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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
21 July 2011 (21.07.2011)

PCT

(10) International Publication Number
WO 2011/088226 A2

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

(21) International Application Number:
PCT/US2011/021160

(22) International Filing Date:
13 January 2011 (13.01.2011)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/294,819 13 January 2010 (13.01.2010) US

(71) Applicant *(for all designated States except US)*: **CARIS LIFE SCIENCES LUXEMBOURG HOLDINGS**; 102 rue de Maraîchers, L2124 Luxembourg, Grand- Duché de Luxembourg (DE).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **KUSLICH, Christine** [US/US]; 1185 W. Spur Avenue, Gilbert, AZ 85233 (US). **PAWLOWSKI, Traci** [US/US]; 2014 N Milkweed Loop, Phoenix, AZ 85037 (US). **SPETZLER, David** [US/US]; 13539 N. 95th Way, Scottsdale, AZ 85260 (US).

(74) Agent: **AKHAVAN, Ramin**; Caris Life Sciences, Inc., 6655 N. Macarthur Blvd., Irving, TX 75039 (US).

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

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

Published:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

WO 2011/088226 A2

(54) Title: DETECTION OF GASTROINTESTINAL DISORDERS

(57) Abstract: Provided herein are methods and systems for characterizing a phenotype by detecting microRNAs, vesicles, or biomarkers that are indicative of disease or disease progress. The disease can be a gastrointestinal disorder, such as colorectal cancer. The microRNAs, vesicles, or biomarkers can be detected in a bodily fluid.

DETECTION OF GASTROINTESTINAL DISORDERS

CROSS-REFERENCE

[0001] This application claims the benefit of priority to U.S. Provisional Application No. 61/294,819, filed January 13, 2010, which application is incorporated herein by reference in its entirety.

BACKGROUND

[0002] Biomarkers for conditions and diseases such as cancer include biological molecules such as proteins, peptides, lipids, RNAs, DNA and variations and modifications thereof.

[0003] The identification of specific biomarkers, such as DNA, RNA and proteins, can provide biosignatures that are used for the diagnosis, prognosis, or theranosis of conditions or diseases. MicroRNAs are short RNAs that regulate the transcription and degradation of messenger RNAs. MicroRNAs have been found in bodily fluids and have been observed as a component within vesicles shed from tumor cells. Vesicles shed from cells have been found in a number of body fluids, including blood plasma, breast milk, bronchoalveolar lavage fluid and urine. Vesicles are known to take part in the communication between cells, as transport vehicles for proteins, RNAs, DNAs, viruses, and prions. The analysis of microRNA and/or vesicles associated with diseases such as cancer can aid in detection of disease or severity thereof, determining predisposition to a disease, as well as making treatment decisions.

[0004] Vesicles present in a biological sample provide a source of biomarkers, e.g., the markers are present within a vesicle (vesicle payload), or are present on the surface of a vesicle. Characteristics of vesicles (e.g., size, surface antigens, determination of cell-of-origin, payload) can also provide a diagnostic, prognostic or theranostic readout. There remains a need to identify biomarkers that can be used to detect and treat disease. microRNA and other biomarkers associated with vesicles as well as the characteristics of a vesicle can provide a diagnosis, prognosis, or theranosis.

[0005] The present invention provides methods and systems for characterizing a phenotype by detecting microRNAs and/or vesicles that are indicative of disease or disease progress.

SUMMARY

[0006] Provided herein are methods for assessing biosignature to characterize a phenotype of a biological sample. The biosignature can comprise microRNA, vesicles and/or additional biomarkers.

[0007] In an aspect, the invention provides a method for characterizing a phenotype in a sample, comprising: i) identifying a biosignature in the sample, wherein the biosignature comprises a level of one or more microRNA selected from the group consisting of miR-548c-5p, miR-362-3p, miR-422a, miR-597, miR-429, miR-200a, and miR-200b; and ii) comparing the biosignature to a reference, wherein a difference in the level of the one or more microRNA in the sample compared to the reference indicate the phenotype, thereby characterizing the phenotype. In some embodiments, the one or more microRNA consist of mature microRNA.

[0008] The sample under assessment can be a biological sample from a cell line. The sample can also be a biological sample from a subject. Identifying the biosignature can be performed in vitro.

[0009] In some embodiments, the phenotype comprises cancer. The characterizing can be a diagnosis, prognosis, determination of drug efficacy, monitoring the status of the subject's response or resistance to a treatment or selection of a treatment for the cancer. In an embodiment, the characterizing comprises diagnosing

and the phenotype comprises a cancer. Thus, the invention provides a method of diagnosing, prognosing or theranosing a cancer by assessing a biosignature.

[0010] The characterizing includes theranosis. Theranostics include use of diagnostic tests that can identify subjects are suited for a drug and/or provide feedback on how well the drug is working in a subject. In an embodiment, the subject is non-responsive to a treatment being administered to the subject. The treatment can be administering a cancer therapeutic.

[0011] In some embodiments, the sample comprises a bodily fluid. The bodily fluid can be peripheral blood, sera, plasma, ascites, urine, cerebrospinal fluid (CSF), sputum, saliva, bone marrow, synovial fluid, aqueous humor, amniotic fluid, cerumen, breast milk, broncheoalveolar lavage fluid, semen, prostatic fluid, cowper's fluid or pre-ejaculatory fluid, female ejaculate, sweat, fecal matter, hair, tears, cyst fluid, pleural and peritoneal fluid, pericardial fluid, lymph, chyme, chyle, bile, interstitial fluid, menses, pus, sebum, vomit, vaginal secretions, mucosal secretion, stool, stool water, pancreatic juice, lavage fluids from sinus cavities, bronchopulmonary aspirates, blastocyl cavity fluid, or umbilical cord blood. The bodily fluid can be peripheral blood, sera, plasma, saliva or stool.

[0012] The cancer that is characterized according to the invention can be an acute lymphoblastic leukemia; acute myeloid leukemia; adrenocortical carcinoma; AIDS-related cancers; AIDS-related lymphoma; anal cancer; appendix cancer; astrocytomas; atypical teratoid/rhabdoid tumor; basal cell carcinoma; bladder cancer; brain stem glioma; brain tumor (including brain stem glioma, central nervous system atypical teratoid/rhabdoid tumor, central nervous system embryonal tumors, astrocytomas, craniopharyngioma, ependymoblastoma, ependymoma, medulloblastoma, medulloepithelioma, pineal parenchymal tumors of intermediate differentiation, supratentorial primitive neuroectodermal tumors and pineoblastoma); breast cancer; bronchial tumors; Burkitt lymphoma; cancer of unknown primary site; carcinoid tumor; carcinoma of unknown primary site; central nervous system atypical teratoid/rhabdoid tumor; central nervous system embryonal tumors; cervical cancer; childhood cancers; chordoma; chronic lymphocytic leukemia; chronic myelogenous leukemia; chronic myeloproliferative disorders; colon cancer; colorectal cancer; craniopharyngioma; cutaneous T-cell lymphoma; endocrine pancreas islet cell tumors; endometrial cancer; ependymoblastoma; ependymoma; esophageal cancer; esthesioneuroblastoma; Ewing sarcoma; extracranial germ cell tumor; extragonadal germ cell tumor; extrahepatic bile duct cancer; gallbladder cancer; gastric (stomach) cancer; gastrointestinal carcinoid tumor; gastrointestinal stromal cell tumor; gastrointestinal stromal tumor (GIST); gestational trophoblastic tumor; glioma; hairy cell leukemia; head and neck cancer; heart cancer; Hodgkin lymphoma; hypopharyngeal cancer; intraocular melanoma; islet cell tumors; Kaposi sarcoma; kidney cancer; Langerhans cell histiocytosis; laryngeal cancer; lip cancer; liver cancer; malignant fibrous histiocytoma bone cancer; medulloblastoma; medulloepithelioma; melanoma; Merkel cell carcinoma; Merkel cell skin carcinoma; mesothelioma; metastatic squamous neck cancer with occult primary; mouth cancer; multiple endocrine neoplasia syndromes; multiple myeloma; multiple myeloma/plasma cell neoplasm; mycosis fungoides; myelodysplastic syndromes; myeloproliferative neoplasms; nasal cavity cancer; nasopharyngeal cancer; neuroblastoma; Non-Hodgkin lymphoma; nonmelanoma skin cancer; non-small cell lung cancer; oral cancer; oral cavity cancer; oropharyngeal cancer; osteosarcoma; other brain and spinal cord tumors; ovarian cancer; ovarian epithelial cancer; ovarian germ cell tumor; ovarian low malignant potential tumor; pancreatic cancer; papillomatosis; paranasal sinus cancer; parathyroid cancer; pelvic cancer; penile cancer; pharyngeal cancer; pineal parenchymal

tumors of intermediate differentiation; pineoblastoma; pituitary tumor; plasma cell neoplasm/multiple myeloma; pleuropulmonary blastoma; primary central nervous system (CNS) lymphoma; primary hepatocellular liver cancer; prostate cancer; rectal cancer; renal cancer; renal cell (kidney) cancer; renal cell cancer; respiratory tract cancer; retinoblastoma; rhabdomyosarcoma; salivary gland cancer; Sézary syndrome; small cell lung cancer; small intestine cancer; soft tissue sarcoma; squamous cell carcinoma; squamous neck cancer; stomach (gastric) cancer; supratentorial primitive neuroectodermal tumors; T-cell lymphoma; testicular cancer; throat cancer; thymic carcinoma; thymoma; thyroid cancer; transitional cell cancer; transitional cell cancer of the renal pelvis and ureter; trophoblastic tumor; ureter cancer; urethral cancer; uterine cancer; uterine sarcoma; vaginal cancer; vulvar cancer; Waldenström macroglobulinemia; or Wilm's tumor. In some embodiments, the cancer comprises a gastrointestinal cancer, gastric cancer, hepatocellular carcinoma, liver cancer, gastrointestinal stromal tumor (GIST), esophageal cancer, pancreatic cancer or colorectal cancer, e.g., the cancer can be colorectal cancer. The colorectal cancer can be Dukes B, Dukes C or Dukes D.

