRP	KIT	EFNA5	C9	FN1	0.849
85	KLK3-SERPINA3 SERPINA4	KIT	EFNA5	EGFR	BMP1	0.849
86	KLK3-SERPINA3 DDC	KIT	EFNA5	C9	BMP1	0.849
87	KLK3-SERPINA3 ALB	KIT	EFNA5	C9	BMPER	0.849
88	KLK3-SERPINA3 ALB	KIT	EFNA5	EGFR	FN1	0.849
89	KLK3-SERPINA3 CRP	KIT	EFNA5	BMP1	DDC	0.849
90	BDNF SERPINA4	KIT	EFNA5	EGFR	C9	0.849
91	KLK3-SERPINA3 ITIH4	EFNA5	EGFR	BMP1	CRP	0.849
92	KLK3-SERPINA3 DDC	BDNF	KIT	EGFR	C9	0.849
93	KLK3-SERPINA3 CRP	BDNF	KIT	EGFR	BMP1	0.849
94	KLK3-SERPINA3 CDK5-CDK5R1	BDNF	KIT	C9	CRP	0.849
95	KLK3-SERPINA3 AFM	BDNF	KIT	EFNA5	C9	0.848
96	KLK3-SERPINA3 CDK8-CCNC	BDNF	KIT	EGFR	C9	0.848
97	KLK3-SERPINA3 CRP	BDNF	KIT	C9	BMP1	0.848
98	KLK3-SERPINA3 CDK5-CDK5R1	KIT	EFNA5	BMP1	CRP	0.848

(continued)

Markers						Mean CV AUC
99	KLK3-SERPINA3 FN1	KIT	EFNA5	C9	BMP1	0.848
100	KLK3-SERPINA3 ALB	BDNF	KIT	EFNA5	SERPINA4	0.848

Table 26: Panels of 7 Biomarkers

Markers						Mean CV AUC
1	KLK3-SERPINA3 CRP	KIT ITIH4	EFNA5	EGFR	FN1	0.855
2	KLK3-SERPINA3 CDK5-CDK5R1	BDNF CRP	KIT	EFNA5	C9	0.855
3	KLK3-SERPINA3 DDC	KIT ITIH4	EFNA5	EGFR	C9	0.855
4	KLK3-SERPINA3 FN1	KIT ALB	EFNA5	EGFR	C9	0.855
5	KLK3-SERPINA3 C9	BDNF CRP	KIT	EFNA5	EGFR	0.855
6	KLK3-SERPINA3 BMP1	BDNF CRP	KIT	EFNA5	C9	0.854
7	KLK3-SERPINA3 FN1	BDNF CRP	KIT	EFNA5	C9	0.854
8	KLK3-SERPINA3 CRP	BDNF ITIH4	KIT	EFNA5	C9	0.854
9	KLK3-SERPINA3 C9	BDNF ITIH4	KIT	EFNA5	EGFR	0.854
10	KLK3-SERPINA3 CRP	KIT ITIH4	EFNA5	EGFR	BMP1	0.854
11	KLK3-SERPINA3 CRP	BDNF ITIH4	KIT	EFNA5	EGFR	0.854
12	KLK3-SERPINA3 CDK8-CCNC	BDNF CRP	KIT	EFNA5	C9	0.854
13	KLK3-SERPINA3 CDK5-CDK5R1	KIT ITIH4	EFNA5	EGFR	C9	0.854
14	KLK3-SERPINA3 BMP1	KIT ALB	EFNA5	EGFR	C9	0.854
15	KLK3-SERPINA3 CRP	KIT ITIH4	EFNA5	EGFR	C9	0.854
16	KLK3-SERPINA3 ALB	BDNF CRP	KIT	EFNA5	C9	0.854
17	KLK3-SERPINA3 BMPER	BDNF CRP	KIT	EFNA5	C9	0.854

(continued)

Markers						Mean CV AUC
18	KLK3-SERPINA3 ALB	KIT ITIH4	EFNA5	EGFR	C9	0.854
19	KLK3-SERPINA3 FN1	KIT ITIH4	EFNA5	EGFR	C9	0.853
20	KLK3-SERPINA3 DDC	BDNF CRP	KIT	EFNA5	C9	0.853
21	KLK3-SERPINA3 BMPER	KIT ITIH4	EFNA5	EGFR	C9	0.853
22	KLK3-SERPINA3 C9	BDNF SERPINA4	KIT	EFNA5	EGFR	0.853
23	KLK3-SERPINA3 SERPINA4	KIT ITIH4	EFNA5	EGFR	FN1	0.853
24	KLK3-SERPINA3 FN1	BDNF CRP	KIT	EFNA5	BMP1	0.853
25	KLK3-SERPINA3 C9	BDNF FN1	KIT	EFNA5	EGFR	0.853
26	KLK3-SERPINA3 FN1	KIT BMPER	EFNA5	EGFR	C9	0.853
27	KLK3-SERPINA3 CDK5-CDK5R1	KIT ITIH4	EFNA5	C9	DDC	0.853
28	KLK3-SERPINA3 SERPINA4	BDNF CRP	KIT	EFNA5	C9	0.853
29	KLK3-SERPINA3 CRP	KIT ITIH4	EFNA5	EGFR	DDC	0.853
30	KLK3-SERPINA3 CRP	KIT ITIH4	EFNA5	EGFR	BMPER	0.853
31	KLK3-SERPINA3 BMPER	KIT ITIH4	EFNA5	C9	DDC	0.853
32	KLK3-SERPINA3 CDK5-CDK5R1	KIT ITIH4	EFNA5	EGFR	CRP	0.853
33	KLK3-SERPINA3 CDK5-CDK5R1	KIT CRP	EFNA5	EGFR	C9	0.853
34	KLK3-SERPINA3 ALB	KIT ITIH4	EFNA5	EGFR	FN1	0.853
35	KLK3-SERPINA3 CDK5-CDK5R1	BDNF ITIH4	KIT	EFNA5	CRP	0.853
36	KLK3-SERPINA3 SERPINA4	KIT ALB	EFNA5	EGFR	C9	0.853
37	KLK3-SERPINA3 BMP1	KIT CRP	EFNA5	EGFR	C9	0.853
38	KLK3-SERPINA3 CRP	KIT ITIH4	EFNA5	BMP1	DDC	0.853

(continued)

Markers						Mean CV AUC
39	KLK3-SERPINA3 BMPER	KIT CRP	EFNA5	EGFR	C9	0.852
40	KLK3-SERPINA3 CDK5-CDK5R1	KIT ITIH4	EFNA5	DDC	CRP	0.852
41	KLK3-SERPINA3 CRP	KIT ITIH4	EFNA5	BMP1	FN1	0.852
42	KLK3-SERPINA3 CRP	BDNF ITIH4	KIT	EFNA5	BMP1	0.852
43	KLK3-SERPINA3 C9	BDNF CRP	KIT	VEGFA	EFNA5	0.852
44	KLK3-SERPINA3 FN1	KIT SERPINA4	EFNA5	EGFR	C9	0.852
45	KLK3-SERPINA3 SERPINA4	BDNF ALB	KIT	EFNA5	C9	0.852
46	KLK3-SERPINA3 C9	BDNF ALB	KIT	EFNA5	EGFR	0.852
47	KLK3-SERPINA3 ALB	KIT CRP	EFNA5	EGFR	C9	0.852
48	KLK3-SERPINA3 FN1	KIT ITIH4	EFNA5	EGFR	BMP1	0.852
49	KLK3-SERPINA3 SERPINA4	KIT ALB	EFNA5	EGFR	BMP1	0.852
50	KLK3-SERPINA3 SERPINA4	KIT BMPER	EFNA5	C9	FN1	0.852
51	KLK3-SERPINA3 FN1	BDNF ITIH4	KIT	EFNA5	EGFR	0.852
52	KLK3-SERPINA3 ALB	KIT CRP	EFNA5	EGFR	BMP1	0.852
53	KLK3-SERPINA3 FN1	BDNF SERPINA4	KIT	EFNA5	C9	0.852
54	KLK3-SERPINA3 CRP	KIT ITIH4	EFNA5	EGFR	ALB	0.852
55	KLK3-SERPINA3 ALB	KIT CRP	EFNA5	EGFR	FN1	0.852
56	KLK3-SERPINA3 C9	BDNF BMP1	KIT	EFNA5	EGFR	0.852
57	KLK3-SERPINA3 FN1	KIT ALB	EFNA5	EGFR	BMP1	0.852
58	KLK3-SERPINA3 SERPINA4	KIT BMPER	EFNA5	EGFR	C9	0.852
59	KLK3-SERPINA3 FN1	BDNF ALB	KIT	EFNA5	C9	0.852

(continued)

Markers						Mean CV AUC
60	KLK3-SERPINA3 FN1	KIT CRP	EFNA5	EGFR	BMP1	0.852
61	KLK3-SERPINA3 FN1	KIT CRP	EFNA5	EGFR	C9	0.852
62	KLK3-SERPINA3 SERPINA4	BDNF ITIH4	KIT	EFNA5	EGFR	0.852
63	KLK3-SERPINA3 CDK5-CDK5R1	KIT FN1	EFNA5	EGFR	C9	0.852
64	KLK3-SERPINA3 DDC	KIT BMPER	EFNA5	EGFR	C9	0.852
65	KLK3-SERPINA3 DDC	BDNF ITIH4	KIT	EFNA5	C9	0.852
66	KLK3-SERPINA3 CRP	BDNF ITIH4	KIT	EFNA5	DDC	0.852
67	KLK3-SERPINA3 BMPER	KIT ITIH4	EFNA5	EGFR	SERPINA4	0.852
68	KLK3-SERPINA3 SERPINA4	BDNF CRP	KIT	EFNA5	BMP1	0.852
69	KLK3-SERPINA3 SERPINA4	KIT ITIH4	EFNA5	EGFR	C9	0.852
70	KLK3-SERPINA3 DDC	KIT ITIH4	EFNA5	EGFR	BMP1	0.852
71	KLK3-SERPINA3 DDC	BDNF ITIH4	KIT	EFNA5	EGFR	0.851
72	KLK3-SERPINA3 ALB	KIT AFM	EFNA5	EGFR	C9	0.851
73	KLK3-SERPINA3 DDC	KIT BMPER	EFNA5	C9	FN1	0.851
74	KLK3-SERPINA3 SERPINA4	BDNF BMPER	KIT	EFNA5	C9	0.851
75	KLK3-SERPINA3 BMP1	BDNF CRP	KIT	VEGFA	EFNA5	0.851
76	KLK3-SERPINA3 DDC	KIT ITIH4	EFNA5	C9	FN1	0.851
77	KLK3-SERPINA3 FN1	BDNF SERPINA4	KIT	EFNA5	EGFR	0.851
78	KLK3-SERPINA3 FN1	KIT SERPINA4	EFNA5	EGFR	BMP1	0.851
79	KLK3-SERPINA3 ALB	BDNF AFM	KIT	EFNA5	C9	0.851
80	KLK3-SERPINA3 BMP1	BDNF CRP	KIT	EFNA5	EGFR	0.851

(continued)

Markers						Mean CV AUC
81	KLK3-SERPINA3 FN1	BDNF ALB	KIT	EGFR	C9	0.851
82	KLK3-SERPINA3 C9	BDNF BMPER	KIT	EFNA5	EGFR	0.851
83	KLK3-SERPINA3 CRP	BDNF ITIH4	KIT	EFNA5	FN1	0.851
84	KLK3-SERPINA3 CRP	BDNF AFM	KIT	EFNA5	C9	0.851
85	KLK3-SERPINA3 BMPER	KIT CRP	EFNA5	BMP1	FN1	0.851
86	KLK3-SERPINA3 BMPER	KIT CRP	EFNA5	C9	FN1	0.851
87	KLK3-SERPINA3 ALB	BDNF CRP	KIT	EFNA5	BMP1	0.851
88	KLK3-SERPINA3 DDC	KIT CRP	EFNA5	BMP1	FN1	0.851
89	KLK3-SERPINA3 ALB	BDNF ITIH4	KIT	EFNA5	C9	0.851
90	KLK3-SERPINA3 C9	BDNF CDK8-CCNC	KIT	EFNA5	EGFR	0.851
91	KLK3-SERPINA3 BMPER	KIT CRP	EFNA5	EGFR	BMP1	0.851
92	KLK3-SERPINA3 CRP	KIT ITIH4	VEGFA	EFNA5	EGFR	0.851
93	KLK3-SERPINA3 SERPINA4	KIT CRP	EFNA5	EGFR	BMP1	0.851
94	KLK3-SERPINA3 CRP	KIT ITIH4	EFNA5	BMP1	BMPER	0.851
95	KLK3-SERPINA3 FN1	BDNF CRP	KIT	EFNA5	EGFR	0.851
96	KLK3-SERPINA3 ALB	KIT ITIH4	EFNA5	EGFR	BMP1	0.851
97	KLK3-SERPINA3 SERPINA4	KIT ALB	EFNA5	EGFR	FN1	0.851
98	KLK3-SERPINA3 BMP1	KIT FN1	EFNA5	EGFR	C9	0.851
99	KLK3-SERPINA3 ALB	BDNF ITIH4	KIT	EGFR	C9	0.851
100	KLK3-SERPINA3 CRP	KIT ITIH4	EFNA5	C9	DDC	0.851

Table 27: Panels of 8 Biomarkers

Markers						Mean CV AUC
1	KLK3-SERPINA3 C9	BDNF CRP	KIT ITIH4	EFNA5	EGFR	0.856
2	KLK3-SERPINA3 FN1	KIT ALB	EFNA5 ITIH4	EGFR	C9	0.856
3	KLK3-SERPINA3 C9	BDNF FN1	KIT SERPINA4	EFNA5	EGFR	0.855
4	KLK3-SERPINA3 DDC	KIT CRP	EFNA5 ITIH4	EGFR	BMP1	0.855
5	KLK3-SERPINA3 FN1	KIT CRP	EFNA5 ITIH4	EGFR	BMP1	0.855
6	KLK3-SERPINA3 FN1	KIT SERPINA4	EFNA5 ALB	EGFR	C9	0.855
7	KLK3-SERPINA3 CDK5-CDK5R1	KIT CRP	EFNA5 ITIH4	EGFR	FN1	0.855
8	KLK3-SERPINA3 FN1	KIT SERPINA4	EFNA5 ALB	EGFR	BMP1	0.855
9	KLK3-SERPINA3 CDK5-CDK5R1	KIT CRP	EFNA5 ITIH4	EGFR	C9	0.855
10	KLK3-SERPINA3 FN1	BDNF CRP	KIT ITIH4	EFNA5	EGFR	0.855
11	KLK3-SERPINA3 C9	BDNF FN1	KIT ITIH4	EFNA5	EGFR	0.855
12	KLK3-SERPINA3 CDK5-CDK5R1	KIT DDC	EFNA5 ITIH4	EGFR	C9	0.855
13	KLK3-SERPINA3 DDC	KIT BMPER	EFNA5 ITIH4	EGFR	C9	0.855
14	KLK3-SERPINA3 FN1	KIT CRP	EFNA5 ITIH4	EGFR	C9	0.855
15	KLK3-SERPINA3 BMP1	KIT FN1	EFNA5 ALB	EGFR	C9	0.855
16	KLK3-SERPINA3 BMP1	KIT SERPINA4	EFNA5 ALB	EGFR	C9	0.854
17	KLK3-SERPINA3 C9	BDNF FN1	KIT ALB	EFNA5	EGFR	0.854
18	KLK3-SERPINA3 C9	BDNF BMP1	KIT CRP	EFNA5	EGFR	0.854
19	KLK3-SERPINA3 CDK5-CDK5R1	KIT FN1	EFNA5 ITIH4	EGFR	C9	0.854
20	KLK3-SERPINA3 C9	BDNF FN1	KIT CRP	EFNA5	EGFR	0.854
21	KLK3-SERPINA3 FN1	KIT SERPINA4	EFNA5 ITIH4	EGFR	BMP1	0.854