[0013] In an embodiment, the reference that is compared with the biosignature comprises a different individual or group of individuals as compared to the subject. In another embodiment, the reference comprises samples obtained from the subject over a time course. The reference can also include both samples from a different individual or group of individuals as compared to the subject, or samples from the same subject.

[0014] In some embodiments of the methods of the invention, the sample comprises a vesicle. The vesicle can have a diameter of about 30 nm to about 800 nm, e.g., the vesicle can have a diameter of about 30 nm to about 200 nm.

[0015] Vesicles can be analyzed as part of the biosignature. In some embodiments, vesicle components can be part of the biosignature. The vesicles can be isolated prior to analysis thereof. The vesicle can be isolated prior to step identifying a biosignature using size exclusion chromatography, density gradient centrifugation, differential centrifugation, flow cytometry, high pressure liquid chromatography, flow pressure liquid chromatography, membrane ultrafiltration, affinity capture, microfluidic device, or combinations thereof.

[0016] In some embodiments, the vesicle is a cell-of-origin specific vesicle. In some embodiment, the cell-of-origin is a tumor or cancer cell. In other embodiments, the cell-of-origin is a lung, pancreas, stomach, intestine, bladder, kidney, ovary, testis, skin, colorectal, breast, prostate, brain, esophagus, liver, placenta, or fetal cell. In still other embodiments, the cell-of-origin is a stomach, intestine, colorectal, esophagus, or liver cell.

[0017] The one or more microRNA used to characterize the phenotype can be present in the vesicle as vesicle payload. One embodiment of identifying the biosignature comprises: i) isolating the vesicle by: a. filtering the biological sample with a filter that retains the vesicle; b. capturing the retained vesicle with a capture binding agent; c. contacting the captured vesicle with a labeled binding agent; and d. isolating the labeled vesicle; ii) extracting nucleic acid from the isolated vesicle; and iii) detecting the one or more microRNA in the extracted nucleic acid. The capture agent can be a binding agent to one or more vesicle marker such as DR3, STEAP, epha2, TMEM211, unc93A, A33, CD24, NGAL, EpCam, MUC17, TROP2, and/or TETS. These markers can also be used as an antigen for the labeled binding agent. The labeled binding agent can be to a general vesicle marker, e.g., to one or more of CD9, CD63 and CD81. The labeled binding agent can be to DR3, STEAP, epha2, TMEM211, unc93A, A33, CD24, NGAL, EpCam, MUC17, TROP2, and/or TETS.

[0018] The capture binding agent and/or labeled binding agent used in the method of the invention can be an antibody, antibody fragment or an aptamer. In some embodiments, the capture binding agent is tethered to a

substrate. The substrate can be an array, well, or particle. In an embodiment, the substrate can be a labeled particle, e.g., a bead comprising a label. The label can be fluorescent label. The labeled binding agent, also referred to as a detection binding agent, can also be fluorescently labeled. The labeled vesicle can be isolated using flow cytometry. The substrate particle can also be magnetic. Thus, vesicles captured by the substrate can be isolated using magnetic force.

[0019] In some embodiments of the method of the invention, the biosignature further comprises a level or presence of one or more general vesicle biomarker and a level or presence of one or more cell-of-origin biomarker. In other embodiments, the biosignature further comprises a level or presence of one or more general vesicle biomarker, and a level or presence of one or more disease specific biomarker. In still other embodiments, the biosignature further comprises a level or presence of one or more general vesicle biomarker, a level or presence of one or more cell-of-origin biomarker, and a level or presence of one or more disease specific biomarker. In an embodiment, the biosignature further comprises a level or presence of one or more cell-of-origin biomarker, and a level or presence of one or more disease specific biomarker. The one or more general vesicle biomarker can be one or more of CD9, CD63, CD81, CD37, CD53, CD82, or Rab-5b. The one or more disease specific biomarker can be a cancer biomarker. The cell-of-origin can be a cell derived from the gastrointestinal tract, e.g., a gastrointestinal cell, gastric cell, hepatocellular cell, liver cell, gastrointestinal stromal cell, esophageal cell, pancreatic cell or colorectal cell, including a cell derived from the duodenum, small intestine, large intestine, or rectum.

[0020] In some embodiments, the biosignature identified by the method further comprises a level or presence of one or more additional biomarker selected from the group consisting of DR3, STEAP, epha2, TMEM211, unc93A, A33, CD24, NGAL, EpCam, MUC17, TROP2, and TETS. The one or more additional biomarker can be at least TMEM211. The one or more additional biomarker can be at least CD24. In some embodiments, the one or more additional biomarker includes both TMEM211 and CD24. The biomarker can be assessed at the nucleic acid level, e.g., DNA or mRNA, or protein level, e.g., as a circulating biomarker or as a vesicle antigen. In an embodiment, the one or more additional biomarker comprises an mRNA, a circulating biomarker, or a protein. The one or more additional biomarker can also be associated with a vesicle. For example, the one or more additional biomarker can be a surface antigen of the vesicle. The one or more additional biomarker can also be a component of the payload within the vesicle.

[0021] In some embodiments, the biosignature identified by the method comprises an expression level, presence, absence, mutation, copy number variation, truncation, duplication, insertion, modification, sequence variation, or molecular association of one or more biomarker. Such biomarker can be a nucleic acid, peptide, protein, lipid, antigen, carbohydrate, proteoglycan, vesicle, or a combination thereof. Detection of the biomarker can use microarray analysis, PCR, hybridization with allele-specific probes, enzymatic mutation detection, ligation chain reaction (LCR), oligonucleotide ligation assay (OLA), flow-cytometric heteroduplex analysis, chemical cleavage of mismatches, mass spectrometry, nucleic acid sequencing, single strand conformation polymorphism (SSCP), denaturing gradient gel electrophoresis (DGGE), temperature gradient gel electrophoresis (TGGE), restriction fragment polymorphisms, serial analysis of gene expression (SAGE), image cytometry, qRT-PCR, real-time PCR, PCR, flow cytometry, mass spectrometry, or a combination thereof. The one or more biomarker can also detected using a binding agent. The binding agent can be selected from the group consisting of: DNA, RNA, monoclonal antibodies, polyclonal antibodies, Fabs, Fab', single chain

antibodies, synthetic antibodies, DNA aptamers, RNA aptamers, lectin, peptoids, zDNA, peptide nucleic acids (PNAs), locked nucleic acids (LNAs), synthetic occurring chemical compounds, naturally occurring chemical compounds, dendrimers, and combinations thereof.

[0022] In another aspect, the invention provides use of an agent for determining the level of at least one microRNA in preparation of a composition for carrying out a method of characterizing a phenotype as described above.

INCORPORATION BY REFERENCE

[0023] All publications, patents, and patent applications mentioned in this specification are herein incorporated by reference to the same extent as if each individual publication, patent, or patent application was specifically and individually indicated to be incorporated by reference.

BRIEF DESCRIPTION OF THE DRAWINGS

[0024] The novel features of the invention are set forth with particularity in the appended claims. A better understanding of the features and advantages of the present invention will be obtained by reference to the following detailed description that sets forth illustrative embodiments, in which the principles of the invention are utilized, and the accompanying drawings of which:

[0025] **FIG. 1A** depicts a method of identifying a biosignature comprising microRNAs to characterize a phenotype. **FIG. 1B** depicts a method of identifying a biosignature of a vesicle to characterize a phenotype by isolating a vesicle.

[0026] **FIG. 2** illustrates methods of characterizing a phenotype by assessing vesicle biosignatures. **FIG. 2A** is a schematic of a planar substrate coated with a capture antibody, which captures vesicles expressing that protein. The capture antibody is for a vesicle protein that is specific or not specific for vesicles derived from diseased cells ("disease vesicle"). The detection antibody binds to the captured vesicle and provides a fluorescent signal. The detection antibody can detect an antigen that is generally associated with vesicles, or is associated with a cell-of-origin or a disease, e.g., a cancer. **FIG. 2B** is a schematic of a bead coated with a capture antibody, which captures vesicles expressing that protein. The capture antibody is for a vesicle protein that is specific or not specific for vesicles derived from diseased cells ("disease vesicle"). The detection antibody binds to the captured vesicle and provides a fluorescent signal. The detection antibody can detect an antigen that is generally associated with vesicles, or is associated with a cell-of-origin or a disease, e.g., a cancer. **FIG. 2C** is an example of a screening scheme that can be performed by multiplexing using the beads as shown in **FIG. 2B**. **FIG. 2D** presents illustrative schemes for capturing and detecting vesicles to characterize a phenotype. **FIG. 2E** presents illustrative schemes for assessing vesicle payload to characterize a phenotype.

[0027] **FIG. 3** illustrates protein expression patterns on a vesicle. Different proteins are typically not distributed evenly or uniformly on a vesicle shell. Vesicle-specific proteins, e.g., CD9, CD63 or CD81, are typically more common, while cancer-specific proteins, e.g., CD66 or EpCam are less common. Capture of a vesicle can be more accomplished using a more common, less cancer-specific protein, and cancer-specific proteins used in the detection phase. Capture of a vesicle can also be more accomplished using a less common, cancer-specific protein and using more common, less cancer-specific proteins used in the detection phase to increase the signal of the captured vesicles.

[0028] **FIG. 4** illustrates colon cancer vesicle biosignatures. **FIG. 4A** depicts histograms of intensity values collected from various multiplexing experiments using a bead-based platform, where beads were functionalized with a capture antibody, incubated with vesicles purified form patient plasma, and then labeled with a detector antibody. The darker shaded bars (blue) represent the population from normals and the lighter shaded bars (green) are from colon cancer patients. **FIG. 4B** shows a normalized graph for each of the histograms shown in **FIG. 4A**. **FIG. 4C** depicts a histogram of intensity values collected from a multiplexing experiment where beads where functionalized with CD66 antibody (the capture antibody), incubated with vesicles purified from patient plasma, and then labeled with a PE conjugated EpCam antibody (the detector antibody). The red population is from six normals and the green is from 21 colon cancer patients. Data from each individual was normalized to account for variation in the number of beads, added together, and then normalized again to account for the different number of samples in each population.

[0029] **FIG. 5** illustrates a colon cancer biosignature for colon cancer by stage, using CD63 detector and CD63 capture. The histograms of intensities from vesicles captured with CD63 coated beads and labeled with CD63 conjugated PE. There are 6 patients in the control group **(A)**, 4 in stage I **(B),** 5 in stage II **(C),** 8 in stage III **(D),** and 4 stage IV **(E).** Data from each individual was normalized to account for variation in the number of beads, added together, and then normalized again to account for the different number of samples in each population **(F).**

[0030] **FIG. 6** illustrates colon cancer biosignature for colon cancer by stage, using EpCam detector and CD9 capture. The histograms of intensities are from vesicles captured with CD9 coated beads and labeled with EpCam. There are patients in the **(A)** control group, **(B)** stage I, **(C)** stage II, **(D)** stage III, and **(E)** stage IV. Data from each individual was normalized to account for variation in the number of beads, added together, and then normalized again to account for the different number of samples in each population **(F).**

[0031] **FIG. 7** illustrates TaqMan Low Density Array (TLDA) miRNA card comparison of colorectal cancer (CRC) cell lines versus normal vesicles. The CRC cell lines are indicated to the right of the plot. The Y-axis shows a fold-change in expression in the CRC cell lines compared to normal controls. The miRNAs surveyed are indicated on the X-axis. These miRNAs were not overexpressed in normal or melanoma cells. For each miRNA, the CRC cell lines from left to right are LOVO, HT29, SW260, COLO205, HCT116 and RKO.

[0032] **FIG. 8A** illustrates the sensitivity and the confidence level for detecting colon cancer using antibodies to the listed proteins listed as the detector and capture antibodies. CD63, CD9 are general markers, EpCam is a cancer marker, and CD66 is a colon marker. The individual results are depicted in **FIG. 8B** for EpCam versus CD63, with 99% confidence, 95% (n=20) cancer patient samples were different from the Generalized Normal Distribution; with 99% confidence, 100% (n=6) normal patient samples were not different from the Generalized Normal Distribution; **FIG. 8C** for EpCam versus CD9, with 99% confidence, 90% (n=20) cancer patient samples were different from the Generalized Normal Distribution; with 99% confidence, 77% (n=6) normal patient samples were not different from the Generalized Normal Distribution; **FIG. 8D** for CD63 versus CD63, with 99% confidence, 60% (n=20) cancer patient samples were different from the Generalized Normal Distribution; with 99% confidence, 80% (n=6) normal patient samples were not different from the Generalized Normal Distribution; **FIG. 8E** for CD9 versus CD63, with 99% confidence, 90% (n=20) cancer patient samples were different from the Generalized Normal Distribution; with 99% confidence, 77% (n=6) normal patient samples were not different from the Generalized Normal Distribution; and **FIG. 8F** for CD66 versus CD9, with

99% confidence, 90% (n=20) cancer patient samples were different from the Generalized Normal Distribution; with 99% confidence, 77% (n=6) normal patient samples were not different from the Generalized Normal Distribution.

[0033] **FIG. 9** illustrates a schematic for a vesicle PCa assay (**FIG. 9A**), which leads to a decision tree (**FIG. 9B** and **FIG. 9C**).

[0034] **FIG. 10** illustrates the ability of a vesicle bio-signature to discriminate between normal prostate and PCa samples. Cancer markers included EpCam and B7H3. General vesicle markers included CD9, CD81 and CD63. Prostate specific markers included PCSA. The test was found to be 98% sensitive and 95% specific for PCa vs normal samples. **FIG. 10B** illustrates mean fluorescence intensity (MFI) on the Y axis for vesicle markers of **FIG. 10A** in normal and prostate cancer patients.

[0035] **FIG. 11** illustrates the top ten differentially expressed microRNAs between VCaP prostate cancer cell derived vesicles and normal plasma vesicles. VCAP cell line vesicles and vesicles from normal plasma were isolated via ultracentrifugation followed by RNA isolation. MicroRNAs were profiled using qRT-PCR analysis. Prostate cancer cell line derived vesicles have higher levels (lower CT values) of the indicated microRNAs as depicted in the bar graph.

[0036] **FIG. 12A** depicts a bar graph of miR-21 expression with CD9 bead capture. 1 ml of plasma from prostate cancer patients, 250 ng/ml of LNCaP, or normal purified vesicles were incubated with CD9 coated Dynal beads. The RNA was isolated from the beads and the bead supernatant. One sample (#6) was also uncaptured for comparison. MiR-21 expression was measured with qRT-PCR and the mean CT values for each sample compared. CD9 capture improves the detection of miR-21 in prostate cancer samples. **FIG. 12B** depicts a bar graph of miR-141 expression with CD9 bead capture. The experiment was performed as in **FIG. 12A**, with miR-141 expression measured with qRT-PCR instead of miR-21.

DETAILED DESCRIPTION OF THE INVENTION

[0037] Disclosed herein are methods and systems for isolating, storing, and analyzing microRNA to characterize a phenotype of biological sample, e.g., a sample from a cell culture, an organism, or a subject. MicroRNA can be obtained directly from a biological sample, such as a bodily fluid or culture medium. The microRNAs of the invention can also be obtained from t vesicles, which are isolated from a subject. Unless otherwise specified, the terms "purified" or "isolated" as used herein in reference to vesicles or biomarker components mean partial or complete purification or isolation of such components from a cell or organism.

[0038] A method of characterizing a phenotype by analyzing microRNA is depicted in scheme **100A** of **FIG. 1A**. In a first step **101**, a biological sample is obtained, e.g., a bodily fluid, tissue sample or cell culture. Nucleic acids are isolated from the sample **103**. The nucleic acid can be DNA or RNA, e.g., microRNA. Such nucleic acids can provide a biosignature for a phenotype. By sampling the nucleic acids associated with target phenotype (e.g., disease versus healthy, pre- and post-treatment), one or more nucleic acid markers that are indicative of the phenotype can be determined. Various aspects of the present invention are directed to biosignatures determined by assessing one or more nucleic acid molecules (e.g., microRNA) present in the sample **105**, where the biosignature corresponds to a predetermined phenotype **107**. **FIG. 1B** illustrates a scheme **100B** of using vesicles to isolate the nucleic acid molecules. In one example, a biological sample is obtained **102**, and one or more vesicles, e.g., vesicles from a desirable cell-of-origin, are isolated from the sample **104**. The vesicles are analyzed **106**, which can include characterizing surface antigens associated with

the vesicles and/or determining the presence or levels of components present within the vesicles ("payload"). The payload may be protein, including peptides and polypeptides, and/or nucleic acids such as DNA and RNAs. RNA payload includes messenger RNA (mRNA) and microRNA (also referred to herein as miRNA or miR). A phenotype is characterized based on the biosignature of the vesicles **108**.

MicroRNA

[0039] MicroRNAs, also referred to herein as miRNAs or miRs, are short RNA strands approximately 21-23 nucleotides in length. MiRNAs are encoded by genes that are transcribed from DNA but are not translated into protein and thus comprise non-coding RNA. The miRs are processed from primary transcripts known as pri-miRNA to short stem-loop structures called pre-miRNA and finally to the resulting single strand miRNA. The pre-miRNA typically forms a structure that fold backs on itself in self-complementary regions. These structures are then processed by the nuclease Dicer in animals or DCL1 in plants. Mature miRNA molecules are partially complementary to one or more messenger RNA (mRNA) molecules and can function to regulate translation of proteins. Identified sequences of miRNA can be accessed at publicly available databases, such as www.microRNA.org, www.mirbase.org, or www.mirz.unibas.ch/cgi/miRNA.cgi.

[0040] miRNAs are generally assigned a number according to the naming convention " mir-[number]." The number of a miRNA is assigned according to its order of discovery relative to previously identified miRNA species. For example, if the last published miRNA was mir-121, the next discovered miRNA will be named mir-122, etc. When a miRNA is discovered that is homologous to a known miRNA from a different organism, the name can be given an optional organism identifier, of the form [organism identifier]- mir-[number]. Identifiers include hsa for Homo sapiens and mmu for Mus Musculus. For example, a human homolog to mir-121 might be referred to as hsa-mir-121 whereas the mouse homolog can be referred to as mmu-mir-121.

[0041] Mature microRNA is commonly designated with the prefix "miR" whereas the gene or precursor miRNA is designated with the prefix "mir." For example, mir-121 is a precursor for miR-121. When differing miRNA genes or precursors are processed into identical mature miRNAs, the genes/precursors can be delineated by a numbered suffix. For example, mir-121-1 and mir-121-2 can refer to distinct genes or precursors that are processed into miR-121. Lettered suffixes are used to indicate closely related mature sequences. For example, mir-121a and mir-121b can be processed to closely related miRNAs miR-121a and miR-121b, respectively. In the context of the invention, any microRNA (miRNA or miR) designated herein with the prefix mir-* or miR-* is understood to encompass both the precursor and/or mature species, unless otherwise explicitly stated otherwise.