(continued)

Markers						Mean CV AUC
22	KLK3-SERPINA3 FN1	KIT DDC	EFNA5 ITIH4	EGFR	C9	0.854
23	KLK3-SERPINA3 FN1	KIT BMPER	EFNA5 ITIH4	EGFR	C9	0.854
24	KLK3-SERPINA3 C9	BDNF DDC	KIT ITIH4	EFNA5	EGFR	0.854
25	KLK3-SERPINA3 DDC	KIT CRP	EFNA5 ITIH4	EGFR	C9	0.854
26	KLK3-SERPINA3 FN1	KIT SERPINA4	EFNA5 ITIH4	EGFR	C9	0.854
27	KLK3-SERPINA3 ALB	KIT CRP	EFNA5 ITIH4	EGFR	BMP1	0.854
28	KLK3-SERPINA3 FN1	KIT SERPINA4	EFNA5 BMPER	EGFR	BMP1	0.854
29	KLK3-SERPINA3 FN1	BDNF CRP	KIT ITIH4	EFNA5	C9	0.854
30	KLK3-SERPINA3 C9	BDNF ALB	KIT ITIH4	EFNA5	EGFR	0.854
31	KLK3-SERPINA3 FN1	KIT CRP	VEGFA ITIH4	EFNA5	EGFR	0.854
32	KLK3-SERPINA3 BMP1	BDNF ALB	KIT CRP	EFNA5	C9	0.854
33	KLK3-SERPINA3 BMPER	KIT CRP	EFNA5 ITIH4	EGFR	BMP1	0.854
34	KLK3-SERPINA3 FN1	KIT SERPINA4	EFNA5 CRP	EGFR	BMP1	0.854
35	KLK3-SERPINA3 BMP1	BDNF FN1	KIT CRP	EFNA5	C9	0.854
36	KLK3-SERPINA3 BMP1	KIT ALB	EFNA5 CRP	EGFR	C9	0.854
37	KLK3-SERPINA3 C9	BDNF ALB	KIT CRP	EFNA5	EGFR	0.854
38	KLK3-SERPINA3 FN1	BDNF SERPINA4	KIT ITIH4	EFNA5	EGFR	0.854
39	KLK3-SERPINA3 FN1	BDNF SERPINA4	KIT CRP	EFNA5	C9	0.854
40	KLK3-SERPINA3 DDC	BDNF CRP	KIT ITIH4	EFNA5	EGFR	0.854
41	KLK3-SERPINA3 DDC	KIT CRP	EFNA5 ITIH4	EGFR	FN1	0.854
42	KLK3-SERPINA3 FN1	KIT ALB	EFNA5 CRP	EGFR	C9	0.854

(continued)

Markers						Mean CV AUC
43	KLK3-SERPINA3 CDK5-CDK5R1	KIT FN1	EFNA5 CRP	EGFR	C9	0.854
44	KLK3-SERPINA3 BMPER	BDNF CRP	KIT ITIH4	EFNA5	C9	0.854
45	KLK3-SERPINA3 C9	BDNF SERPINA4	KIT ALB	EFNA5	EGFR	0.854
46	KLK3-SERPINA3 ALB	KIT CRP	EFNA5 ITIH4	EGFR	FN1	0.853
47	KLK3-SERPINA3 BMPER	KIT CRP	EFNA5 ITIH4	EGFR	FN1	0.853
48	KLK3-SERPINA3 C9	BDNF SERPINA4	KIT CRP	EFNA5	EGFR	0.853
49	KLK3-SERPINA3 ALB	KIT CRP	EFNA5 ITIH4	EGFR	C9	0.853
50	KLK3-SERPINA3 FN1	KIT ALB	EFNA5 ITIH4	EGFR	BMP1	0.853
51	KLK3-SERPINA3 C9	KIT FN1	VEGFA ITIH4	EFNA5	EGFR	0.853
52	KLK3-SERPINA3 FN1	KIT BMPER	EFNA5 CRP	EGFR	C9	0.853
53	KLK3-SERPINA3 SERPINA4	KIT ALB	EFNA5 ITIH4	EGFR	FN1	0.853
54	KLK3-SERPINA3 CDK5-CDK5R1	BDNF FN1	KIT CRP	EFNA5	C9	0.853
55	KLK3-SERPINA3 CDK5-CDK5R1	KIT CRP	EFNA5 ITIH4	EGFR	DDC	0.853
56	KLK3-SERPINA3 SERPINA4	KIT BMPER	EFNA5 ITIH4	EGFR	C9	0.853
57	KLK3-SERPINA3 DDC	KIT BMPER	EFNA5 ITIH4	C9	FN1	0.853
58	KLK3-SERPINA3 SERPINA4	KIT BMPER	EFNA5 ITIH4	EGFR	FN1	0.853
59	KLK3-SERPINA3 CDK5-CDK5R1	BDNF CRP	KIT ITIH4	EFNA5	C9	0.853
60	KLK3-SERPINA3 SERPINA4	KIT ALB	EFNA5 ITIH4	EGFR	C9	0.853
61	KLK3-SERPINA3 BMP1	KIT ALB	EFNA5 ITIH4	EGFR	C9	0.853
62	KLK3-SERPINA3 C9	BDNF FN1	KIT BMPER	EFNA5	EGFR	0.853
63	KLK3-SERPINA3 FN1	KIT SERPINA4	EFNA5 BMPER	EGFR	C9	0.853

(continued)

Markers						Mean CV AUC
64	KLK3-SERPINA3 SERPINA4	BDNF CRP	KIT ITIH4	EFNA5	EGFR	0.853
65	KLK3-SERPINA3 EGFR	BDNF C9	KIT CRP	VEGFA	EFNA5	0.853
66	KLK3-SERPINA3 BMP1	KIT CRP	EFNA5 ITIH4	EGFR	C9	0.853
67	KLK3-SERPINA3 BMP1	BDNF SERPINA4	KIT CRP	EFNA5	C9	0.853
68	KLK3-SERPINA3 FN1	BDNF SERPINA4	KIT ALB	EFNA5	EGFR	0.853
69	KLK3-SERPINA3 FN1	BDNF BMPER	KIT CRP	EFNA5	C9	0.853
70	KLK3-SERPINA3 BMP1	BDNF DDC	KIT CRP	EFNA5	C9	0.853
71	KLK3-SERPINA3 C9	BDNF BMP1	KIT FN1	EFNA5	EGFR	0.853
72	KLK3-SERPINA3 BMPER	KIT CRP	EFNA5 ITIH4	EGFR	C9	0.853
73	KLK3-SERPINA3 DDC	BDNF CRP	KIT ITIH4	EFNA5	C9	0.853
74	KLK3-SERPINA3 SERPINA4	KIT DDC	EFNA5 ITIH4	EGFR	C9	0.853
75	KLK3-SERPINA3 FN1	KIT BMPER	EFNA5 ALB	EGFR	C9	0.853
76	KLK3-SERPINA3 FN1	BDNF ALB	KIT CRP	EFNA5	C9	0.853
77	KLK3-SERPINA3 C9	BDNF SERPINA4	KIT BMPER	EFNA5	EGFR	0.853
78	KLK3-SERPINA3 CDK5-CDK5R1	KIT FN1	EFNA5 ALB	EGFR	C9	0.853
79	KLK3-SERPINA3 EGFR	BDNF C9	KIT FN1	VEGFA	EFNA5	0.853
80	KLK3-SERPINA3 BMP1	KIT FN1	EFNA5 CRP	EGFR	C9	0.853
81	KLK3-SERPINA3 CDK8-CCNC	BDNF FN1	KIT CRP	EFNA5	C9	0.853
82	KLK3-SERPINA3 DDC	KIT ALB	EFNA5 ITIH4	EGFR	C9	0.853
83	KLK3-SERPINA3 C9	BDNF SERPINA4	KIT ITIH4	EFNA5	EGFR	0.853
84	KLK3-SERPINA3 CDK5-CDK5R1	KIT CRP	EFNA5 ITIH4	C9	DDC	0.853

(continued)

Markers						Mean CV AUC
85	KLK3-SERPINA3 C9	BDNF ALB	KIT AFM	EFNA5	EGFR	0.853
86	KLK3-SERPINA3 FN1	KIT DDC	EFNA5 ALB	EGFR	C9	0.853
87	KLK3-SERPINA3 FN1	KIT ALB	EFNA5 AFM	EGFR	C9	0.853
88	KLK3-SERPINA3 SERPINA4	KIT CRP	EFNA5 ITIH4	EGFR	FN1	0.853
89	KLK3-SERPINA3 C9	BDNF CRP	KIT ITIH4	VEGFA	EFNA5	0.853
90	KLK3-SERPINA3 CRP	KIT AFM	EFNA5 ITIH4	EGFR	BMPER	0.853
91	KLK3-SERPINA3 CDK5-CDK5R1	KIT CRP	EFNA5 ITIH4	BMP1	DDC	0.853
92	KLK3-SERPINA3 CDK5-CDK5R1	KIT CRP	EFNA5 ITIH4	EGFR	BMP1	0.853
93	KLK3-SERPINA3 FN1	BDNF SERPINA4	KIT ALB	EFNA5	C9	0.853
94	KLK3-SERPINA3 CDK5-CDK5R1	KIT ALB	EFNA5 ITIH4	EGFR	C9	0.853
95	KLK3-SERPINA3 CDK5-CDK5R1	KIT FN1	EFNA5 SERPINA4	EGFR	C9	0.853
96	KLK3-SERPINA3 DDC	KIT BMPER	EFNA5 ITIH4	EGFR	FN1	0.853
97	KLK3-SERPINA3 ALB	BDNF CRP	KIT ITIH4	EFNA5	EGFR	0.852
98	KLK3-SERPINA3 SERPINA4	KIT BMPER	EFNA5 ITIH4	C9	FN1	0.852
99	KLK3-SERPINA3 CDK5-CDK5R1	BDNF C9	KIT CRP	EFNA5	EGFR	0.852
100	KLK3-SERPINA3 FN1	KIT ALB	EFNA5 CRP	EGFR	BMP1	0.852

Table 28: Panels of 9 Biomarkers

Markers						Mean CV AUC
1	KLK3-SERPINA3 FN1	KIT SERPINA4	EFNA5 ALB	EGFR ITIH4	C9	0.856
2	KLK3-SERPINA3 CDK5-CDK5R1	KIT FN1	EFNA5 CRP	EGFR ITIH4	C9	0.856
3	KLK3-SERPINA3 C9	BDNF FN1	KIT CRP	EFNA5 ITIH4	EGFR	0.856

(continued)

Markers						Mean CV AUC
4	KLK3-SERPINA3 FN1	KIT BMPER	EFNA5 CRP	EGFR ITIH4	C9	0.855
5	KLK3-SERPINA3 FN1	KIT SERPINA4	EFNA5 BMPER	EGFR ITIH4	C9	0.855
6	KLK3-SERPINA3 C9	BDNF ALB	KIT CRP	EFNA5 ITIH4	EGFR	0.855
7	KLK3-SERPINA3 FN1	KIT SERPINA4	EFNA5 ALB	EGFR ITIH4	BMP1	0.855
8	KLK3-SERPINA3 FN1	KIT SERPINA4	EFNA5 CRP	EGFR ITIH4	BMP1	0.855
9	KLK3-SERPINA3 C9	BDNF FN1	KIT SERPINA4	EFNA5 ALB	EGFR	0.855
10	KLK3-SERPINA3 CDK5-CDK5R1	KIT DDC	EFNA5 CRP	EGFR ITIH4	C9	0.855
11	KLK3-SERPINA3 CDK5-CDK5R1	KIT FN1	EFNA5 ALB	EGFR ITIH4	C9	0.855
12	KLK3-SERPINA3 BMP1	KIT FN1	EFNA5 SERPINA4	EGFR ALB	C9	0.855
13	KLK3-SERPINA3 BMP1	KIT FN1	EFNA5 ALB	EGFR ITIH4	C9	0.855
14	KLK3-SERPINA3 FN1	BDNF SERPINA4	KIT ALB	EFNA5 ITIH4	EGFR	0.855
15	KLK3-SERPINA3 FN1	KIT DDC	EFNA5 CRP	EGFR ITIH4	C9	0.855
16	KLK3-SERPINA3 C9	BDNF FN1	KIT SERPINA4	EFNA5 ITIH4	EGFR	0.855
17	KLK3-SERPINA3 BMP1	KIT ALB	EFNA5 CRP	EGFR ITIH4	C9	0.855
18	KLK3-SERPINA3 C9	BDNF BMP1	KIT SERPINA4	EFNA5 CRP	EGFR	0.855
19	KLK3-SERPINA3 FN1	KIT DDC	EFNA5 CRP	EGFR ITIH4	BMP1	0.855
20	KLK3-SERPINA3 C9	BDNF FN1	KIT SERPINA4	EFNA5 CRP	EGFR	0.854
21	KLK3-SERPINA3 C9	BDNF SERPINA4	KIT CRP	EFNA5 ITIH4	EGFR	0.854
22	KLK3-SERPINA3 FN1	KIT ALB	EFNA5 CRP	EGFR ITIH4	BMP1	0.854
23	KLK3-SERPINA3 FN1	KIT ALB	EFNA5 CRP	EGFR ITIH4	C9	0.854
24	KLK3-SERPINA3 C9	BDNF FN1	KIT ALB	EFNA5 ITIH4	EGFR	0.854

(continued)

Markers						Mean CV AUC
25	KLK3-SERPINA3 C9	KIT FN1	VEGFA ALB	EFNA5 ITIH4	EGFR	0.854
26	KLK3-SERPINA3 SERPINA4	KIT DDC	EFNA5 BMPER	EGFR ITIH4	C9	0.854
27	KLK3-SERPINA3 FN1	KIT DDC	EFNA5 BMPER	EGFR ITIH4	C9	0.854
28	KLK3-SERPINA3 FN1	KIT BMPER	EFNA5 ALB	EGFR ITIH4	C9	0.854
29	KLK3-SERPINA3 FN1	KIT ALB	EFNA5 AFM	EGFR ITIH4	C9	0.854
30	KLK3-SERPINA3 C9	BDNF BMP1	KIT FN1	EFNA5 CRP	EGFR	0.854
31	KLK3-SERPINA3 DDC	KIT BMPER	EFNA5 AFM	EGFR ITIH4	C9	0.854
32	KLK3-SERPINA3 C9	BDNF ALB	KIT AFM	EFNA5 ITIH4	EGFR	0.854
33	KLK3-SERPINA3 C9	BDNF BMP1	KIT ALB	EFNA5 CRP	EGFR	0.854
34	KLK3-SERPINA3 C9	BDNF FN1	KIT AFM	EFNA5 ITIH4	EGFR	0.854
35	KLK3-SERPINA3 C9	BDNF FN1	KIT BMPER	EFNA5 CRP	EGFR	0.854
36	KLK3-SERPINA3 C9	BDNF FN1	KIT SERPINA4	EFNA5 BMPER	EGFR	0.854
37	KLK3-SERPINA3 BMP1	KIT FN1	EFNA5 ALB	EGFR CRP	C9	0.854
38	KLK3-SERPINA3 FN1	KIT SERPINA4	EFNA5 BMPER	EGFR ITIH4	BMP1	0.854
39	KLK3-SERPINA3 FN1	BDNF ALB	KIT CRP	EFNA5 ITIH4	EGFR	0.854
40	KLK3-SERPINA3 C9	BDNF FN1	KIT ALB	EFNA5 CRP	EGFR	0.854
41	KLK3-SERPINA3 C9	BDNF BMPER	KIT CRP	EFNA5 ITIH4	EGFR	0.854
42	KLK3-SERPINA3 FN1	BDNF SERPINA4	KIT CRP	EFNA5 ITIH4	EGFR	0.854
43	KLK3-SERPINA3 CDK5-CDK5R1	BDNF C9	KIT CRP	EFNA5 ITIH4	EGFR	0.854
44	KLK3-SERPINA3 BMP1	KIT SERPINA4	EFNA5 BMPER	EGFR ALB	C9	0.854
45	KLK3-SERPINA3 DDC	KIT BMPER	EFNA5 CRP	EGFR ITIH4	FN1	0.854

(continued)

Markers						Mean CV AUC
46	KLK3-SERPINA3 DDC	KIT BMPER	EFNA5 CRP	EGFR ITIH4	C9	0.854
47	KLK3-SERPINA3 CDK5-CDK5R1	BDNF C9	KIT FN1	EFNA5 CRP	EGFR	0.854
48	KLK3-SERPINA3 CDK5-CDK5R1	KIT ALB	EFNA5 CRP	EGFR ITIH4	FN1	0.854
49	KLK3-SERPINA3 C9	BDNF DDC	KIT CRP	EFNA5 ITIH4	EGFR	0.854
50	KLK3-SERPINA3 DDC	KIT BMPER	EFNA5 AFM	EGFR ITIH4	FN1	0.854
51	KLK3-SERPINA3 C9	BDNF FN1	KIT DDC	EFNA5 ITIH4	EGFR	0.854
52	KLK3-SERPINA3 SERPINA4	KIT DDC	EFNA5 BMPER	C9 ITIH4	FN1	0.854
53	KLK3-SERPINA3 CDK5-CDK5R1	KIT FN1	EFNA5 SERPINA4	EGFR ITIH4	C9	0.854
54	KLK3-SERPINA3 C9	KIT FN1	VEGFA CRP	EFNA5 ITIH4	EGFR	0.854
55	KLK3-SERPINA3 FN1	BDNF CRP	KIT AFM	EFNA5 ITIH4	EGFR	0.854
56	KLK3-SERPINA3 C9	BDNF SERPINA4	KIT ALB	EFNA5 ITIH4	EGFR	0.854
57	KLK3-SERPINA3 FN1	KIT BMPER	EFNA5 CRP	EGFR ITIH4	BMP1	0.854
58	KLK3-SERPINA3 FN1	KIT SERPINA4	EFNA5 BMPER	EGFR CRP	C9	0.854
59	KLK3-SERPINA3 C9	BDNF BMP1	KIT FN1	EFNA5 ALB	EGFR	0.854
60	KLK3-SERPINA3 C9	BDNF FN1	KIT ALB	EFNA5 AFM	EGFR	0.854
61	KLK3-SERPINA3 BMP1	KIT FN1	VEGFA CRP	EFNA5 ITIH4	EGFR	0.854
62	KLK3-SERPINA3 BMP1	BDNF FN1	KIT CRP	EFNA5 ITIH4	EGFR	0.854
63	KLK3-SERPINA3 FN1	BDNF SERPINA4	KIT DDC	EFNA5 ITIH4	EGFR	0.854
64	KLK3-SERPINA3 EGFR	BDNF C9	KIT CRP	VEGFA ITIH4	EFNA5	0.854
65	KLK3-SERPINA3 BMPER	KIT CRP	EFNA5 AFM	EGFR ITIH4	FN1	0.854
66	KLK3-SERPINA3 CDK5-CDK5R1	BDNF C9	KIT FN1	VEGFA CRP	EFNA5	0.854

(continued)

Markers						Mean CV AUC
67	KLK3-SERPINA3 CDK5-CDK5R1	KIT FN1	EFNA5 CRP	EGFR ITIH4	BMP1	0.854
68	KLK3-SERPINA3 FN1	KIT SERPINA4	EFNA5 ALB	EGFR CRP	C9	0.854
69	KLK3-SERPINA3 SERPINA4	KIT BMPER	EFNA5 AFM	EGFR ITIH4	FN1	0.854
70	KLK3-SERPINA3 DDC	KIT BMPER	EFNA5 CRP	C9 ITIH4	FN1	0.854
71	KLK3-SERPINA3 C9	BDNF SERPINA4	KIT ALB	EFNA5 CRP	EGFR	0.854
72	KLK3-SERPINA3 CDK5-CDK5R1	KIT CRP	EFNA5 AFM	EGFR ITIH4	FN1	0.854
73	KLK3-SERPINA3 C9	BDNF CDK8-CCNC	KIT CRP	EFNA5 ITIH4	EGFR	0.854
74	KLK3-SERPINA3 EGFR	BDNF C9	KIT FN1	VEGFA CRP	EFNA5	0.854
75	KLK3-SERPINA3 FN1	KIT DDC	EFNA5 ALB	EGFR ITIH4	C9	0.854
76	KLK3-SERPINA3 CDK5-CDK5R1	KIT SERPINA4	EFNA5 DDC	EGFR ITIH4	C9	0.853
77	KLK3-SERPINA3 C9	BDNF FN1	KIT SERPINA4	EFNA5 DDC	EGFR	0.853
78	KLK3-SERPINA3 C9	BDNF CRP	KIT AFM	EFNA5 ITIH4	EGFR	0.853
79	KLK3-SERPINA3 FN1	KIT SERPINA4	EFNA5 DDC	EGFR ITIH4	C9	0.853
80	KLK3-SERPINA3 EGFR	BDNF C9	KIT FN1	VEGFA SERPINA4	EFNA5	0.853
81	KLK3-SERPINA3 FN1	KIT DDC	EFNA5 BMPER	EGFR CRP	C9	0.853
82	KLK3-SERPINA3 CDK5-CDK5R1	KIT FN1	EFNA5 DDC	EGFR ITIH4	C9	0.853
83	KLK3-SERPINA3 SERPINA4	KIT DDC	EFNA5 CRP	EGFR ITIH4	BMP1	0.853
84	KLK3-SERPINA3 C9	BDNF DDC	KIT ALB	EFNA5 ITIH4	EGFR	0.853
85	KLK3-SERPINA3 DDC	KIT BMPER	EFNA5 CRP	EGFR ITIH4	BMP1	0.853
86	KLK3-SERPINA3 BMP1	KIT DDC	EFNA5 CRP	EGFR ITIH4	C9	0.853
87	KLK3-SERPINA3 FN1	KIT DDC	EFNA5 BMPER	EGFR ALB	C9	0.853

(continued)

Markers						Mean CV AUC
88	KLK3-SERPINA3 BMP1	KIT DDC	EFNA5 BMPER	EGFR ITIH4	C9	0.853
89	KLK3-SERPINA3 C9	BDNF BMP1	KIT SERPINA4	EFNA5 ALB	EGFR	0.853
90	KLK3-SERPINA3 CDK5-CDK5R1	KIT SERPINA4	EFNA5 ALB	EGFR ITIH4	C9	0.853
91	KLK3-SERPINA3 C9	BDNF FN1	KIT BMPER	EFNA5 AFM	EGFR	0.853
92	KLK3-SERPINA3 FN1	KIT SERPINA4	EFNA5 ALB	EGFR CRP	BMP1	0.853
93	KLK3-SERPINA3 FN1	KIT SERPINA4	EFNA5 DDC	EGFR BMPER	C9	0.853
94	KLK3-SERPINA3 CDK5-CDK5R1	KIT FN1	EFNA5 SERPINA4	EGFR ALB	C9	0.853
95	KLK3-SERPINA3 SERPINA4	KIT ALB	EFNA5 AFM	EGFR ITIH4	FN1	0.853
96	KLK3-SERPINA3 FN1	BDNF SERPINA4	KIT BMPER	EFNA5 ITIH4	EGFR	0.853
97	KLK3-SERPINA3 BMP1	KIT FN1	EFNA5 CRP	EGFR ITIH4	C9	0.853
98	KLK3-SERPINA3 BMP1	KIT DDC	EFNA5 ALB	EGFR ITIH4	C9	0.853
99	KLK3-SERPINA3 BMP1	KIT SERPINA4	EFNA5 ALB	EGFR ITIH4	C9	0.853
100	KLK3-SERPINA3 CDK5-CDK5R1	BDNF DDC	KIT CRP	EFNA5 ITIH4	C9	0.853

Table 29: Panels of 10 Biomarkers

Markers						Mean CV AUC
1	KLK3-SERPINA3 BMP1	KIT FN1	EFNA5 SERPINA4	EGFR ALB	C9 ITIH4	0.856
2	KLK3-SERPINA3 FN1	KIT DDC	EFNA5 BMPER	EGFR CRP	C9 ITIH4	0.856
3	KLK3-SERPINA3 C9	BDNF FN1	KIT SERPINA4	EFNA5 ALB	EGFR ITIH4	0.856
4	KLK3-SERPINA3 CDK5-CDK5R1	KIT FN1	EFNA5 DDC	EGFR CRP	C9 ITIH4	0.856
5	KLK3-SERPINA3 CDK5-CDK5R1	KIT FN1	EFNA5 ALB	EGFR CRP	C9 ITIH4	0.856

(continued)

Markers						Mean CV AUC
6 7	KLK3-SERPINA3 C9	BDNF FN1	KIT ALB	EFNA5 CRP	EGFR ITIH4	0.856
	KLK3-SERPINA3 C9	BDNF BMP1	KIT SERPINA4	EFNA5 ALB	EGFR CRP	0.855
8	KLK3-SERPINA3 C9	BDNF FN1	KIT ALB	EFNA5 AFM	EGFR ITIH4	0.855
9	KLK3-SERPINA3 CDK5-CDK5R1	KIT FN1	EFNA5 SERPINA4	EGFR ALB	C9 ITIH4	0.855
10	KLK3-SERPINA3 FN1	KIT SERPINA4	EFNA5 ALB	EGFR CRP	BMP1 ITIH4	0.855
11	KLK3-SERPINA3 C9	BDNF BMP1	KIT FN1	EFNA5 SERPINA4	EGFR ALB	0.855
12	KLK3-SERPINA3 FN1	KIT SERPINA4	EFNA5 BMPER	EGFR ALB	C9 ITIH4	0.855
13	KLK3-SERPINA3 C9	BDNF SERPINA4	KIT ALB	EFNA5 CRP	EGFR ITIH4	0.855
14	KLK3-SERPINA3 C9	BDNF BMP1	KIT FN1	EFNA5 ALB	EGFR CRP	0.855
15	KLK3-SERPINA3 FN1	KIT DDC	EFNA5 BMPER	EGFR CRP	BMP1 ITIH4	0.855
16	KLK3-SERPINA3 FN1	KIT SERPINA4	EFNA5 ALB	EGFR CRP	C9 ITIH4	0.855
17	KLK3-SERPINA3 FN1	KIT SERPINA4	EFNA5 BMPER	EGFR CRP	C9 ITIH4	0.855
18	KLK3-SERPINA3 FN1	KIT BMPER	EFNA5 CRP	EGFR AFM	C9 ITIH4	0.855
19	KLK3-SERPINA3 C9	BDNF FN1	KIT BMPER	EFNA5 CRP	EGFR ITIH4	0.855
20	KLK3-SERPINA3 FN1	KIT SERPINA4	EFNA5 DDC	EGFR BMPER	C9 ITIH4	0.855
21	KLK3-SERPINA3 CDK5-CDK5R1	KIT C9	VEGFA FN1	EFNA5 CRP	EGFR ITIH4	0.855
22	KLK3-SERPINA3 C9	BDNF FN1	KIT SERPINA4	EFNA5 ALB	EGFR AFM	0.855
23	KLK3-SERPINA3 C9	BDNF FN1	KIT SERPINA4	EFNA5 CRP	EGFR ITIH4	0.855
24	KLK3-SERPINA3 FN1	KIT SERPINA4	EFNA5 DDC	EGFR ALB	C9 ITIH4	0.855
25	KLK3-SERPINA3 FN1	KIT DDC	EFNA5 BMPER	EGFR AFM	C9 ITIH4	0.855
26	KLK3-SERPINA3 C9	KIT FN1	VEGFA ALB	EFNA5 CRP	EGFR ITIH4	0.854

(continued)

Markers						Mean CV AUC
27	KLK3-SERPINA3 C9	BDNF BMP1	KIT FN1	EFNA5 SERPINA4	EGFR CRP	0.854
28	KLK3-SERPINA3 BMP1	KIT FN1	EFNA5 ALB	EGFR CRP	C9 ITIH4	0.854
29	KLK3-SERPINA3 C9	BDNF FN1	KIT SERPINA4	EFNA5 DDC	EGFR ITIH4	0.854
30	KLK3-SERPINA3 C9	BDNF BMP1	KIT ALB	EFNA5 CRP	EGFR ITIH4	0.854
31	KLK3-SERPINA3 CDK5-CDK5R1	KIT FN1	EFNA5 SERPINA4	EGFR CRP	C9 ITIH4	0.854
32	KLK3-SERPINA3 FN1	KIT SERPINA4	EFNA5 DDC	EGFR CRP	BMP1 ITIH4	0.854
33	KLK3-SERPINA3 CDK5-CDK5R1	BDNF C9	KIT FN1	EFNA5 CRP	EGFR ITIH4	0.854
34	KLK3-SERPINA3 C9	BDNF FN1	KIT SERPINA4	EFNA5 BMPER	EGFR AFM	0.854
35	KLK3-SERPINA3 C9	BDNF FN1	KIT SERPINA4	EFNA5 ALB	EGFR CRP	0.854
36	KLK3-SERPINA3 FN1	KIT SERPINA4	EFNA5 BMPER	EGFR AFM	C9 ITIH4	0.854
37	KLK3-SERPINA3 C9	BDNF FN1	KIT SERPINA4	EFNA5 BMPER	EGFR ITIH4	0.854
38	KLK3-SERPINA3 EGFR	BDNF C9	KIT FN1	VEGFA CRP	EFNA5 ITIH4	0.854
39	KLK3-SERPINA3 FN1	KIT SERPINA4	EFNA5 BMPER	EGFR AFM	BMP1 ITIH4	0.854
40	KLK3-SERPINA3 CDK5-CDK5R1	KIT FN1	EFNA5 BMPER	EGFR CRP	C9 ITIH4	0.854
41	KLK3-SERPINA3 FN1	KIT BMPER	EFNA5 ALB	EGFR CRP	C9 ITIH4	0.854
42	KLK3-SERPINA3 BMP1	BDNF FN1	KIT SERPINA4	EFNA5 CRP	EGFR ITIH4	0.854
43	KLK3-SERPINA3 FN1	BDNF BMPER	KIT CRP	EFNA5 AFM	EGFR ITIH4	0.854
44	KLK3-SERPINA3 C9	KIT FN1	VEGFA SERPINA4	EFNA5 ALB	EGFR ITIH4	0.854
45	KLK3-SERPINA3 BMP1	KIT SERPINA4	EFNA5 DDC	EGFR BMPER	C9 ITIH4	0.854
46	KLK3-SERPINA3 C9	BDNF FN1	KIT SERPINA4	EFNA5 AFM	EGFR ITIH4	0.854
47	KLK3-SERPINA3 FN1	BDNF SERPINA4	KIT ALB	EFNA5 AFM	EGFR ITIH4	0.854