[0042] Sometimes it is observed that two mature miRNA sequences originate from the same precursor. When one of the sequences is more abundant that the other, a "*" suffix can be used to designate the less common variant. For example, miR-121 would be the predominant product whereas miR-121* is the less common variant found on the opposite arm of the precursor. If the predominant variant is not identified, the miRs can be distinguished by the suffix "5p" for the variant from the 5' arm of the precursor and the suffix "3p" for the variant from the 3' arm. For example, miR-121-5p originates from the 5' arm of the precursor whereas miR-121-3p originates from the 3' arm. Less commonly, the 5p and 3p variants are referred to as the sense ("s") and anti-sense ("as") forms, respectively. For example, miR-121-5p may be referred to as miR-121-s whereas miR-121-3p may be referred to as miR-121-as.

[0043] The above naming conventions have evolved over time and are general guidelines rather than absolute rules. For example, the let- and lin- families of miRNAs continue to be referred to by these monikers. The mir/miR convention for precursor/mature forms is also a guideline and context should be taken into account to determine which form is referred to. Further details of miR naming can be found at www.mirbase.org or Ambros et al., A uniform system for microRNA annotation, RNA **9**:277-279 (2003).

[0044] Plant miRNAs follow a different naming convention as described in Meyers et al., Plant Cell. 2008 20(12):3186-3190.

[0045] A number of miRNAs are involved in gene regulation, and miRNAs are part of a growing class of non-coding RNAs that is now recognized as a major tier of gene control. In some cases, miRNAs can interrupt translation by binding to regulatory sites embedded in the 3′-UTRs of their target mRNAs, leading to the repression of translation. Target recognition involves complementary base pairing of the target site with the miRNA's seed region (positions 2–8 at the miRNA's 5′ end), although the exact extent of seed complementarity is not precisely determined and can be modified by 3′ pairing. In other cases, miRNAs function like small interfering RNAs (siRNA) and bind to perfectly complementary mRNA sequences to destroy the target transcript.

[0046] Characterization of a number of miRNAs indicates that they influence a variety of processes, including early development, cell proliferation and cell death, apoptosis and fat metabolism. For example, some miRNAs, such as lin-4, let-7, mir-14, mir-23, and bantam, have been shown to play critical roles in cell differentiation and tissue development. Others are believed to have similarly important roles because of their differential spatial and temporal expression patterns.

[0047] The miRNA database available at miRBase (www.mirbase.org) comprises a searchable database of published miRNA sequences and annotation. Further information about miRBase can be found in the following articles, each of which is incorporated by reference in its entirety herein: Griffiths-Jones et al., miRBase: tools for microRNA genomics. NAR 2008 36(Database Issue):D154-D158; Griffiths-Jones et al., miRBase: microRNA sequences, targets and gene nomenclature. NAR 2006 34(Database Issue):D140-D144; and Griffiths-Jones, S. The microRNA Registry. NAR 2004 32(Database Issue):D109-D111. Representative miRNAs contained in Release 16 of miRBase, made available September 2010.

[0048] Techniques to isolate and characterize vesicles and miRs are known to those of skill in the art. See, e.g., PCT Patent Application PCT/US09/06095, entitled "METHODS AND SYSTEMS OF USING EXOSOMES FOR DETERMINING PHENOTYPES" and filed November 12, 2009; and U.S. Patent Application Serial No. 12/609,847, entitled "METHODS FOR ASSESSING RNA PATTERNS" and filed October 30, 2009; both of which applications are incorporated by reference herein in their entirety.

Vesicles

[0049] Methods of the invention can include assaying one or more vesicles. A vesicle, as used herein, is a membrane vesicle that is shed from cells. Vesicles or membrane vesicles include without limitation: microvesicle, exosome, nanovesicle, dexosome, bleb, blebby, prostasome, microparticle, intralumenal vesicle, membrane fragment, intralumenal endosomal vesicle, endosomal-like vesicle, exocytosis vehicle, endosome vesicle, endosomal vesicle, apoptotic body, multivesicular body, secretory vesicle, phospholipid vesicle, liposomal vesicle, argosome, texasome, secresome, tolerosome, melanosome, oncosome, or exocytosed vehicle. Unless otherwise specified, methods that make use of a species of vesicle can be applied to other types of

vesicles. Vesicles comprise spherical structures with a lipid bilayer similar to cell membranes which surrounds an inner compartment which can contain soluble components, sometimes referred to as the payload. In some embodiments, the methods of the invention make use of exosomes, which are small secreted vesicles of about 40–100 nm in diameter. For a review of membrane vesicles, including types and characterizations, see *Thery et al., Nat Rev Immunol. 2009 Aug;9(8):581-93*. Some properties of different types of vesicles include those in **Table 1**:

Table 1: Vesicle Properties

Feature	Exosomes	Microvesicles	Ectosomes	Membrane particles	Exosome-like vesicles	Apoptotic vesicles
Size	50-100 nm	100-1,000 nm	50-200 nm	50-80 nm	20-50 nm	50-500 nm
Density in sucrose	1.13-1.19 g/ml			1.04-1.07 g/ml	1.1 g/ml	1.16-1.28 g/ml
EM appearance	Cup shape	Irregular shape, electron dense	Bilamellar round structures	Round	Irregular shape	Heterogeneous
Sedimentation	100,000 g	10,000 g	160,000-200,000 g	100,000-200,000 g	175,000 g	1,200 g, 10,000 g, 100,000 g
Lipid composition	Enriched in cholesterol, sphingomyelin and ceramide; contains lipid rafts; expose PPS	Expose PPS	Enriched in cholesterol and diacylglycerol; expose PPS		No lipid rafts	
Major protein markers	Tetraspanins (CD63, CD9), Alix, TSG101	Integrins, selectins and CD40 ligand	CR1 and proteolytic enzymes; no CD63	CD133; no CD63	TNFRI	Histones
Intracellular origin	Internal compartments (endosomes)	Plasma membrane	Plasma membrane	Plasma membrane		

Abbreviations: phosphatidylserine (PPS); electron microscopy (EM)

(a) Vesicles include shed membrane bound particles, or "microparticles," that are derived from either the plasma membrane or an internal membrane. Vesicles can be released into the extracellular environment from cells. Cells releasing vesicles include without limitation cells that originate from, or are derived from, the ectoderm, endoderm, or mesoderm. The cells may have undergone genetic, environmental, and/or any other variations or alterations. For example, the cell can be tumor cells. A vesicle can reflect any changes in the source cell, and thereby reflect changes in the originating cells, e.g., cells having various genetic mutations. In one mechanism, a vesicle is generated intracellularly when a segment of the cell membrane spontaneously invaginates and is ultimately exocytosed (see for example, *Keller et al., Immunol. Lett. 107 (2): 102–8 (2006)*). Vesicles also include cell-derived structures bounded by a lipid bilayer membrane arising from both herniated evagination (blebbing) separation and sealing of portions of the plasma membrane or from the export of any intracellular membrane-bounded vesicular structure containing various membrane-associated proteins of tumor origin, including surface-bound molecules derived from the host circulation that bind selectively to the tumor-derived proteins together with molecules contained in the vesicle lumen, including but not limited to tumor-derived microRNAs or intracellular proteins. Blebs and blebbing are further described in *Charras et al., Nature Reviews Molecular and Cell Biology, Vol. 9, No. 11, p. 730-736 (2008)*. A vesicle

shed into circulation or bodily fluids from tumor cells may be referred to as a "circulating tumor-derived vesicle." When such vesicle is an exosome, it may be referred to as a circulating-tumor derived exosome (CTE). In some instances, a vesicle can be derived from a specific cell of origin. CTE, as with a cell-of-origin specific vesicle, typically have one or more unique biomarkers that permit isolation of the CTE or cell-of-origin specific vesicle, e.g., from a bodily fluid and sometimes in a specific manner. For example, a cell or tissue specific markers are utilized to identify the cell of origin. Examples of such cell or tissue specific markers are disclosed herein and can further be accessed in the Tissue-specific Gene Expression and Regulation (TiGER) Database, available at bioinfo.wilmer.jhu.edu/tiger/; Liu et al. (2008) TiGER: a database for tissue-specific gene expression and regulation. BMC Bioinformatics. 9:271; TissueDistributionDBs, available at genome.dkfz-heidelberg.de/menu/tissue_db/index.html.

[0050] A vesicle can have a diameter of greater than about 10, 20, or 30 nm. They can have a diameter of about 30-1000 nm, about 30-800 nm, about 30-200 nm, or about 30-100 nm. In some embodiments, the vesicle has a diameter of less than 10,000 nm, 1000 nm, 800 nm, 500 nm, 200 nm, 100 nm or 50 nm. As used herein the term "about" in reference to a numerical value means that variations of 10% above or below the numerical value are within the range ascribed to the specified value. Typical sizes for various types of vesicles are shown in **Table 1**.

[0051] Vesicles can be directly assayed from a biological sample. The level or amount of vesicles in the sample can itself provide a biosignature. Methods can be used to determine a vesicle biosignature without prior isolation, purification, or concentration of a vesicle from a biological sample, i.e., directly from a sample. Alternatively, the vesicle in the sample may be isolated, purified, or concentrated from a sample prior to analysis. The isolation can be performed using various techniques as described herein, e.g., chromatography, filtration, centrifugation, flow cytometry, affinity capture (e.g., to a planar surface or bead), and/or using microfluidics.

[0052] Vesicles such as exosomes can be assessed to provide a phenotypic characterization by comparing vesicle characteristics to a reference. In some embodiments, surface antigens on a vesicle are assessed. The surface antigens can provide an indication of the anatomical origin of the vesicles and other phenotypic information, e.g., tumor status. For example, wherein vesicles found in a patient sample, e.g., a bodily fluid such as blood, serum or plasma, are assessed for surface antigens indicative of colorectal origin and the presence of cancer. The surface antigens may comprise any informative biological entity that can be detected on the vesicle membrane surface, including without limitation surface proteins, lipids, carbohydrates, and other membrane components. Positive detection of colon derived vesicles expressing tumor antigens can indicate that the patient has colorectal cancer. One of skill will understand that this approach can be applied to other anatomical or cellular origins beyond colorectal and to other disease states beyond cancer. In another embodiment, vesicle payload is assessed to provide a phenotypic characterization. The payload with a vesicle comprises any informative biological entity that can be detected encapsulated with the vesicle, including without limitation proteins and nucleic acids, e.g., mRNA and microRNAs (miRs). In still another embodiment, vesicle surface antigens and vesicle payload are both assessed to provide a phenotypic characterization. For example, vesicles can be isolated using antibodies against specific surface antigens, and the payload of the isolated vesicles is assessed. The payload may comprise detecting the level of specific miRs. As described herein, the levels of vesicles with surface antigens of interest or with payload of interest can be compared to a reference to

characterize a phenotype. For example, overexpression in a sample of cancer-related surface antigens or vesicle payload, e.g., a tumor associated mRNA or microRNA, as compared to a reference can indicate the presence of cancer in the sample.

Circulating Biomarkers

[0053] Circulating biomarkers include biomarkers that are detectable in body fluids, such as blood, plasma, serum. Examples of circulating cancer biomarkers include cardiac troponin T (cTnT), prostate specific antigen (PSA) for prostate cancer and CA125 for ovarian cancer. Circulating biomarkers according to the invention include any appropriate biomarker that can be detected in bodily fluid, including without limitation protein, nucleic acids, e.g., DNA, mRNA and microRNA, lipids, carbohydrates and metabolites. Circulating biomarkers can include biomarkers that are not associated with cells, such as biomarkers that are membrane associated, embedded in membrane fragments, part of a biological complex, or free in solution. Circulating biomarkers have been identified for use in characterization of various phenotypes. See, e.g., Ahmed N, et al., Proteomic-based identification of haptoglobin-1 precursor as a novel circulating biomarker of ovarian cancer. Br. J. Cancer 2004; Mathelin et al., Circulating proteinic biomarkers and breast cancer, Gynecol Obstet Fertil. 2006 Jul-Aug;34(7-8):638-46. Epub 2006 Jul 28; Ye et al., Recent technical strategies to identify diagnostic biomarkers for ovarian cancer. Expert Rev Proteomics. 2007 Feb;4(1):121-31; Carney, Circulating oncoproteins HER2/neu, EGFR and CAIX (MN) as novel cancer biomarkers. Expert Rev Mol Diagn. 2007 May;7(3):309-19; Gagnon, Discovery and application of protein biomarkers for ovarian cancer, Curr Opin Obstet Gynecol. 2008 Feb;20(1):9-13; Pasterkamp et al., Immune regulatory cells: circulating biomarker factories in cardiovascular disease. Clin Sci (Lond). 2008 Aug;115(4):129-31; PCT Patent Publication WO/2007/088537; U.S. Patents 7,745,150 and 7,655,479; U.S. Patent Publications 20110008808, 20100330683, 20100248290, 20100222230, 20100203566, 20100173788, 20090291932, 20090239246, 20090226937, 20090111121, 20090004687, 20080261258, 20080213907, 20060003465, 20050124071, and 20040096915, each of which applications is incorporated herein by reference in its entirety.

Samples

[0054] MicroRNA and/or vesicles can be isolated from a biological sample obtained from the subject. A subject or patient can include, but is not limited to, mammals such as bovine, avian, canine, equine, feline, ovine, porcine, or primate animals (including humans and non-human primates). A subject may also include mammals of importance due to being endangered, such as Siberian tigers; or economic importance, such as animals raised on farms for consumption by humans, or animals of social importance to humans such as animals kept as pets or in zoos. Examples of such animals include but are not limited to: carnivores such as cats and dogs; swine including pigs, hogs and wild boars; ruminants or ungulates such as cattle, oxen, sheep, giraffes, deer, goats, bison, camels or horses. Also included are birds that are endangered or kept in zoos, as well as fowl and more particularly domesticated fowl, i.e. poultry, such as turkeys and chickens, ducks, geese, guinea fowl. Also included are domesticated swine and horses (including race horses). In addition, any animal species connected to commercial activities are also included such as those animals connected to agriculture and aquaculture and other activities in which disease monitoring, diagnosis, and therapy selection are routine practice in husbandry for economic productivity and/or safety of the food chain.

[0055] The subject can have a pre-existing disease or condition, such as cancer. Alternatively, the subject may not have any known pre-existing condition. The subject may also be non-responsive to an existing or past

treatment, such as a treatment for cancer. According to the invention, non-responsive includes complete or partial lack of positive response to the treatment.

[0056] The biological sample obtained from the subject may be any appropriate bodily fluid. For example, the biological sample can be peripheral blood, sera, plasma, ascites, urine, cerebrospinal fluid (CSF), sputum, saliva, bone marrow, synovial fluid, aqueous humor, amniotic fluid, cerumen, breast milk, broncheoalveolar lavage fluid, semen (including prostatic fluid), Cowper's fluid or pre-ejaculatory fluid, female ejaculate, sweat, fecal matter, hair, tears, cyst fluid, pleural and peritoneal fluid, pericardial fluid, lymph, chyme, chyle, bile, interstitial fluid, menses, pus, sebum, vomit, vaginal secretions, mucosal secretion, stool water, pancreatic juice, lavage fluids from sinus cavities, bronchopulmonary aspirates or other lavage fluids. A biological sample may also include the blastocyl cavity, umbilical cord blood, or maternal circulation which may be of fetal or maternal origin. The microRNA, vesicle or other biomarker of the invention can be detected in a serum or plasma sample.

[0057] The biological sample may also be a tissue sample or biopsy, from which microRNA and/or vesicles can be obtained. For example, if the sample is a solid sample, cells from the sample can be cultured and vesicle product induced.

[0058] The biological sample may be obtained through a third party, such as a party not performing the analysis of the microRNA. For example, the sample may be obtained through a clinician, physician, or other health care manager of a subject from which the sample is derived. In some embodiments, the biological sample is obtained by the same party analyzing the microRNA.

[0059] The volume of the biological sample used for analyzing microRNA can be in the range of between 0.1-20 mL, such as less than about 20, 15, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1 or 0.1 mL. In some embodiments, the sample is about 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.5, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 mL. In some embodiments, the sample is about 1,000, 900, 800, 700, 600, 500, 400, 300, 250, 200, 150, 100, 75, 50, 25 or 10 μl. For example, a small volume sample could be obtained by a prick or swab.

[0060] In some embodiments, analysis of one or more microRNA in a biological sample is used to determine whether an additional biological sample should be obtained for analysis. For example, analysis of one or more microRNA in a serum sample can be used to determine whether a biopsy should be obtained. Similarly, analysis of one or more microRNA in a plasma sample can be used to determine whether a biopsy should be obtained.

Vesicle Isolation

[0061] In some embodiments, the microRNAs of the invention are isolated from vesicles. A vesicle may be purified or concentrated prior to analysis. Analysis of a vesicle can include quantitiating the amount one or more vesicle populations of a biological sample. For example, a heterogeneous population of vesicles can be quantitated, or a homogeneous population of vesicles, such as a population of vesicles with a particular biomarker profile, a particular biosignature, or derived from a particular cell type (cell-of-origin specific vesicles) can be isolated from a heterogeneous population of vesicles and quantitated. Analysis of a vesicle can also include detecting, quantitatively or qualitatively, a particular biomarker profile or a biosignature, of a vesicle, as described below.

[0062] A vesicle can be stored and archived, such as in a bio-fluid bank and retrieved for analysis as necessary. A vesicle may also be isolated from a biological sample that has been previously harvested and

stored from a living or deceased subject. In addition, a vesicle may be isolated from a biological sample which has been collected as described in *King et al., Breast Cancer Res 7(5): 198-204 (2005)*. A vesicle may be isolated from an archived or stored sample. Alternatively, a vesicle may be isolated from a biological sample and analyzed without storing or archiving of the sample. Furthermore, a third party may obtain or store the biological sample, or obtain or store the vesicles for analysis.

[0063] An enriched population of vesicles can be obtained from a biological sample. For example, vesicles may be concentrated or isolated from a biological sample using size exclusion chromatography, density gradient centrifugation, differential centrifugation, nanomembrane ultrafiltration, immunoabsorbent capture, affinity purification, microfluidic separation, or combinations thereof.

[0064] Size exclusion chromatography, such as gel permeation columns, centrifugation or density gradient centrifugation, and filtration methods can be used. For example, vesicles can be isolated by differential centrifugation, anion exchange and/or gel permeation chromatography (for example, as described in US Patent Nos. 6,899,863 and 6,812,023), sucrose density gradients, organelle electrophoresis (for example, as described in U.S. Patent No. 7,198,923), magnetic activated cell sorting (MACS), or with a nanomembrane ultrafiltration concentrator. Various combinations of isolation or concentration methods can be used.