(continued)

Markers						Mean CV AUC
48	KLK3-SERPINA3 FN1	KIT DDC	EFNA5 ALB	EGFR CRP	C9 ITIH4	0.854
49	KLK3-SERPINA3 C9	BDNF SERPINA4	KIT DDC	EFNA5 CRP	EGFR ITIH4	0.854
50	KLK3-SERPINA3 C9	BDNF FN1	KIT CRP	EFNA5 AFM	EGFR ITIH4	0.854
51	KLK3-SERPINA3 C9	BDNF CDK8-CCNC	KIT FN1	EFNA5 CRP	EGFR ITIH4	0.854
52	KLK3-SERPINA3 DDC	KIT BMPER	EFNA5 CRP	EGFR AFM	FN1 ITIH4	0.854
53	KLK3-SERPINA3 C9	BDNF FN1	KIT SERPINA4	EFNA5 BMPER	EGFR ALB	0.854
54	KLK3-SERPINA3 SERPINA4	KIT BMPER	EFNA5 ALB	EGFR AFM	FN1 ITIH4	0.854
55	KLK3-SERPINA3 FN1	KIT SERPINA4	EFNA5 ALB	EGFR AFM	C9 ITIH4	0.854
56	KLK3-SERPINA3 FN1	KIT DDC	EFNA5 BMPER	EGFR ALB	C9 ITIH4	0.854
57	KLK3-SERPINA3 BMP1	KIT FN1	EFNA5 BMPER	EGFR CRP	C9 ITIH4	0.854
58	KLK3-SERPINA3 DDC	KIT BMPER	EFNA5 CRP	EGFR AFM	C9 ITIH4	0.854
59	KLK3-SERPINA3 C9	BDNF FN1	KIT DDC	EFNA5 CRP	EGFR ITIH4	0.854
60	KLK3-SERPINA3 FN1	KIT SERPINA4	EFNA5 BMPER	EGFR CRP	BMP1 ITIH4	0.854
61	KLK3-SERPINA3 C9	BDNF FN1	KIT SERPINA4	EFNA5 BMPER	EGFR CRP	0.854
62	KLK3-SERPINA3 BMP1	KIT FN1	EFNA5 DDC	EGFR ALB	C9 ITIH4	0.854
63	KLK3-SERPINA3 C9	BDNF SERPINA4	KIT ALB	EFNA5 AFM	EGFR ITIH4	0.854
64	KLK3-SERPINA3 BMP1	BDNF FN1	KIT SERPINA4	EFNA5 ALB	EGFR ITIH4	0.854
65	KLK3-SERPINA3 C9	BDNF FN1	KIT BMPER	EFNA5 AFM	EGFR ITIH4	0.854
66	KLK3-SERPINA3 C9	BDNF BMP1	KIT FN1	EFNA5 ALB	EGFR ITIH4	0.854
67	KLK3-SERPINA3 FN1	KIT BMPER	EFNA5 CRP	EGFR AFM	BMP1 ITIH4	0.854
68	KLK3-SERPINA3 FN1	KIT DDC	EFNA5 CRP	EGFR AFM	BMP1 ITIH4	0.854

(continued)

Markers						Mean CV AUC
69	KLK3-SERPINA3 FN1	KIT BMPER	EFNA5 ALB	EGFR AFM	C9 ITIH4	0.854
70	KLK3-SERPINA3 FN1	KIT SERPINA4	EFNA5 BMPER	EGFR ALB	BMP1 ITIH4	0.854
71	KLK3-SERPINA3 EGFR	BDNF C9	KIT FN1	VEGFA SERPINA4	EFNA5 ALB	0.854
72	KLK3-SERPINA3 FN1	BDNF SERPINA4	KIT ALB	EFNA5 CRP	EGFR ITIH4	0.854
73	KLK3-SERPINA3 BMP1	KIT FN1	EFNA5 ALB	EGFR AFM	C9 ITIH4	0.854
74	KLK3-SERPINA3 C9	BDNF ALB	KIT CRP	EFNA5 AFM	EGFR ITIH4	0.854
75	KLK3-SERPINA3 FN1	KIT SERPINA4	EFNA5 ALB	EGFR AFM	BMP1 ITIH4	0.854
76	KLK3-SERPINA3 FN1	KIT DDC	EFNA5 BMPER	EGFR AFM	BMP1 ITIH4	0.854
77	KLK3-SERPINA3 BMP1	KIT FN1	EFNA5 BMPER	EGFR ALB	C9 ITIH4	0.854
78	KLK3-SERPINA3 BMP1	BDNF FN1	KIT CRP	EFNA5 AFM	EGFR ITIH4	0.854
79	KLK3-SERPINA3 FN1	BDNF SERPINA4	KIT BMPER	EFNA5 AFM	EGFR ITIH4	0.854
80	KLK3-SERPINA3 BMP1	KIT FN1	VEGFA ALB	EFNA5 CRP	EGFR ITIH4	0.854
81	KLK3-SERPINA3 C9	BDNF DDC	KIT BMPER	EFNA5 AFM	EGFR ITIH4	0.854
82	KLK3-SERPINA3 SERPINA4	KIT DDC	EFNA5 BMPER	EGFR AFM	FN1 ITIH4	0.854
83	KLK3-SERPINA3 CDK5-CDK5R1	KIT FN1	EFNA5 DDC	EGFR BMPER	C9 ITIH4	0.854
84	KLK3-SERPINA3 BMP1	KIT FN1	EFNA5 SERPINA4	EGFR BMPER	C9 ITIH4	0.854
85	KLK3-SERPINA3 BMP1	KIT FN1	EFNA5 SERPINA4	EGFR BMPER	C9 ALB	0.854
86	KLK3-SERPINA3 FN1	KIT SERPINA4	EFNA5 DDC	EGFR BMPER	BMP1 ITIH4	0.853
87	KLK3-SERPINA3 CDK5-CDK5R1	BDNF C9	KIT FN1	EFNA5 SERPINA4	EGFR CRP	0.853
88	KLK3-SERPINA3 FN1	BDNF SERPINA4	KIT BMPER	EFNA5 CRP	EGFR ITIH4	0.853
89	KLK3-SERPINA3 C9	BDNF FN1	KIT DDC	EFNA5 ALB	EGFR ITIH4	0.853

(continued)

	Markers					Mean CV AUC
90	KLK3-SERPINA3 DDC	KIT BMPER	EFNA5 CRP	EGFR AFM	BMP1 ITIH4	0.853
91	KLK3-SERPINA3 CDK5-CDK5R1	BDNF C9	KIT FN1	VEGFA CRP	EFNA5 ITIH4	0.853
92	KLK3-SERPINA3 BMP1	KIT FN1	EFNA5 SERPINA4	EGFR CRP	C9 ITIH4	0.853
93	KLK3-SERPINA3 CDK5-CDK5R1	KIT FN1	EFNA5 SERPINA4	EGFR BMPER	C9 ITIH4	0.853
94	KLK3-SERPINA3 FN1	KIT SERPINA4	EFNA5 BMPER	EGFR ALB	C9 CRP	0.853
95	KLK3-SERPINA3 CDK5-CDK5R1	KIT FN1	EFNA5 CRP	EGFR AFM	C9 ITIH4	0.853
96	KLK3-SERPINA3 BMP1	KIT FN1	EFNA5 DDC	EGFR BMPER	C9 ITIH4	0.853
97	KLK3-SERPINA3 CDK5-CDK5R1	KIT DDC	EFNA5 BMPER	EGFR CRP	C9 ITIH4	0.853
98	KLK3-SERPINA3 C9	BDNF SERPINA4	KIT BMPER	EFNA5 CRP	EGFR ITIH4	0.853
99	KLK3-SERPINA3 C9	BDNF BMP1	KIT ALB	EFNA5 CRP	EGFR AFM	0.853
100	KLK3-SERPINA3 CDK5-CDK5R1	KIT FN1	EFNA5 DDC	EGFR ALB	C9 ITIH4	0.853

Table 30: Counts of markers in biomarker panels

	Panel Size							
Biomarker	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
AFM	149	146	123	138	142	197	262	354
ALB	129	120	129	148	194	258	332	405
APOA1	99	28	12	4	1	1	1	2
BDNF	169	326	447	480	476	478	491	513
BMP1	149	177	205	241	287	318	359	404
BMPER	160	227	236	260	280	311	357	409
C9	199	365	421	475	539	586	648	705
CCL23	98	65	39	33	21	23	14	19
CDK5-CDK5R1	72	58	73	84	129	153	182	223
CDK8-CCNC	98	52	52	61	71	68	87	98
CFHR5	69	6	1	1	0	0	0	0
CRP	181	254	292	342	403	478	549	599
DDC	142	189	192	217	231	274	323	374
EFNA5	157	277	416	566	727	859	931	958

(continued)

	Panel Size							
Biomarker	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
EGFR	171	324	413	496	582	651	744	824
FGA-FGB-FGG	40	0	0	0	0	0	0	0
FN1	130	161	220	289	383	520	619	722
ITIH4	144	155	203	267	369	469	571	691
KIT	166	315	575	769	895	948	985	993
KLK3-SERPINA3	201	490	681	821	897	951	970	980
SERPINA4	162	179	201	226	278	325	407	498
VEGFA	115	86	69	82	95	132	168	229

Table 31: Parameters derived from cancer datasets set for naïve Bayes classifiers

		Mesothelioma		NSCLC		Renal Cell Carc.	
		Control	**Cancer**	**Control**	**Cancer**	**Control**	**Cancer**
ALB	Mean	9.49	9.28	9.76	9.64	9.60	9.37
	SD	0.12	0.30	0.13	0.17	0.13	0.31
BMP1	Mean	8.62	8.30	8.77	8.55	8.72	8.51
	SD	0.27	0.35	0.21	0.23	0.25	0.34
C9	Mean	11.52	11.96	11.72	11.94	11.78	12.10
	SD	0.20	0.29	0.19	0.22	0.23	0.28
EFNA5	Mean	6.70	6.83	6.91	6.99	6.88	7.01
	SD	0.11	0.25	0.11	0.15	0.14	0.20
EGFR	Mean	10.46	10.26	10.58	10.43	10.52	10.38
	SD	0.11	0.21	0.12	0.13	0.14	0.12
FN1	Mean	8.92	8.53	9.29	9.06	9.10	8.94
	SD	0.36	0.38	0.24	0.32	0.19	0.32
ITIH4	Mean	10.18	10.46	10.60	10.74	10.56	10.82
	SD	0.32	0.34	0.12	0.23	0.15	0.20
KIT	Mean	9.39	9.18	9.60	9.50	9.39	9.25
	SD	0.16	0.20	0.14	0.14	0.16	0.19
KLK3-SERPINA3	Mean	8.00	8.51	8.10	8.33	8.09	8.68
	SD	0.16	0.53	0.19	0.33	0.23	0.48
SERPINA4	Mean	10.71	10.40	10.88	10.75	10.78	10.38
	SD	0.13	0.43	0.14	0.22	0.18	0.47

Table 32: Calculations derived from training set for naïve Bayes classifier.

Biomarker	μ_c	μ_d	σ_c	σ_d	$\tilde{x}$	$p(c\|\tilde{x})$	$p(d\|\tilde{x})$	$ln(p(d\|\tilde{x})/p(d\|\tilde{x}))$
EFNA5	6.907	6.994	0.107	0.148	6.974	3.059	2.663	-0.139
KIT	9.603	9.503	0.139	0.141	9.534	2.546	2.767	0.083
FN1	9.286	9.058	0.239	0.325	9.266	1.665	1.000	-0.510

(continued)

Biomarker	μ_c	μ_d	σ_c	σ_d	$\tilde{x}$	$p(c\|\tilde{x})$	$p(d\|\tilde{x})$	$ln(p(d\|\tilde{x})/p(d\|\tilde{x}))$
EGFR	10.578	10.428	0.119	0.135	10.547	3.236	2.003	-0.480
C9	11.715	11.936	0.189	0.223	11.715	2.114	1.096	-0.657
ALB	9.763	9.640	0.130	0.166	9.617	1.636	2.381	0.375
SERPINA4	10.881	10.745	0.144	0.223	10.905	2.728	1.384	-0.679
KLK3-SERPINA3	8.102	8.327	0.194	0.330	7.909	1.253	0.542	-0.838
BMP1	8.766	8.548	0.213	0.234	8.725	1.837	1.282	-0.360
ITIH4	10.596	10.738	0.121	0.227	10.600	3.301	1.460	-0.816

[0358] The following numbered paragraphs contain statements of the broad combinations of the inventive technical features herein disclosed:

1. A method for diagnosing that an individual does or does not have mesothelioma, the method comprising:

    detecting, in a biological sample from an individual, biomarker values that each correspond to one of at least N biomarkers selected from Table 1, wherein said individual is classified as having or not having mesothelioma, or the likelihood of the individual having mesothelioma is determined, based on said biomarker values, and wherein N =2 - 66.

2. The method of paragraph 1, wherein the diagnosis comprises the differential diagnosis of mesothelioma from benign conditions found in asbestos exposed individuals.

3. The method of paragraph 1 wherein the individual has a pleural abnormality.

4. The method of paragraph 1, wherein detecting the biomarker values comprises performing an in vitro assay.

5. The method of paragraph 4, wherein said in vitro assay comprises at least one capture reagent corresponding to each of said biomarkers, and further comprising selecting said at least one capture reagent from the group consisting of aptamers, antibodies, and a nucleic acid probe.

6. The method of paragraph 5, wherein said at least one capture reagent is an aptamer.

7. The method of paragraph 1, wherein the biological sample is selected from the group consisting of whole blood, plasma, serum and pleural fluid.

8. The method of paragraph 4, wherein the in vitro assay is selected from the group consisting of an immunoassay, an aptamer-based assay, a histological or cytological assay, and an mRNA expression level assay.

9. The method of paragraph 8, wherein the biological sample is serum.

10. The method paragraph 1, wherein the biological sample is pleural or peritoneal mesothelium tissue and wherein the biomarker values derive from a histological or cytological analysis of said mesothelium tissue.

11. The method of paragraph 1, wherein the individual is a human.

12. The method of paragraph 1, wherein N = 3 - 10.

13. The method of paragraph 1, wherein N = 3 - 15.

14. The method of paragraph 1, wherein N = 2 - 10.

15. The method of paragraph 1, wherein N = 4 - 10.

16. The method of paragraph 1, wherein N = 5 - 10.

17. The method of paragraph 1, wherein the individual is high risk for mesothelioma due to asbestos or related fiber exposure.

18. The method of paragraph 1, wherein the biomarkers are selected from Table 18.

19. A computer-implemented method for indicating a likelihood of mesothelioma, the method comprising:

retrieving on a computer biomarker information for an individual, wherein the biomarker information comprises biomarker values that each correspond to one of at least N biomarkers selected from Table 1;
performing with the computer a classification of each of said biomarker values; and indicating a likelihood that said individual has mesothelioma based upon a plurality of classifications, and wherein N = 2 - 66.

20. The method of paragraph 19, wherein indicating the likelihood that the individual has mesothelioma comprises displaying the likelihood on a computer display.

21.A computer program product for indicating a likelihood of mesothelioma, the computer program product comprising:

a computer readable medium embodying program code executable by a processor of a computing device or system, the program code comprising:

code that retrieves data attributed to a biological sample from an individual, wherein the data comprises biomarker values that each correspond to one of at least N biomarkers selected from Table 1, wherein said biomarkers were detected in the biological sample; and code that executes a classification method that indicates a mesothelioma status of the individual as a function of said biomarker values; and wherein N = 2 -66.