[0065] Highly abundant proteins, such as albumin and immunoglobulin, may hinder isolation of vesicles from a biological sample. For example, vesicles may be isolated from a biological sample using a system that utilizes multiple antibodies that are specific to the most abundant proteins found in blood. Such a system can remove up to several proteins at once, thus unveiling the lower abundance species such as cell-of-origin specific vesicles.

[0066] This type of system can be used for isolation of vesicles from biological samples such as blood, cerebrospinal fluid or urine. The isolation of vesicles from a biological sample may also be enhanced by high abundant protein removal methods as described in *Chromy et al. J Proteome Res 2004; 3:1120-1127*. In another embodiment, the isolation of vesicles from a biological sample may also be enhanced by removing serum proteins using glycopeptide capture as described in *Zhang et al, Mol Cell Proteomics 2005;4:144-155*. In addition, vesicles from a biological sample such as urine may be isolated by differential centrifugation followed by contact with antibodies directed to cytoplasmic or anti-cytoplasmic epitopes as described in *Pisitkun et al., Proc Natl Acad Sci U S A, 2004;101:13368-13373*.

[0067] Isolation or enrichment of vesicles from biological samples can also be enhanced by use of sonication (for example, by applying ultrasound), or the use of detergents, other membrane-active agents, or any combination thereof. For example, ultrasonic energy can be applied to a potential tumor site, and without being bound by theory, release of vesicles from the tissue can be increased, allowing an enriched population of vesicles that can be analyzed or assessed from a biological sample using one or more methods disclosed herein.

Binding Agents

[0068] A vesicle can be isolated using one or more binding agents. A binding agent is an agent that binds to a vesicle component, or vesicle biomarker, which can be any component present in a vesicle or on the vesicle. The vesicle component can be a nucleic acid (e.g. RNA or DNA), protein, peptide, polypeptide, antigen, lipid, carbohydrate, or proteoglycan. The binding agent can be a capture agent, such that a capture agent captures the vesicle by binding to a vesicle target, such as carbohydrate or glycoprotein. The capture agent can be coupled to a substrate and used to isolate the vesicle, such as described herein. A vesicle can be isolated using one or more

binding agents for a vesicle glycoprotein or carbohydrate. For example, the capture agent or binding agent can be a lectin.

[0069] A binding agent can be a lectin, nucleic acid (e.g. DNA, RNA), monoclonal antibody, polyclonal antibody, Fab, Fab', single chain antibody, synthetic antibody, aptamer (DNA/RNA), peptoid, zDNA, peptide nucleic acid (PNA), locked nucleic acid (LNA), synthetic or naturally occurring chemical compound (including but not limited to a drug or labeling reagent), dendrimer, or any combination thereof. For example, the binding agent can be a lectin and used to isolate a vesicle.

[0070] In some instances, a single binding agent is used to isolate or detect a vesicle. In other instances, a combination of different binding agents is used to isolate or detect a vesicle. For example, at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 50, 75 or 100 different binding agents may be used to isolate or detect a vesicle from a biological sample. The one or more different binding agents for a vesicle can form a vesicle biosignature in whole or in part, as further described below.

[0071] Different binding agents can be used for multiplex analysis. In some embodiments, isolation or detection of more than one population of vesicles is performed by isolating or detecting each vesicle population with a different binding agent. Different binding agents can be bound to different particles, wherein the different particles are labeled. The particles can be differently labeled in order to distinguish particles. In another embodiment, an array comprising different binding agents is used for multiplex analysis, wherein the different binding agents are differentially labeled or can be ascertained based on the location of the binding agent on the array. Multiplexing can be accomplished up to the resolution capability of the labels or detection method, as described below.

[0072] The binding agent can be a binding agent that binds vesicle "housekeeping proteins," or general vesicle biomarkers, such as CD63, CD9, CD81, CD82, CD37, CD53, or Rab-5b. Tetraspanins, a family of membrane proteins, can be used as general vesicle markers. The tetraspanins include CD151, CD53, CD37, CD82, CD81, CD9 and CD63. The binding agent can also be an agent that binds to vesicles derived from specific cell types, such as tumor cells (e.g. binding agent for EpCam or CD24) or specific cell-of-origins. The binding agent can be specific for a tumor antigen. For example, the binding agent used to isolate a vesicle may be a binding agent for an antigen selected from **Table 2**.

Table 2: Exemplary cancers by lineage, group comparisons of cells/tissue, and specific disease states and antigens specific to those cancers, group cell/tissue comparisons and specific disease states.

Cancer Lineage, Group Comparison, Disease State	Antigens	References
Colon	CEA	Park et al., 2002
Colon	MUC2	Park et al., 2002
Colon	GPA33	Huber et al., 2005
Colon	CEACAM5	Huber et al., 2005
Colon	ENFB1	Huber et al., 2006
Colon	CCSA-3	Leman et al., 2007
Colon	CCSA-4	Leman et al., 2008
Colon	ADAM 10	Choi et al., 2007, J Ethnopharmacol 110(1): 49-55.
Colon	CD44	Choi et al., 2007, J Ethnopharmacol 110(1): 49-55.
Colon	NG2	Choi et al., 2007, J Ethnopharmacol 110(1): 49-55.
Colon	ephrin-B1	Choi et al., 2007, J Ethnopharmacol 110(1): 49-55.

Colon	plakoglobin	Choi et al., 2007, J Ethnopharmacol 110(1): 49-55.
Colon	galectin-4	Choi et al., 2007, J Ethnopharmacol 110(1): 49-55.
Colon	RACK1	Choi et al., 2007, J Ethnopharmacol 110(1): 49-55
Colon	tetraspanin-8	Choi et al., 2007, J Ethnopharmacol 110(1): 49-55
Colon	FasL	Choi et al., 2007, J Ethnopharmacol 110(1): 49-55
Colon	A33	Choi et al., 2007, J Ethnopharmacol 110(1): 49-55
Colon	CEA	Choi et al., 2007, J Ethnopharmacol 110(1): 49-55
Colon	EGFR	Choi et al., 2007, J Ethnopharmacol 110(1): 49-55
Colon	dipeptidase 1	Choi et al., 2007, J Ethnopharmacol 110(1): 49-55
Colon	PTEN	Frattini et al., 2007
Colon	Na(+)-dependent glucose transporter	Wang Y et al., Pediatr Res. 1994 Oct;36(4):514-21.
Colon	UDP-glucuronosyltransferase 1A	Gong QH et al., Pharmacogenetics 11:357-368(2001).
Hepatocellular Carcinoma	HBxAg	Wang W, et al., 1991
Hepatocellular Carcinoma	HBsAg	Wang W, et al., 1991
Hepatocellular Carcinoma	NLT	Simonson GD et al., Journal of Cell Science 107, 1065-1072 (1994)
Irritable Bowel Disease	II-16	Seegert D, et al., 2001
Irritable Bowel Disease	5-HT	Kerckhoffs AP et al., Neurogastroenterol Motil. 2008 Aug;20(8):900-7.
Irritable Bowel Disease	II-1beta	Seegert D, et al., 2001
Irritable Bowel Disease	II-12	Seegert D, et al., 2001
Irritable Bowel Disease	TNF-alpha	Seegert D, et al., 2001
Irritable Bowel Disease	interferon gamma	Seegert D, et al., 2001
Irritable Bowel Disease	II-6	Seegert D, et al., 2001
Irritable Bowel Disease	Rantes	Seegert D, et al., 2001
Irritable Bowel Disease	MCP-1	Seegert D, et al., 2001
Barrett's Esophagus	p53	Hamelin R, et al., 1994
Barrett's Esophagus	MUC1	Burjonrappa SC et al., Indian J Cancer. 2007 Jan-Mar;44(1):1-5.
Barrett's Esophagus	MUC6	Glickman JN et al., Am J Surg Pathol. 2003 Oct;27(10):1357-65
GIST	PDGFRA	Yang J et al., ncer. 2008 Oct 1;113(7):1532-43
GIST	c-kit	Yang J et al., ncer. 2008 Oct 1;113(7):1532-43
GIST	NHE-3	Kulaksiz H et al., Cell Tissue Res. 2001 Mar;303(3):337-43.
Cirrhosis	NLT	Simonson GD et al., Journal of Cell Science 107,1065-1072 (1994)
Cirrhosis	HBsAg	Wang, W. et al., 1991
Esophageal cancer	CaSR	Justinich CJ et al., Am J Physiol Gastrointest Liver Physiol. 2008 Jan;294(1):G120-9.

[0073] The binding agent can be for an antigen such as 5T4, B7H3, caveolin, CD63, CD9, E-Cadherin, MFG-E8, PSCA, PSMA, Rab-5B, STEAP, TNFR1, CD81, EpCam, CD59, or CD66. The binding agent can be for an

antigen such as DR3, STEAP, epha2, TMEM211, unc93A, A33, CD24, NGAL, EpCam, MUC17, TROP2, TETS. One or more binding agents, such as one or more binding agents for two or more of the antigens, can be used for isolating a vesicle. The binding agent used can be selected based on the desire of isolating vesicles derived from particular cell types, or cell-of-origin specific vesicles.

[0074] The binding agent for a vesicle can also be selected from those listed in **Table 3**.

Table 3: Exemplary cancers by lineage, group comparisons of cells/tissue, and specific disease states and binding agents specific to those cancers, group cell/tissue comparisons and specific disease states

Cancer Lineage, Group Comparison, Other Significant Disease State	Binding Agents	Reference(s)
Colon	angiopoietin 2 specific aptamer	Sarraf-Yazdi et al., J SURG Res 146(1)16-23, 2008.
Colon	beta-catenin aptamer	Lee et al., Cancer Research 66(21)10560-6, 2006.
Colon	TCF1 aptamer	Choi et al., Mol Caner Therapy (9)2428-34, 2006.
Colon	anti-Derlin1 antibody	Ran et al., Clin Cancer Res 14(206538-45, 2008
Colon	anti-RAGE antibody	Turovskaya et al., Carcinogenesis 29(10)2035-2043, 2008.
Colon	monoclonal antibody gb3.1	Turovskaya et al., Carcinogenesis 29(10)2035-2043, 2008.
Colon	Galectin-3 binding agent	Greco et al., Glycobiology 14(9)783-92, 2004.
Colon	Cetuximab	Giuliani F, Colucci G et al., Int J Biol Markers. 2007 Jan-Mar;22(1 Suppl 4):S62-70
Colon	Panitumumab	Chua YJ, Cunningham D, Clin Colorectal Cancer. 2005 Nov;5 Suppl 2:S81-8.
Colon	Matuzumab	Chua YJ, Cunningham D, Clin Colorectal Cancer. 2005 Nov;5 Suppl 2:S81-8.
Colon	Bevacizumab	Majer M et al., Anticancer Agents Med Chem. 2007 Sep;7(5):492-503
Colon	Mac-2 binding agent	Lotz MM et al., Proc Natl Acad Sci U S A. 1993 90(18): 8319-23, "Mitogen-activated protein kinases p42mapk and p44mapk are required for fibroblast proliferation."
Adenoma versus colorectal cancer (CRC)	Complement C3	Qui et al , J of Proteome Res 7(4)1693-1703, 2008
Adenoma versus CRC	histidine-rich glycoprotein binding agent	Qui et al , J of Proteome Res 7(4)1693-1703, 2008
Adenoma versus CRC	kininogen-1 binding agent	Qui et al , J of Proteome Res 7(4)1693-1703, 2008
Adenoma versus CRC	Galectin-3 binding agent	Schoeppner HL et al., Cancer. 1995 Jun 15;75(12):2818-26.
Adenoma with low grade versus high grade dysplasia	Galectin-3 binding agent	Schoeppner HL et al., Cancer. 1995 Jun 15;75(12):2818-26.
CRC versus normal	anti-ODC monoclonal antibody	Hu HY et al., World J Gastroenterol. 2005 Apr 21;11(15):2244-8.
CRC versus normal	anti-CEA monoclonal antibody	Zhang HZ et al., Cancer Res. 1989 Oct 15;49(20):5766-73.
CRC versus normal	Mac-2 binding agent	Lotz MM et al., Proc Natl Acad Sci U S A. 1993 Apr 15;90(8):3466-70.

Pancreatic	H38-15 (HGF aptamer)	Saito T and Tomida M, DNA Cell Biol. 2005 Oct;24(10):624-33.
Pancreatic	H38-21(HGF aptamer)	Saito T and Tomida M, DNA Cell Biol. 2005 Oct;24(10):624-33.
Pancreatic	Matuzumab	Kleepspeies A et al., Clin Cancer Res. 2008 Sep 1;14(17):5426-36
Pancreatic	Cetuximanb	Burris H 3rd et al., Oncologist. 2008 Mar;13(3):289-98.
Pancreatic	Bevacizumab	Burris H 3rd et al., Oncologist. 2008 Mar;13(3):289-98.
Irritable Bowel Disease	ACCA (anti-glycan antibody)	Li X et al., World J Gastroenterol. 2008 Sep 7;14(33):5115-24.
Irritable Bowel Disease	ALCA(anti-glycan antibody)	Li X et al., World J Gastroenterol. 2008 Sep 7;14(33):5115-24.
Irritable Bowel Disease	AMCA (anti-glycan antibody)	Li X et al., World J Gastroenterol. 2008 Sep 7;14(33):5115-24.
GIST	anti-DOG1 antibody	Espinosa F et al., Am J Surg Pathol Feb;32(2)210-8, 2008
Esophageal cancer	CaSR binding agent	Justinich CJ et al., Am J Physiol Gastrointest Liver Physiol. 2008 Jan;294(1):G120-9.
Gastric cancer	Calpain nCL-2 binding agent	Hata et al., J. Biol. Chem., Vol. 281, Issue 16, 11214-11224, April 21, 2006
Gastric cancer	drebrin binding agent	Keon BH et al., Journal of Cell Science, Vol 113, Issue 2 325-336

[0075] The binding agents can be used to detect the vesicles, such as for detecting cell-of-origin specific vesicles. A binding agent or multiple binding agents can themselves form a binding agent profile that provides a biosignature for a vesicle. One or more binding agents can be selected from **Table 3**. For example, if a vesicle population is detected or isolated using two, three, four or more binding agents in a differential detection or isolation of a vesicle from a heterogeneous population of vesicles, the particular binding agent profile for the vesicle population provides a biosignature for the particular vesicle population. The vesicle can be detected using any number of binding agents in a multiplex fashion. Thus, the binding agent can also be used to form a biosignature for a vesicle. The biosignature can be used to characterize a phenotype.

[0076] The binding agent can be a lectin. Lectins are proteins that bind selectively to polysaccharides and glycoproteins and are widely distributed in plants and animals. For example, lectins such as those derived from Galanthus nivalis in the form of Galanthus nivalis agglutinin ("GNA"), Narcissus pseudonarcissus in the form of Narcissus pseudonarcissus agglutinin ("NPA") and the blue green algae Nostoc ellipsosporum called "cyanovirin" (*Boyd et al. Antimicrob Agents Chemother 41(7): 1521 1530, 1997; Hammar et al. Ann N Y Acad Sci 724: 166 169, 1994; Kaku et al. Arch Biochem Biophys 279(2): 298 304, 1990*) can be used to isolate a vesicle. These lectins can bind to glycoproteins having a high mannose content (*Chervenak et al. Biochemistry 34(16): 5685 5695, 1995*). High mannose glycoprotein refers to glycoproteins having mannose-mannose linkages in the form of α-1→3 or α-1→6 mannose-mannose linkages.

[0077] The binding agent can be an agent that binds one or more lectins. Lectin capture can be applied to the isolation of the biomarker cathepsin D since it is a glycosylated protein capable of binding the lectins Galanthus nivalis agglutinin (GNA) and concanavalin A (ConA).

[0078] Methods and devices for using lectins to capture vesicles are described in International Patent Applications PCT/US2010/058461, entitled "METHODS AND SYSTEMS FOR ISOLATING, STORING, AND ANALYZING VESICLES" and filed November 30, 2010; PCT/US2009/066626, entitled "AFFINITY CAPTURE OF CIRCULATING BIOMARKERS" and filed December 3, 2009, and PCT/US2007/006101,

entitled "EXTRACORPOREAL REMOVAL OF MICROVESICULAR PARTICLES" and filed March 9, 2007, each of which applications is incorporated by reference herein in its entirety.

[0079] The binding agent can be an antibody. For example, a vesicle may be isolated using one or more antibodies specific for one or more antigens present on the vesicle. For example, a vesicle can have CD63 on its surface, and an antibody, or capture antibody, for CD63 can be used to isolate the vesicle. Alternatively, a vesicle derived from a tumor cell can express EpCam, the vesicle can be isolated using an antibody for EpCam and CD63. Other antibodies for isolating vesicles can include an antibody, or capture antibody, to CD9, PSCA, TNFR, CD63, B7H3, MFG-E8, EpCam, Rab, CD81, STEAP, PCSA, PSMA, or 5T4. Other antibodies for isolating vesicles can include an antibody, or capture antibody, to DR3, STEAP, epha2, TMEM211, unc93A, A33, CD24, NGAL, EpCam, MUC17, TROP2, TETS.

[0080] The antibodies disclosed herein can be immunoglobulin molecules or immunologically active portions of immunoglobulin molecules, i.e., molecules that contain an antigen binding site that specifically binds an antigen and synthetic antibodies. The immunoglobulin molecules can be of any class (e.g., IgG, IgE, IgM, IgD or IgA) or subclass of immunoglobulin molecule. Antibodies include, but are not limited to, polyclonal, monoclonal, bispecific, synthetic, humanized and chimeric antibodies, single chain antibodies, Fab fragments and $F(ab')_2$ fragments, Fv or Fv' portions, fragments produced by a Fab expression library, anti-idiotypic (anti-Id) antibodies, or epitope-binding fragments of any of the above. An antibody, or generally any molecule, "binds specifically" to an antigen (or other molecule) if the antibody binds preferentially to the antigen, and, e.g., has less than about 30%, 20%, 10%, 5% or 1% cross-reactivity with another molecule.

[0081] The binding agent can also be a polypeptide or peptide. Polypeptide is used in its broadest sense and may include a sequence of subunit amino acids, amino acid analogs, or peptidomimetics. The subunits may be linked by peptide bonds. The polypeptides may be naturally occurring, processed forms of naturally occurring polypeptides (such as by enzymatic digestion), chemically synthesized or recombinantly expressed. The polypeptides for use in the methods of the present invention may be chemically synthesized using standard techniques. The polypeptides may comprise D-amino acids (which are resistant to L- amino acid-specific proteases), a combination of D- and L-amino acids, β amino acids, or various other designer or non-naturally occurring amino acids (e.g., β-methyl amino acids, Cα- methyl amino acids, and Nα-methyl amino acids, etc.) to convey special properties. Synthetic amino acids may include ornithine for lysine, and norleucine for leucine or isoleucine. In addition, the polypeptides can have peptidomimetic bonds, such as ester bonds, to prepare polypeptides with novel properties. For example, a polypeptide may be generated that incorporates a reduced peptide bond, i.e., R_1-CH_2-NH-R_2, where R_1 and R_2 are amino acid residues or sequences. A reduced peptide bond may be introduced as a dipeptide subunit. Such a polypeptide would be resistant to protease activity, and would possess an extended half- live in vivo. Polypeptides can also include peptoids (N-substituted glycines), in which the side chains are appended to nitrogen atoms along the molecule's backbone, rather than to the α-carbons, as in amino acids. Polypeptides and peptides are intended to be used interchangeably throughout this application, i.e. where the term peptide is used, it may also include polypeptides and where the term polypeptides is used, it may also include peptides.