22. The computer program product of paragraph 21, wherein said classification method uses a probability density function.

23. The computer program product of paragraph 22, wherein said classification method uses two or more classes.

24. A method for screening an asymptomatic high risk individual for mesothelioma, the method comprising:

detecting, in a biological sample from an individual, biomarker values that each correspond to one of at least N biomarkers selected from Table 1, wherein said individual is classified as having or not having mesothelioma, or the likelihood of the individual having mesothelioma is determined, based on said biomarker values, and wherein N =2 - 66.

25. The method of paragraph 24, wherein detecting the biomarker values comprises performing an in vitro assay.

26. The method of paragraph 25, wherein said in vitro assay comprises at least one capture reagent corresponding to each of said biomarkers, and further comprising selecting said at least one capture reagent from the group consisting of aptamers, antibodies, and a nucleic acid probe.

27. The method of paragraph 26, wherein said at least one capture reagent is an aptamer.

28. The method of paragraph 27, wherein the in vitro assay is selected from the group consisting of an immunoassay, an aptamer-based assay, a histological or cytological assay, and an mRNA expression level assay.

29. The method of paragraph 24, wherein the biological sample is selected from the group consisting of whole blood, plasma, serum and pleural fluid.

30. The method of paragraph 29, wherein the biological sample is serum.

31. The method paragraph 24, wherein the biological sample is mesothelium tissue and wherein the biomarker values derive from a histological or cytological analysis of said mesothelium tissue.

32. The method of paragraph 24, wherein the individual is a human.

33. The method of paragraph 24, wherein N = 3 - 10.

34. The method of paragraph 24, wherein N = 3 - 15.

35. The method of paragraph 24, wherein N = 2 - 10.

36. The method of paragraph 24, wherein N = 4 - 10.

37. The method of paragraph 24, wherein N = 5 - 10.

38. The method of paragraph 24, wherein the individual is high risk for mesothelioma due to exposure to asbestos or other related fibers.

39. The method of paragraph 24, wherein the biomarkers are selected from Table 18.

40. A computer-implemented method for indicating a likelihood of mesothelioma cancer, the method comprising:

retrieving on a computer biomarker information for an individual, wherein the biomarker information comprises biomarker values that each correspond to one of at least N biomarkers selected from Table 1;
performing with the computer a classification of each of said biomarker values; and indicating a likelihood that said individual has mesothelioma cancer based upon a plurality of classifications, and wherein N = 2 - 66.

41. The method of paragraph 40, wherein indicating the likelihood that the individual has mesothelioma comprises displaying the likelihood on a computer display.

42. A computer program product for indicating a likelihood of mesothelioma, the computer program product comprising:

a computer readable medium embodying program code executable by a processor of a computing device or system, the program code comprising:

code that retrieves data attributed to a biological sample from an individual, wherein the data comprises biomarker values that each correspond to one of at least N biomarkers selected from Table 1, wherein said biomarkers were detected in the biological sample; and
code that executes a classification method that indicates a mesothelioma status of the individual as a function of said biomarker values; and wherein N = 2 - 66.

43. The computer program product of paragraph 42, wherein said classification method uses a probability density function.

44. The computer program product of paragraph 43, wherein said classification method uses two or more classes.

45. A method for diagnosing that an individual does or does not have cancer, the method comprising:

detecting, in a biological sample from an individual, biomarker values that each correspond to one of at least N biomarkers selected from Table 19, wherein said individual is classified as having or not having cancer, or the likelihood of the individual having mesothelioma is determined, based on said biomarker values, and wherein N = 2 - 22.

46. The method of paragraph 45 wherein said diagnosed cancer comprises one or more of lung cancer, renal cell carcinoma, or mesothelioma.

47. The method of paragraph 45, wherein detecting the biomarker values comprises performing an in vitro assay.

48. The method of paragraph 47, wherein said in vitro assay comprises at least one capture reagent corresponding to each of said biomarkers, and further comprising selecting said at least one capture reagent from the group consisting of aptamers, antibodies, and a nucleic acid probe.

49. The method of paragraph 48, wherein said at least one capture reagent is an aptamer.

50. The method of paragraph 47, wherein the in vitro assay is selected from the group consisting of an immunoassay, an aptamer-based assay, a histological or cytological assay, and an mRNA expression level assay.

51. The method of paragraph 45, wherein the biological sample is selected from the group consisting of whole blood, plasma, and serum.

52. The method of paragraph 51, wherein the biological sample is serum.

53. The method of paragraph 45, wherein the individual is a human.

54. The method of paragraph 45, wherein N = 2 - 10.

55. The method of paragraph 45, wherein N = 2 - 15.

56. The method of paragraph 45, wherein N = 3 - 10.

57. The method of paragraph 45, wherein N = 3 - 15.

58. The method of paragraph 45, wherein N = 4 - 10.

59. A computer-implemented method for indicating a likelihood of cancer, the method comprising:

retrieving on a computer biomarker information for an individual, wherein the biomarker information comprises biomarker values that each correspond to one of at least N biomarkers selected from Table 19;
performing with the computer a classification of each of said biomarker values; and indicating a likelihood that said individual has cancer based upon a plurality of classifications, and wherein N = 2 - 22.

60. The method of paragraph 59, wherein indicating the likelihood that the individual has cancer comprises displaying the likelihood on a computer display.

61. A computer program product for indicating a likelihood of cancer, the computer program product comprising:

a computer readable medium embodying program code executable by a processor of a computing device or system, the program code comprising: code that retrieves data attributed to a biological sample from an individual, wherein the data comprises biomarker values that each correspond to one of at least N biomarkers selected from Table 19, wherein said biomarkers were detected in the biological sample; and code that executes a classification method that indicates a cancer status of the individual as a function of said biomarker values; and wherein N = 2 - 22.

62. The computer program product of paragraph 61, wherein said classification method uses a probability density function.

63. The computer program product of paragraph 62, wherein said classification method uses two or more classes.

64. The method according to paragraphs 1 or 25, wherein said individual is classified as having or not having mesothelioma, or the likelihood of the individual having mesothelioma is determined, based on said biomarker values and at least one item of additional biomedical information corresponding to said individual.

65. The method according to paragraph 45, wherein said individual is classified as having or not having cancer, or the likelihood of the individual having cancer is determined, based on said biomarker values and at least one item

of additional biomedical information corresponding to said individual.

66. The method of paragraph 64 or 65, wherein said at least one item of additional biomedical information is independently selected from the group consisting of

(a) information corresponding to the presence or absence of a pleural effusion or mass or other pleural mass,
(b) information corresponding to physical descriptors of said individual,
(c) information corresponding to a change in weight of said individual,
(d) information corresponding to the ethnicity of said individual,
(e) information corresponding to the gender of said individual,
(f) information corresponding to said individual's smoking history,
(g) information corresponding to said individual's asbestos exposure history,
(h) information corresponding to said individual's occupational history,
(i) information corresponding to said individual's family history of mesothelioma or other cancer,
(j) information corresponding to the presence or absence in said individual of at least one genetic marker correlating with a higher risk of mesothelioma or cancer insaid individual or a family member of said individual,
(k) information corresponding to clinical symptoms of said individual,
(l) information corresponding to other laboratory tests,
(m) information corresponding to gene expression values of said individual, and
(n) information corresponding to said individual's exposure to known carcinogens.

67. A classifier comprising the biomarkers of Table 18.

Claims

1. A method for diagnosing mesothelioma in an individual, the method comprising:

detecting biomarker proteins in a biological sample from the individual to give biomarker values that correspond to biomarker F9 and at least N additional biomarkers selected from Table 1, wherein said individual is classified as having or not having mesothelioma, or the likelihood of the individual having mesothelioma is determined, based on said biomarker values, and wherein N is 1.

2. The method of claim 1, wherein the biomarkers are selected from Table 18.

3. The method according to claim 1 or claim 2, wherein the biomarker is CHD1, CRK or MDK.

4. The method of claim 1, wherein N equals 2, 3, 4, 5, 6, 7, 8 or 9.

5. The method of any one of the preceding claims wherein the diagnosis comprises the differential diagnosis of mesothelioma from benign conditions found in asbestos exposed individuals.

6. The method of any one of the preceding claims wherein the individual has a pleural abnormality.

7. The method of any one of the preceding claims wherein detecting the biomarker values comprises performing an in vitro assay.

8. The method of claim 7, wherein said in vitro assay comprises at least one capture reagent corresponding to each of said biomarkers, wherein the capture reagent is selected from the group consisting of aptamers and antibodies.

9. The method of any one of the preceding claims wherein the biological sample is selected from the group consisting of serum, whole blood, plasma, and pleural fluid.

10. The method of claim 7, wherein the in vitro assay is selected from the group consisting of an immunoassay, an aptamer-based assay, a histological or cytological assay.

11. The method of any one of the preceding claims wherein the biological sample is pleural or peritoneal mesothelium tissue and wherein the biomarker values derive from a histological or cytological analysis of said mesothelium tissue.

12. The method of any one of the preceding claims wherein N = 3 - 10, 3-15, 2-10, 2-15, 4-10 or 5-10.

13. The method of any one of the preceding claims wherein the individual is high risk for mesothelioma due to asbestos or related fiber exposure.

14. A computer-implemented method for indicating a likelihood of mesothelioma, the method comprising:

retrieving on a computer biomarker information for an individual, wherein the biomarker information comprises biomarker values that correspond to biomarker F9 and at least N additional biomarkers selected from Table 1;
performing with the computer a classification of each of said biomarker values;
and indicating a likelihood that said individual has mesothelioma based upon a plurality of classifications, and wherein N = 1 - 66.

15. A computer program product for indicating a likelihood of mesothelioma, the computer program product comprising:

a computer readable medium embodying program code executable by a processor of a computing device or system, the program code comprising:

code that retrieves data attributed to a biological sample from an individual, wherein the data comprises biomarker values that correspond to biomarker F9 and at least N additional biomarkers selected from Table 1, wherein said biomarkers were detected in the biological sample; and
code that executes a classification method that indicates a mesothelioma status of the individual as a function of said biomarker values; and wherein N = 1 -66.

16. The computer program product of claim 15, wherein said classification method uses a probability density function or uses two or more classes.

17. The method according to any one of claims 1 to 14, wherein said individual is classified as having or not having mesothelioma, or the likelihood of the individual having mesothelioma is determined, based on said biomarker values and at least one item of additional biomedical information corresponding to said individual.

18. The method of claim 17, wherein said at least one item of additional biomedical information is independently selected from the group consisting of:

(a) information corresponding to the presence or absence of a pleural effusion or mass or other pleural mass,
(b) information corresponding to physical descriptors of said individual,
(c) information corresponding to a change in weight of said individual,
(d) information corresponding to the ethnicity of said individual,
(e) information corresponding to the gender of said individual,
(f) information corresponding to said individual's smoking history,
(g) information corresponding to said individual's asbestos exposure history,
(h) information corresponding to said individual's occupational history,
(i) information corresponding to said individual's family history of mesothelioma or other cancer,
(j) information corresponding to the presence or absence in said individual of at least one genetic marker correlating with a higher risk of mesothelioma or cancer in said individual or a family member of said individual,
(k) information corresponding to clinical symptoms of said individual,
(l) information corresponding to other laboratory tests,
(m) information corresponding to gene expression values of said individual, and;
(n) information corresponding to said individual's exposure to known carcinogens.

19. A classifier comprising the biomarkers of Table 18.

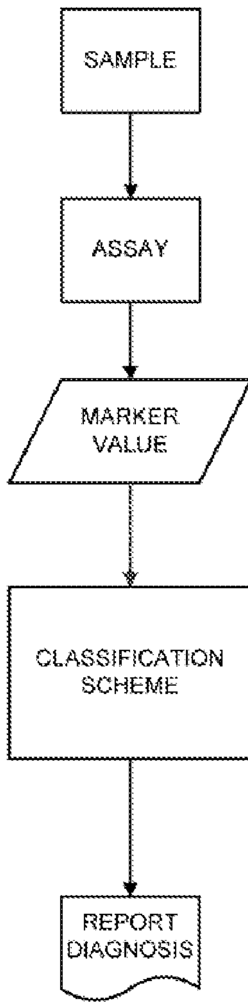


FIG. 1A

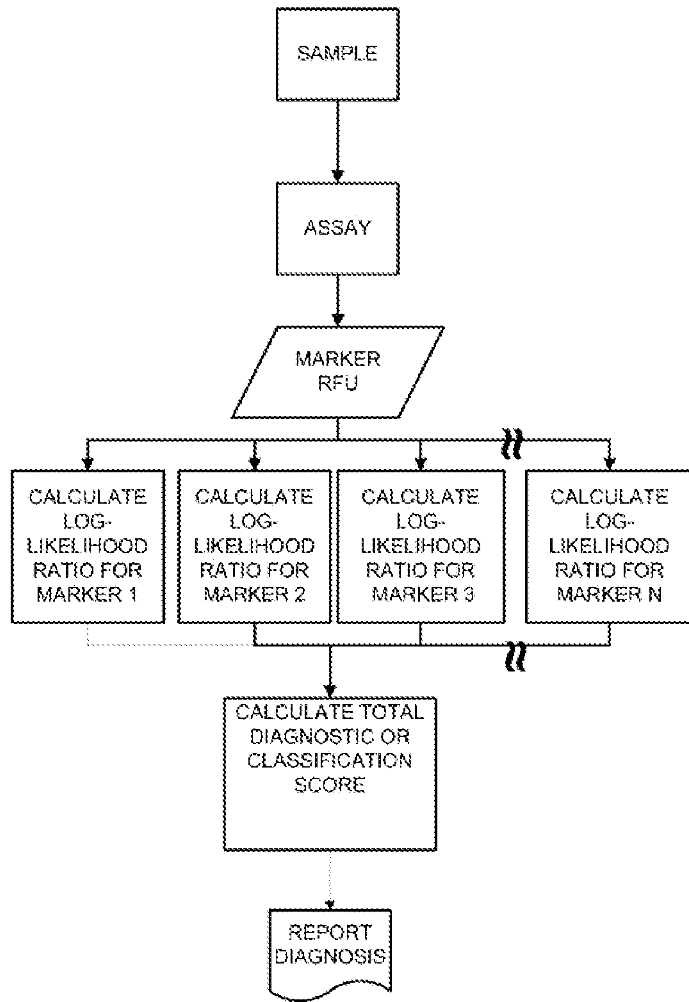


FIG. 1B

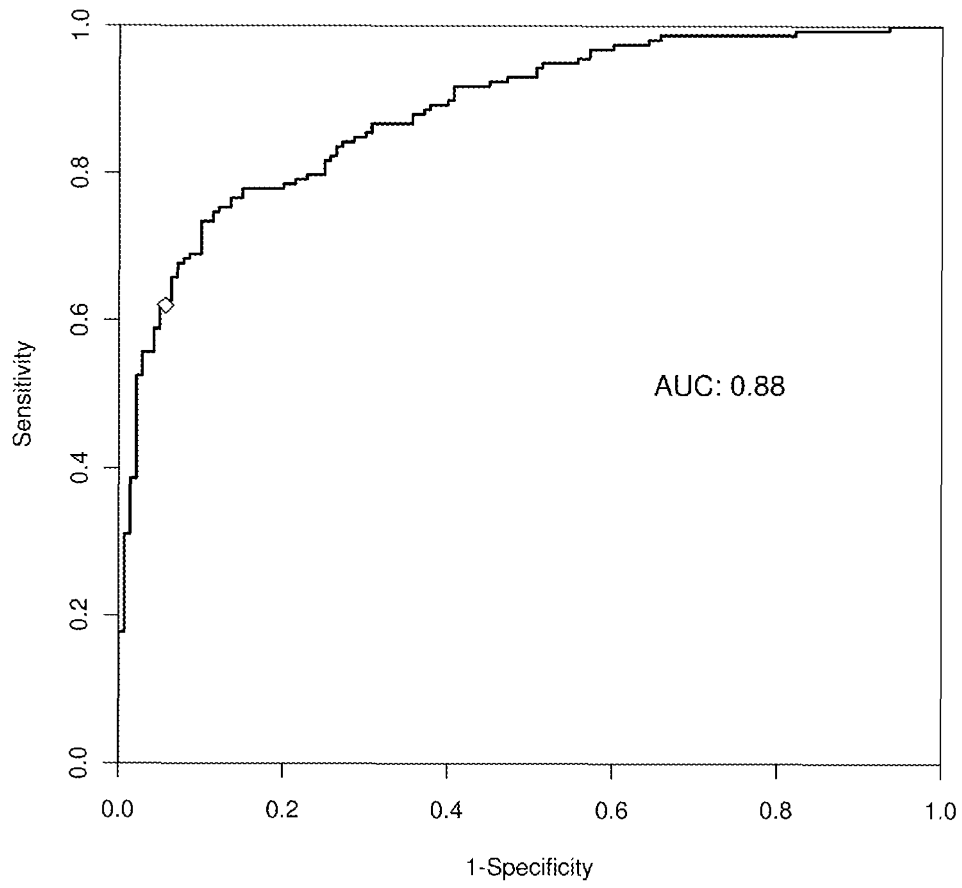


FIG. 2

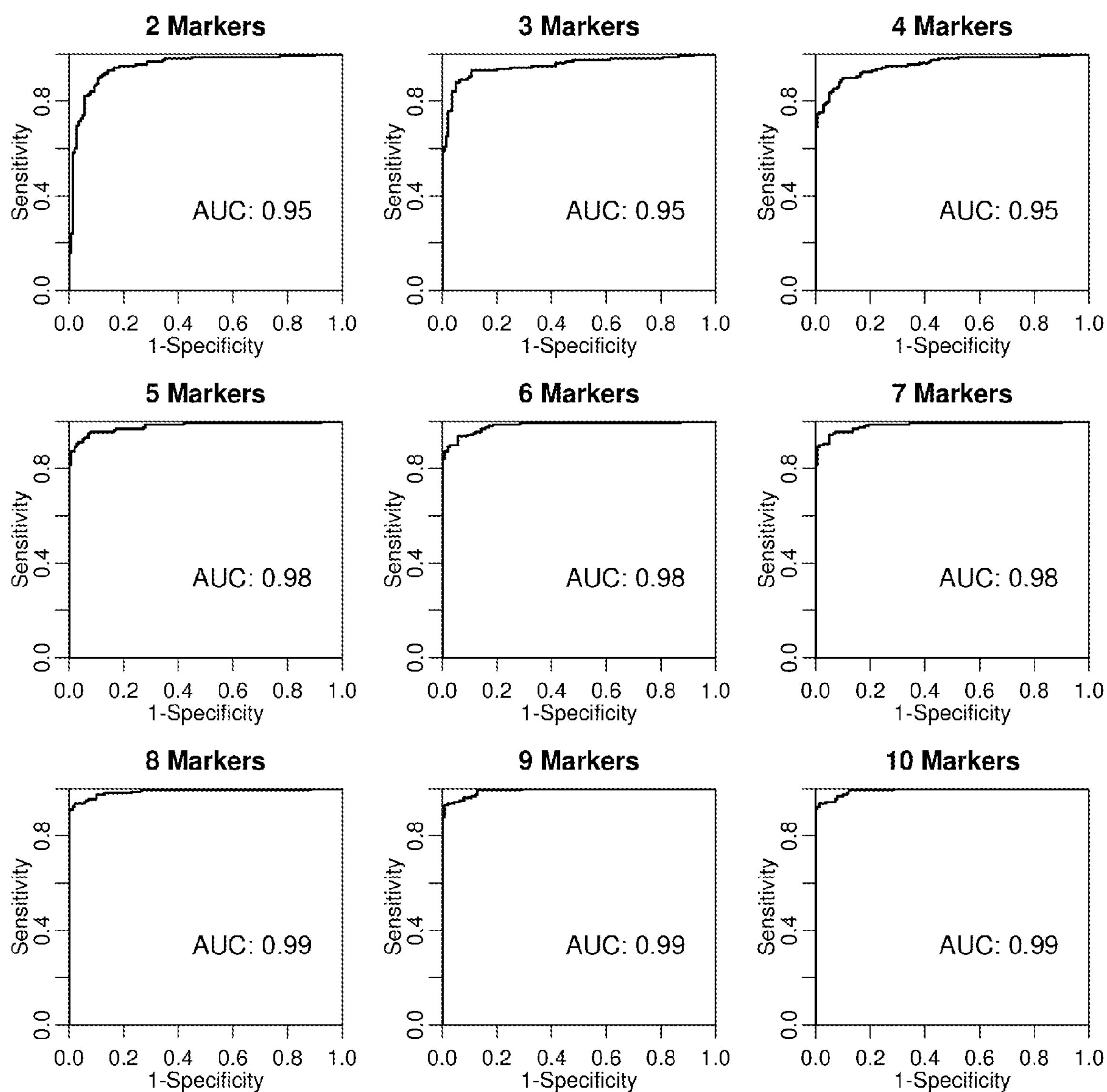


FIG. 3

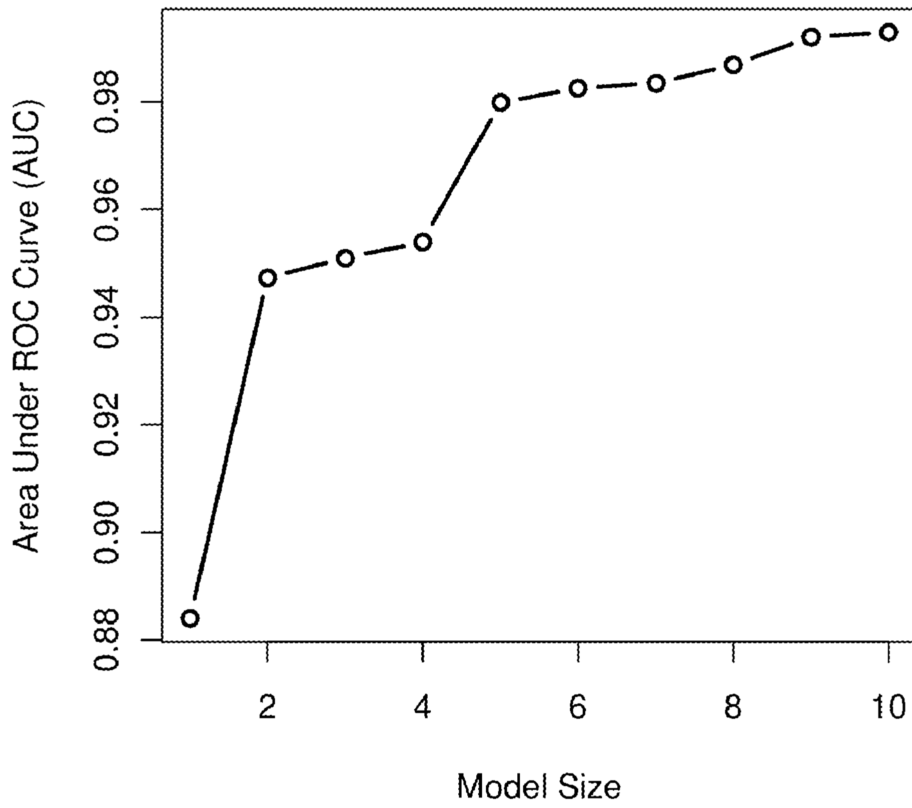


FIG. 4

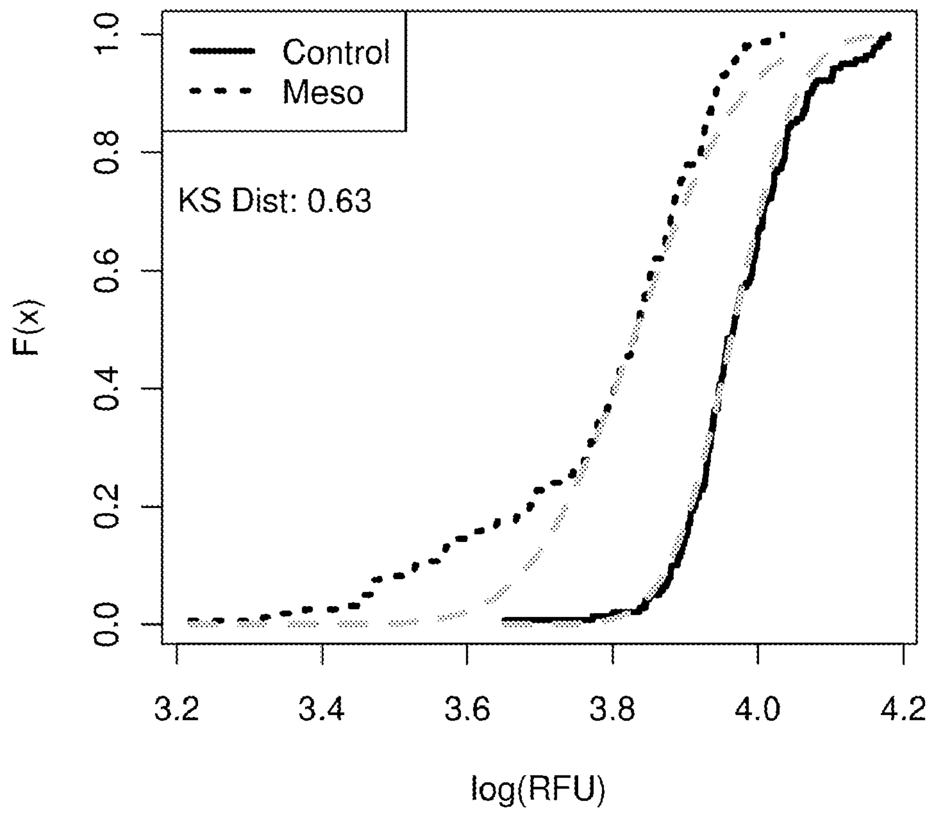


FIG. 5

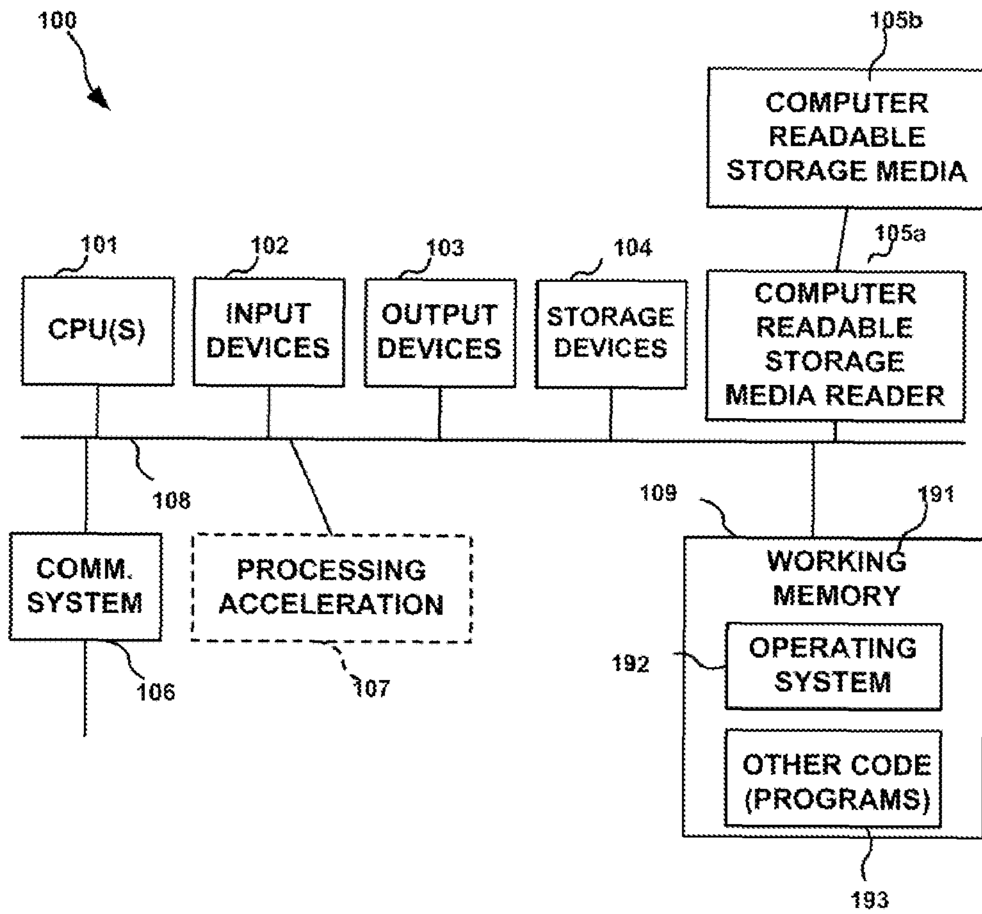


FIG. 6

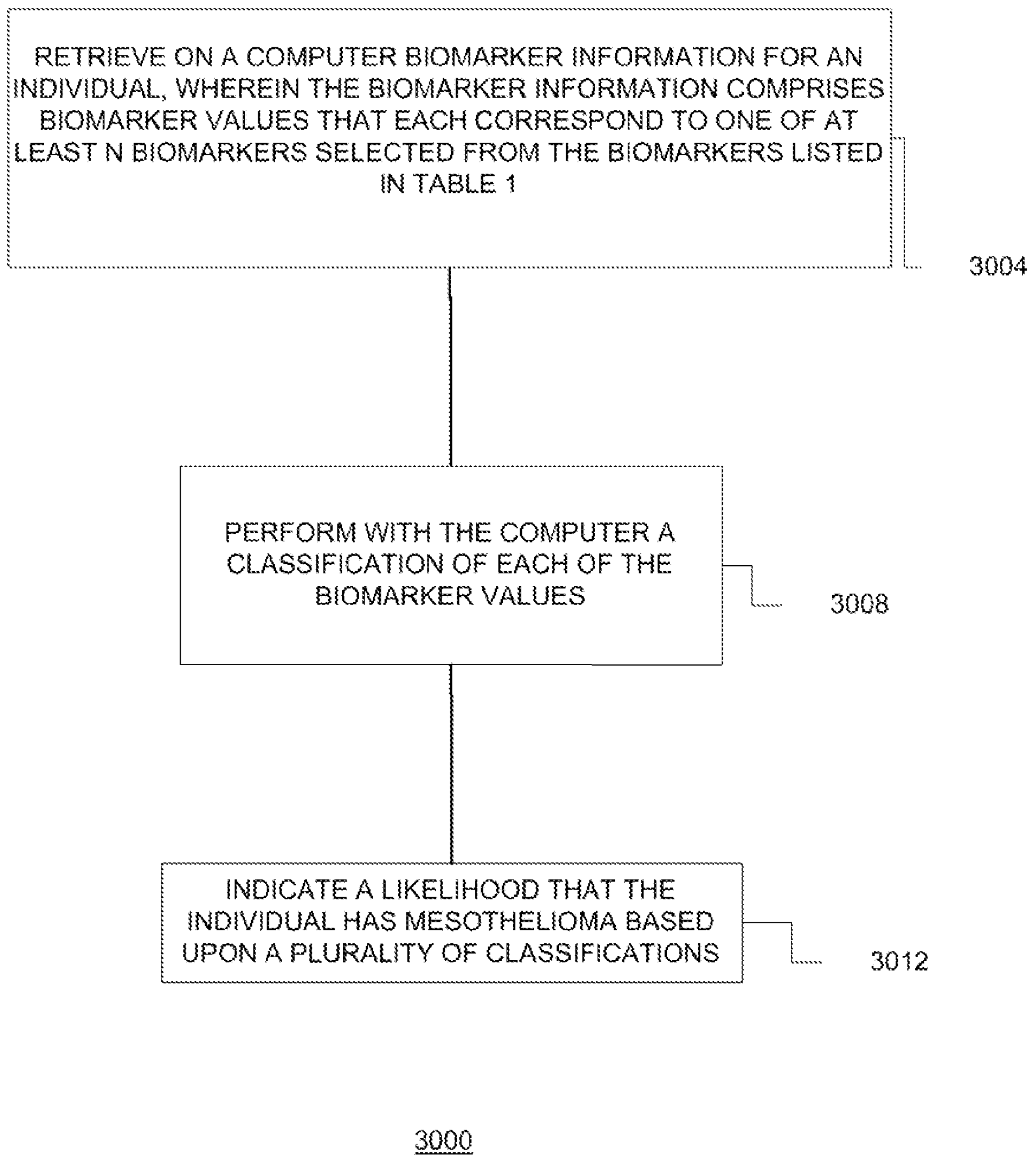


FIG. 7

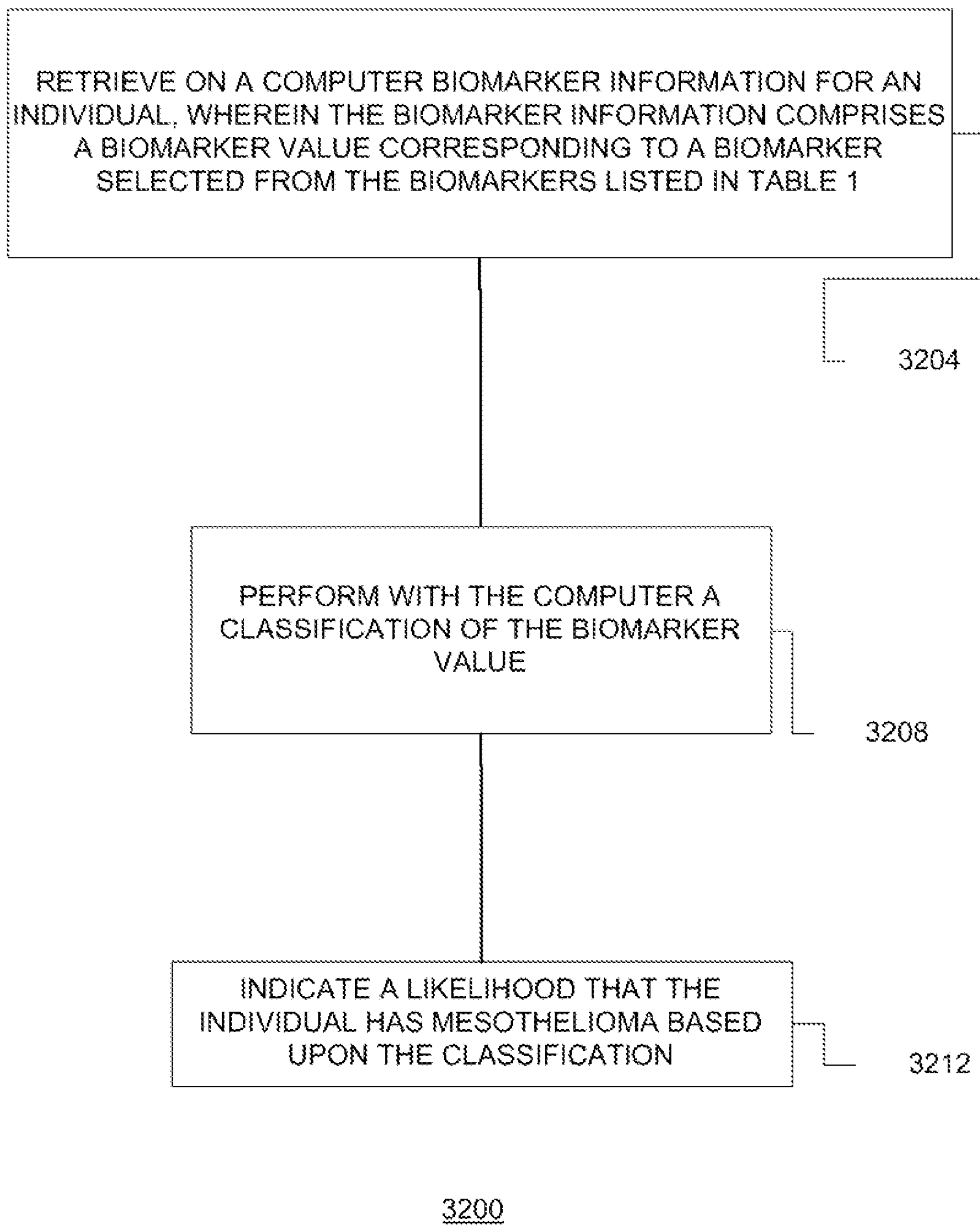


FIG. 8

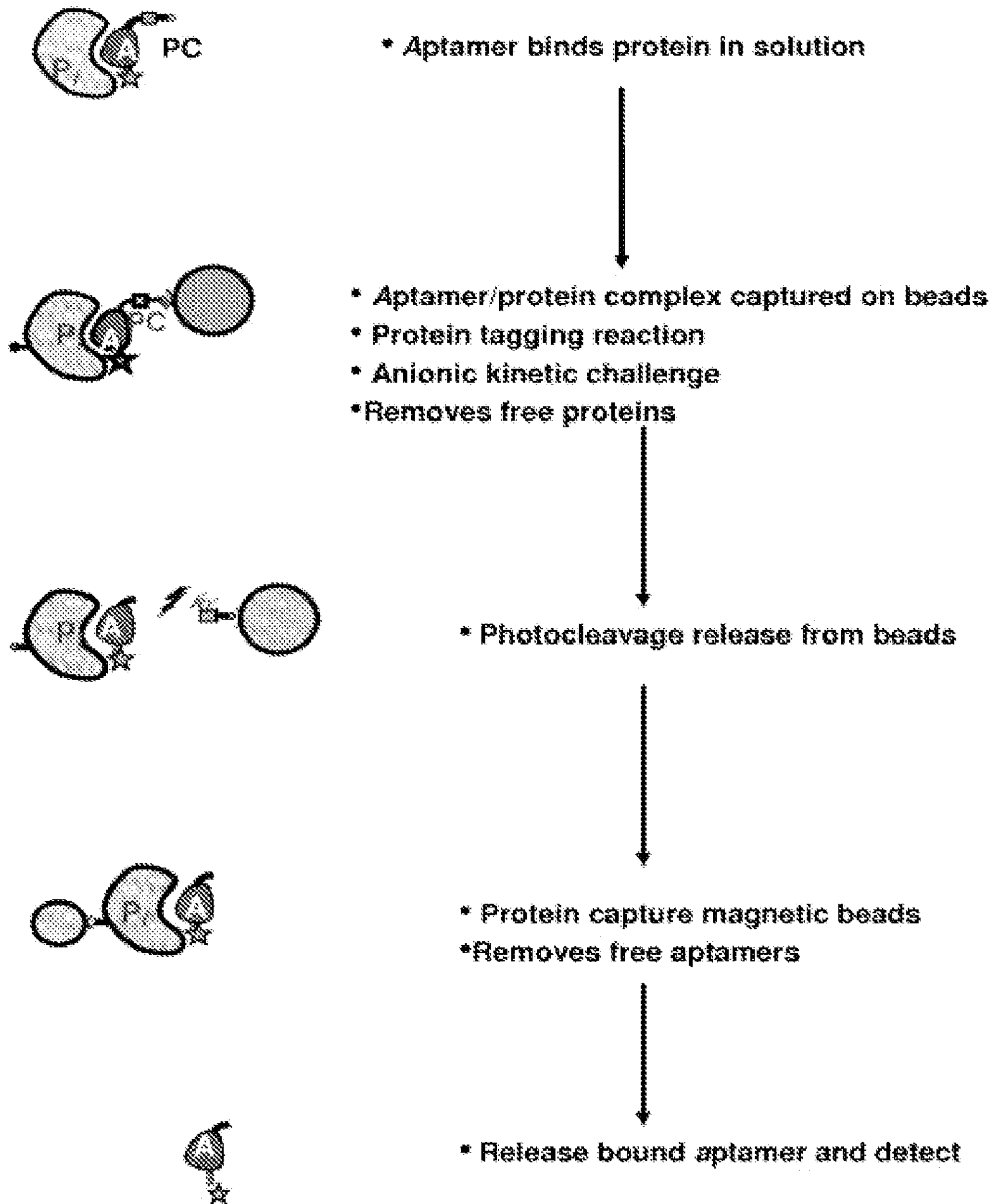


FIG. 9

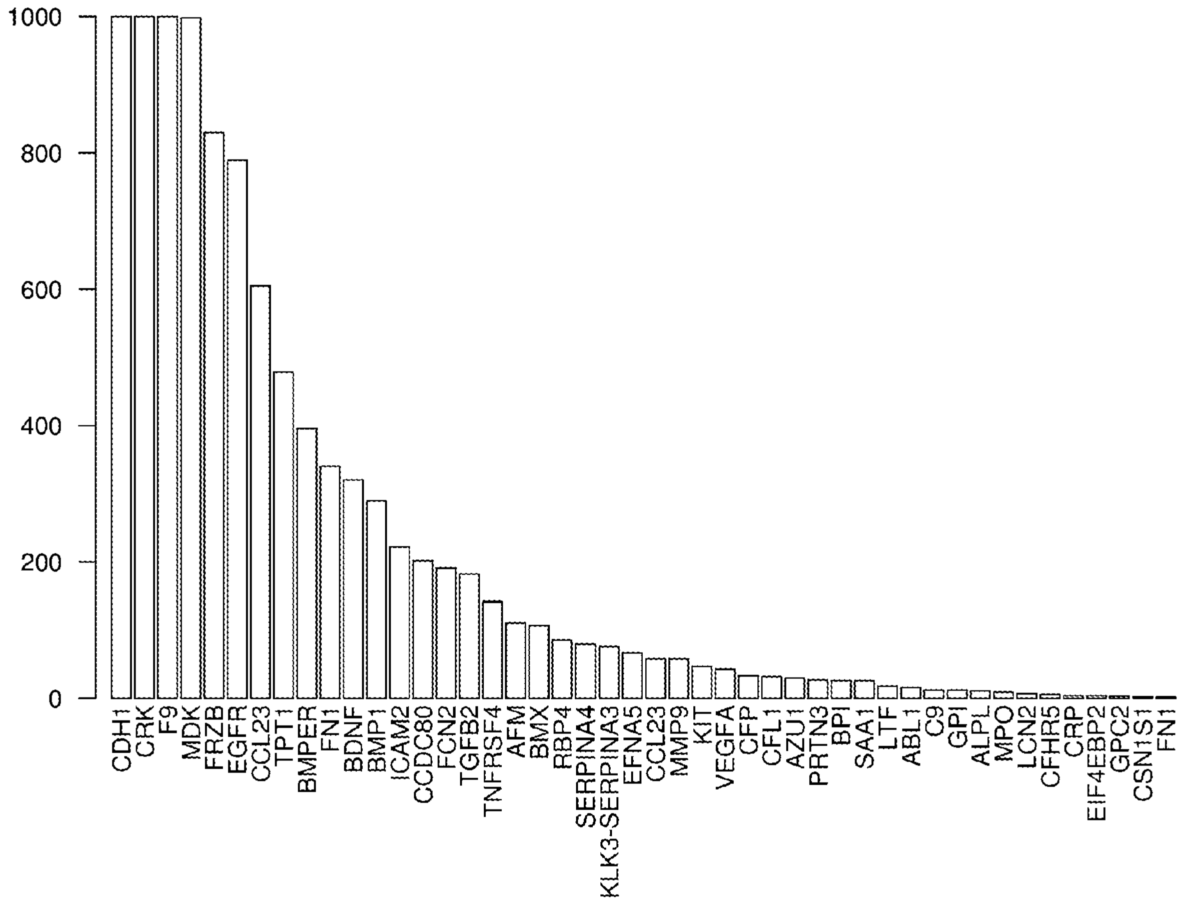


FIG. 10

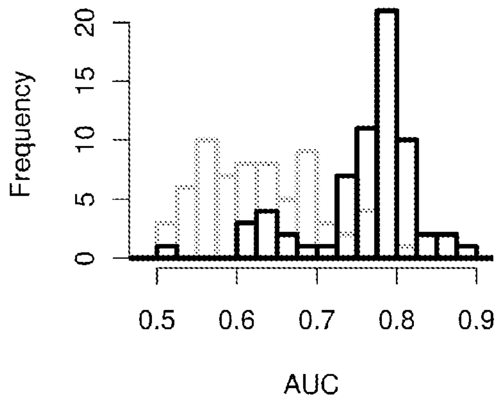


FIG. 11A

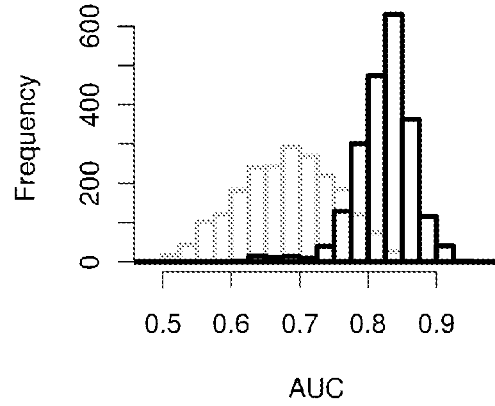


FIG. 11B

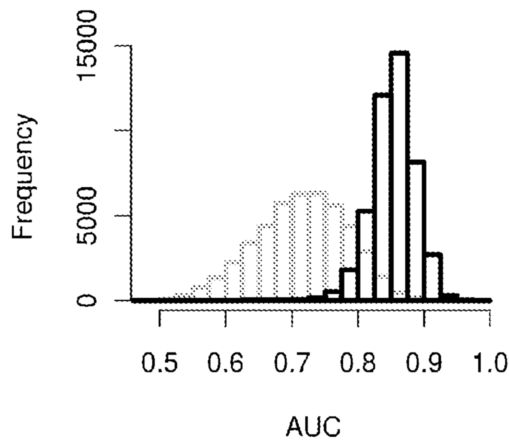


FIG. 11C

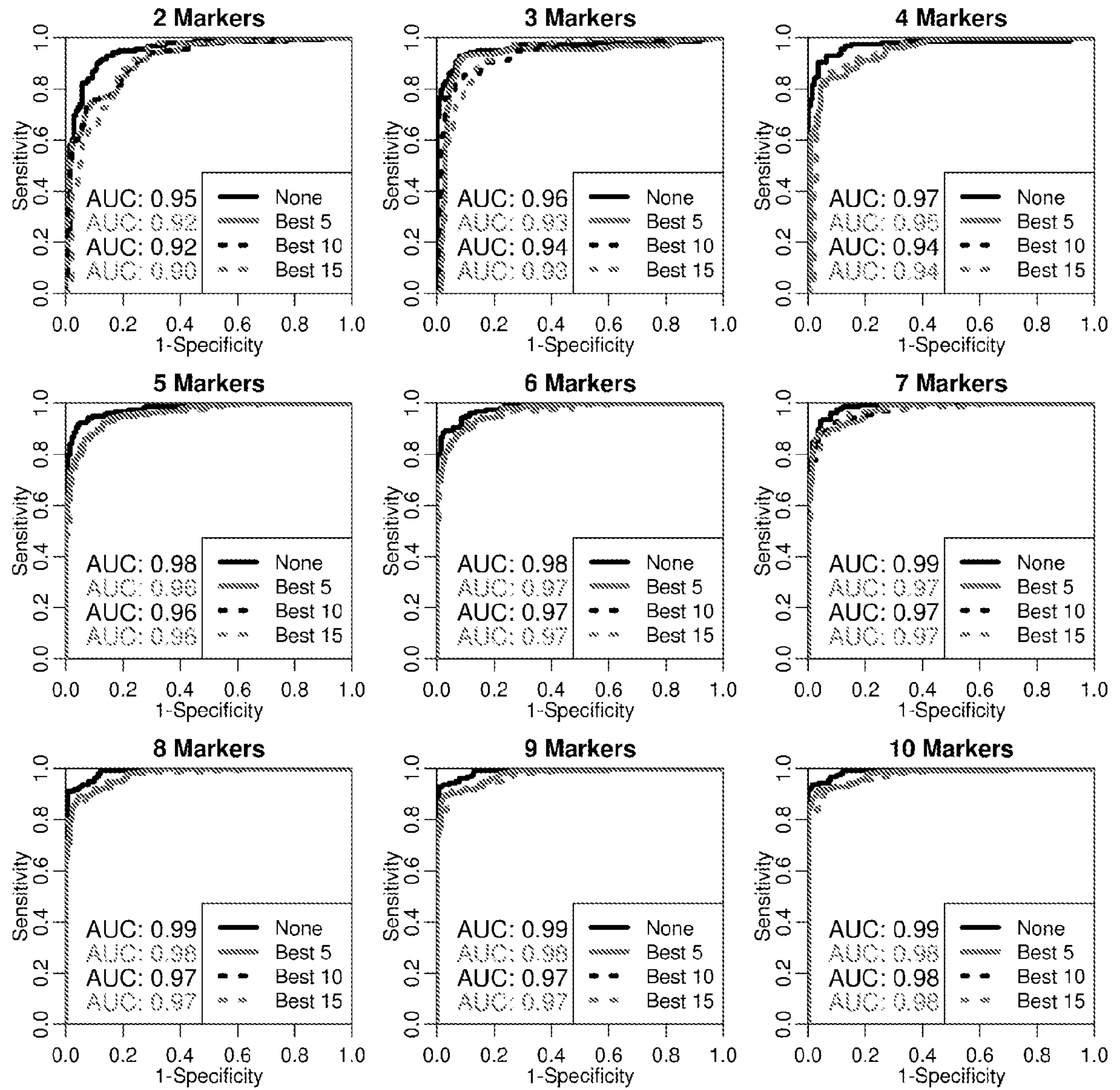


FIG. 12

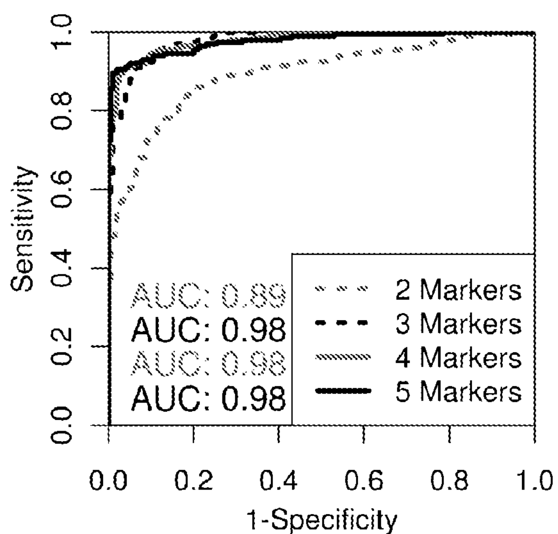


FIG. 13A

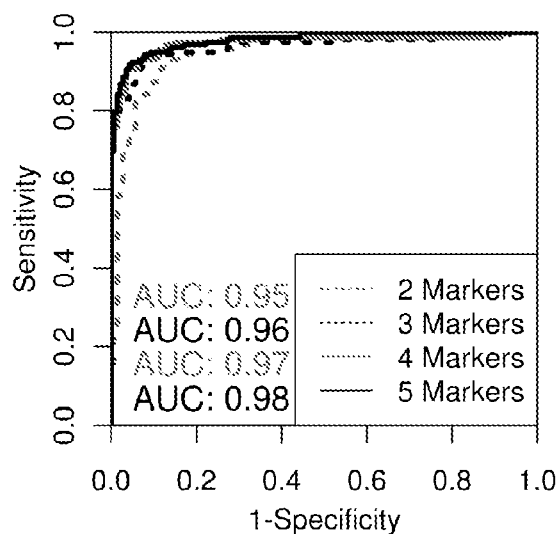


FIG. 13B

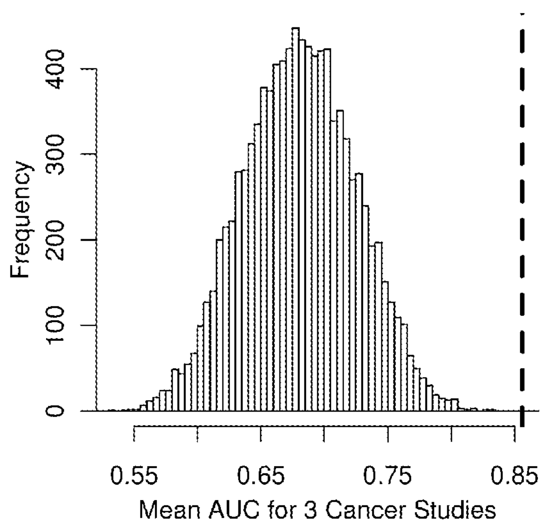


FIG. 14A

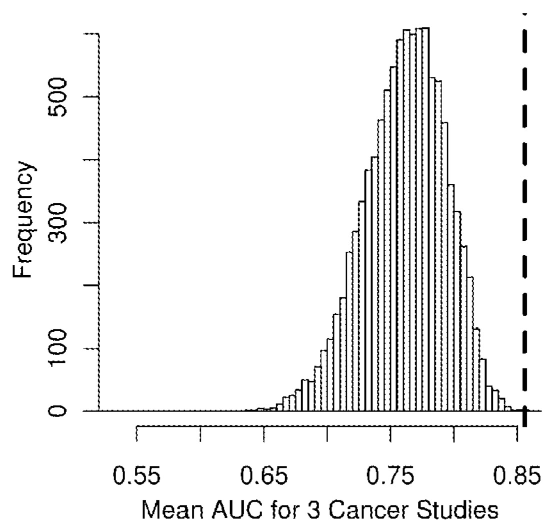


FIG. 14B

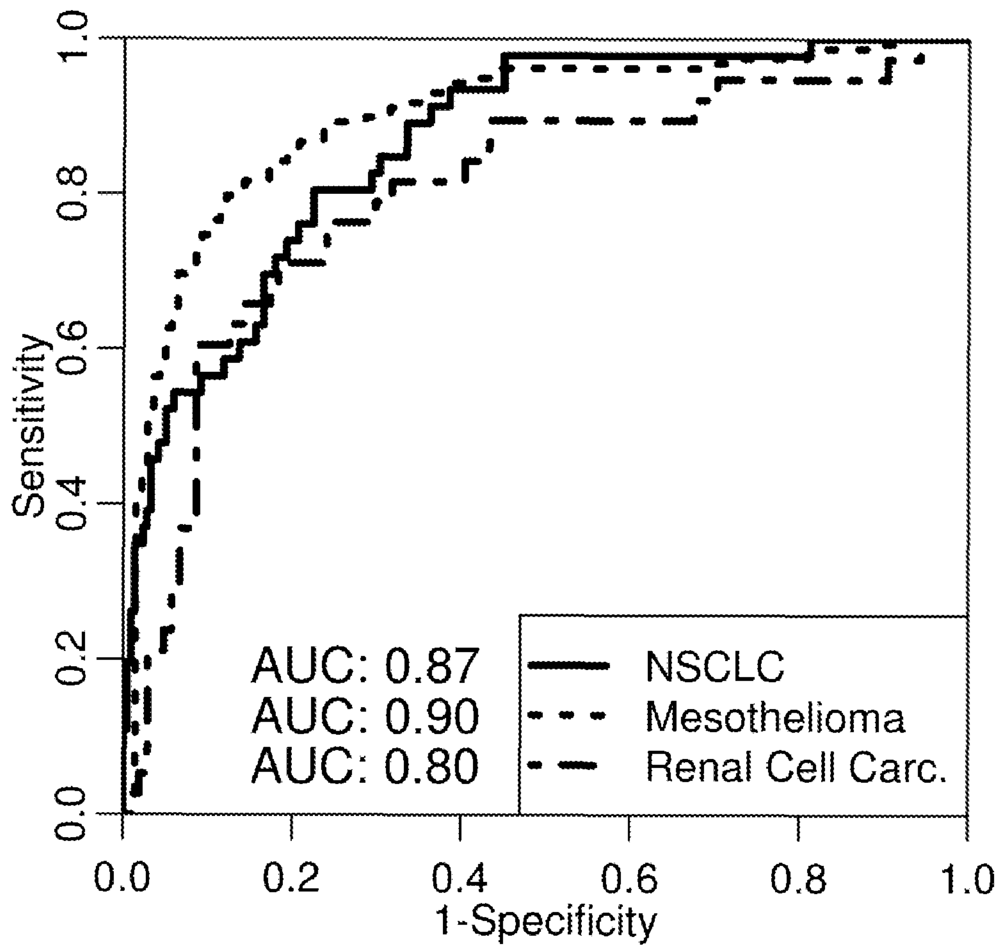


FIG. 15

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- US 61386840 A **[0001]**
- US 61470143 A **[0001]**
- US 5475096 A **[0137] [0142]**
- US 6242246 B **[0137]**
- US 6458543 B **[0137]**
- US 6503715 B **[0137]**
- US 20090004667 A **[0140] [0144]**
- US 5105331 A **[0142]**
- US 5660985 A **[0143]**
- US 5580737 A **[0143]**
- US 20090098549 A **[0143]**
- US 6544776 B **[0145]**
- US 5763177 A **[0145]**
- US 6001577 A **[0145]**
- US 6291184 B **[0145]**
- US 6458539 B **[0145]**
- US 20090042206 A **[0147]**

Non-patent literature cited in the description

- **AMOS et al.** *Nature Genetics,* 2009, vol. 40, 616-622 **[0121]**
- **J. R. LAKOWICZ.** Principles of Fluorescence Spectroscopy. Springer Science + Business Media, Inc, 2004 **[0132]**
- Bioluminescence & Chemiluminescence: Progress & Current Applications. World Scientific Publishing Company, 2002 **[0132]**
- ImmunoAssay: A Practical Guide. Taylor & Francis, Ltd, 2005 **[0158]**
- Gene Expression Profiling: Methods and Protocols. Humana Press, 2004 **[0163]**
- **N. BLOW.** *Nature Methods,* 2009, vol. 6, 465-469 **[0177]**
- **RICHARD G. KESSEL.** Basic Medical Histology: the Biology of Cells, Tissues and Organs. Oxford University Press, 1998 **[0207]**
- **BURLINGAME et al.** *Anal. Chem.,* 1998, vol. 70, 647 R-716R **[0208]**
- Pattern Classification. John Wiley & Sons, 2001 **[0216] [0217]**
- The Elements of Statistical Learning - Data Mining, Inference, and Prediction. LLC, 2009 **[0216] [0217]**
- **SAMBROOK et al.** Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, 2001 **[0268]**
- *PLoS ONE,* 2007, vol. 2 (10), 1002 **[0327]**
- **DABNEY AR ; STOREY JD.** Optimality Driven Nearest Centroid Classification from Genomic Data. *PLoS ONE,* 2007, vol. 2 (10), 1002 **[0351]**

patsnap

专利名称(译)	间皮瘤生物标志物及其用途		
公开(公告)号	EP3029153A2	公开(公告)日	2016-06-08
申请号	EP2015189622	申请日	2011-09-27
[标]申请(专利权)人(译)	私募蛋白质体公司		
申请(专利权)人(译)	SOMALOGIC INC.		
当前申请(专利权)人(译)	SOMALOGIC INC.		
[标]发明人	OSTROFF RACHEL M STEWART ALEX A E WILLIAMS STEPHEN ALARIC BRODY EDWARD N NIKRAD MALTI RIEL MEHAN MICHAEL		
发明人	OSTROFF, RACHEL M. STEWART, ALEX A. E. WILLIAMS, STEPHEN ALARIC BRODY, EDWARD N. NIKRAD, MALTI RIEL-MEHAN, MICHAEL		
IPC分类号	C12Q1/68 G01N33/53 G01N33/574		
优先权	61/386840 2010-09-27 US 61/470143 2011-03-31 US PCT/US2011/053377 2011-09-27 WO		
其他公开文献	EP3029153B1 EP3029153A3		
外部链接	Espacenet		

摘要(译)

本公开内容包括用于一般地检测和诊断癌症和特别是间皮瘤的生物标记物，方法，装置，试剂，系统和试剂盒。在一个方面，本公开内容提供了可以单独使用或以各种组合使用的生物标记物，以特异性地诊断癌症或间皮瘤。另一方面，提供了用于诊断个体间皮瘤的方法，其中所述方法包括在来自个体的生物样品中检测对应于选自表1中提供的生物标志物组的至少一种生物标志物的至少一种生物标志物值，其中，基于至少一种生物标志物值，将个体分类为具有间皮瘤，或确定个体患有间皮瘤的可能性。在另一方面，提供了通常在个体中诊断癌症的方法，其中所述方法包括在来自个体的生物样品中检测对应于选自表中提供的生物标志物组的至少一种生物标志物的至少一种生物标志物值。参照图19，其中基于至少一种生物标记值，将个体一般分类为患有癌症，或确定患有癌症的个体的可能性。

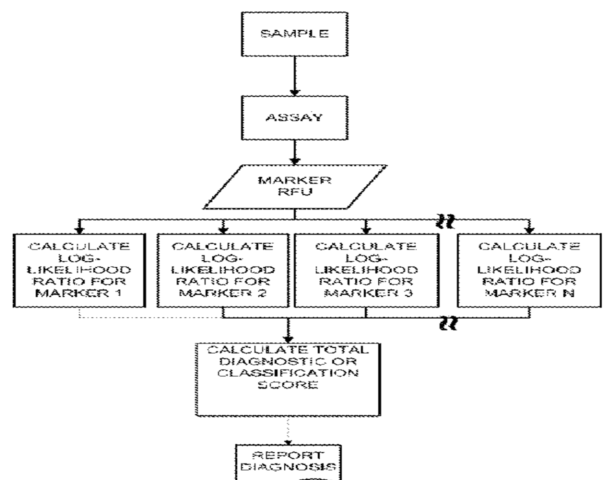


FIG. 1B