[0082] A vesicle may be isolated using a known binding agent. For example, the binding agent can be an agent that binds exosomal "housekeeping proteins," or general vesicle biomarkers, such as CD63, CD9, CD81, CD82, CD37, CD53, or Rab-5b. The binding agent can also be an agent that binds to vesicles derived from

specific cell types, such as tumor cells (e.g. binding agent for EpCam) or specific cell-of-origins, such as described below. For example, the binding agent used to isolate a vesicle may be a binding agent for an antigen selected from **Table 2**. The binding agent for a vesicle can also be selected from those listed in **Table 3**. For example, the binding agent can be for an antigen such as 5T4, B7H3, caveolin, CD63, CD9, E-Cadherin, MFG-E8, PSCA, PSMA, Rab-5B, STEAP, TNFR1, CD81, EpCam, CD59, or CD66. As another example, the binding agent can be for an antigen such as DR3, STEAP, epha2, TMEM211, unc93A, A33, CD24, NGAL, EpCam, MUC17, TROP2, TETS. One or more binding agents, such as one or more binding agents for two or more of the antigens, can be used for isolating a vesicle. The binding agent used can be selected based on the desire of isolating vesicles derived from particular cell types, or cell-of-origin specific vesicles.

[0083] A binding agent can also be linked directly or indirectly to a solid surface or substrate. A solid surface or substrate can be any physically separable solid to which a binding agent can be directly or indirectly attached including, but not limited to, surfaces provided by microarrays and wells, particles such as beads, columns, optical fibers, wipes, glass and modified or functionalized glass, quartz, mica, diazotized membranes (paper or nylon), polyformaldehyde, cellulose, cellulose acetate, paper, ceramics, metals, metalloids, semiconductive materials, quantum dots, coated beads or particles, other chromatographic materials, magnetic particles; plastics (including acrylics, polystyrene, copolymers of styrene or other materials, polypropylene, polyethylene, polybutylene, polyurethanes, TEFLON™, etc.), polysaccharides, nylon or nitrocellulose, resins, silica or silica-based materials including silicon and modified silicon, carbon, metals, inorganic glasses, plastics, ceramics, conducting polymers (including polymers such as polypyrole and polyindole); micro or nanostructured surfaces such as nucleic acid tiling arrays, nanotube, nanowire, or nanoparticulate decorated surfaces; or porous surfaces or gels such as methacrylates, acrylamides, sugar polymers, cellulose, silicates, or other fibrous or stranded polymers. In addition, as is known the art, the substrate may be coated using passive or chemically-derivatized coatings with any number of materials, including polymers, such as dextrans, acrylamides, gelatins or agarose. Such coatings can facilitate the use of the array with a biological sample.

[0084] For example, an antibody used to isolate a vesicle can be bound to a solid substrate such as a well, such as commercially available plates (e.g. from Nunc, Milan Italy). Each well can be coated with the antibody. In some embodiments, the antibody used to isolate a vesicle can be bound to a solid substrate such as an array. The array can have a predetermined spatial arrangement of molecule interactions, binding islands, biomolecules, zones, domains or spatial arrangements of binding islands or binding agents deposited within discrete boundaries. Further, the term array may be used herein to refer to multiple arrays arranged on a surface, such as would be the case where a surface bore multiple copies of an array. Such surfaces bearing multiple arrays may also be referred to as multiple arrays or repeating arrays.

[0085] A binding agent can also be bound to particles such as beads or microspheres. For example, an antibody specific for an exosomal component can be bound to a particle, and the antibody-bound particle is used to isolate vesicles from a biological sample. In some embodiments, the microspheres may be magnetic or fluorescently labeled. In addition, a binding agent for isolating vesicles can be a solid substrate itself. For example, latex beads, such as aldehyde/sulfate beads (Interfacial Dynamics, Portland, OR) can be used.

[0086] A binding agent bound to a magnetic bead can also be used to isolate a vesicle. For example, a biological sample such as serum from a patient can be collected for colon cancer screening. The sample can be incubated with anti-CCSA-3 (Colon Cancer–Specific Antigen) coupled to magnetic microbeads. A low-density

microcolumn can be placed in the magnetic field of a MACS Separator and the column is then washed with a buffer solution such as Tris-buffered saline. The magnetic immune complexes can then be applied to the column and unbound, non-specific material can be discarded. The CCSA-3 selected vesicles can be recovered by removing the column from the separator and placing it on a collection tube. A buffer can be added to the column and the magnetically labeled vesicles can be released by applying the plunger supplied with the column. The isolated vesicles can be diluted in IgG elution buffer and the complex can then be centrifuged to separate the microbeads from the vesicles. The pelleted isolated cell-of-origin specific vesicles can be resuspended in buffer such as phosphate-buffered saline and quantitated. Alternatively, due to the strong adhesion force between the antibody captured cell-of-origin specific vesicles and the magnetic microbeads, a proteolytic enzyme such as trypsin can be used for the release of captured vesicles without the need for centrifugation. The proteolytic enzyme can be incubated with the antibody captured cell-of-origin specific vesicles for at least a time sufficient to release the vesicles.

[0087] A binding agent, such as an antibody, for isolating a vesicle is preferably contacted with the biological sample comprising the vesicle of interest for at least a time sufficient for the binding agent to bind to an exosomal component. For example, an antibody may be contacted with a biological sample for various intervals ranging from seconds days, including but not limited to, about 10 minutes, 30 minutes, 1 hour, 3 hours, 5 hours, 7 hours, 10 hours, 15 hours, 1 day, 3 days, 7 days or 10 days.

[0088] A binding agent, such as an antibody specific to an antigen listed in **Table 2**, or a binding agent listed in **Table 3**, can be labeled with, including but not limited to, a magnetic label, a fluorescent moiety, an enzyme, a chemiluminescent probe, a metal particle, a non-metal colloidal particle, a polymeric dye particle, a pigment molecule, a pigment particle, an electrochemically active species, semiconductor nanocrystal or other nanoparticles including quantum dots or gold particles. The label can be, but not be limited to, fluorophores, quantum dots, or radioactive labels. For example, the label can be a radioisotope (radionuclides), such as ^{3}H, ^{11}C, ^{14}C, ^{18}F, ^{32}P, ^{35}S, ^{64}Cu, ^{68}Ga, ^{86}Y, ^{99}Tc, ^{111}In, ^{123}I, ^{124}I, ^{125}I, ^{131}I, ^{133}Xe, ^{177}Lu, ^{211}At, or ^{213}Bi. The label can be a fluorescent label, such as a rare earth chelate (europium chelate), fluorescein type, such as, but not limited to, FITC, 5-carboxyfluorescein, 6-carboxy fluorescein; a rhodamine type, such as, but not limited to, TAMRA; dansyl; Lissamine; cyanines; phycoerythrins; Texas Red; and analogs thereof.

[0089] A binding agent can be directly or indirectly labeled, e.g., the label is attached to the antibody through biotin-streptavidin. Alternatively, an antibody is not labeled, but is later contacted with a second antibody that is labeled after the first antibody is bound to an antigen of interest.

[0090] For example, various enzyme-substrate labels are available or disclosed (see for example, U.S. Pat. No. 4,275,149). The enzyme generally catalyzes a chemical alteration of a chromogenic substrate that can be measured using various techniques. For example, the enzyme may catalyze a color change in a substrate, which can be measured spectrophotometrically. Alternatively, the enzyme may alter the fluorescence or chemiluminescence of the substrate. Examples of enzymatic labels include luciferases (e.g., firefly luciferase and bacterial luciferase; U.S. Pat. No. 4,737,456), luciferin, 2,3-dihydrophthalazinediones, malate dehydrogenase, urease, peroxidase such as horseradish peroxidase (HRP), alkaline phosphatase (AP), β-galactosidase, glucoamylase, lysozyme, saccharide oxidases (e.g., glucose oxidase, galactose oxidase, and glucose-6-phosphate dehydrogenase), heterocyclic oxidases (such as uricase and xanthine oxidase), lactoperoxidase, microperoxidase, and the like. Examples of enzyme-substrate combinations include, but are

not limited to, horseradish peroxidase (HRP) with hydrogen peroxidase as a substrate, wherein the hydrogen peroxidase oxidizes a dye precursor (e.g., orthophenylene diamine (OPD) or 3,3',5,5'-tetramethylbenzidine hydrochloride (TMB)); alkaline phosphatase (AP) with para-nitrophenyl phosphate as chromogenic substrate; and β-D-galactosidase (β-D-Gal) with a chromogenic substrate (e.g., p-nitrophenyl- β-D-galactosidase) or fluorogenic substrate 4-methylumbelliferyl-β-D-galactosidase.

[0091] Depending on the method of isolation used, the binding agent may be linked to a solid surface or substrate, such as arrays, particles, wells and other substrates described above. Methods for direct chemical coupling of antibodies, to the cell surface are known in the art, and may include, for example, coupling using glutaraldehyde or maleimide activated antibodies. Methods for chemical coupling using multiple step procedures include biotinylation, coupling of trinitrophenol (TNP) or digoxigenin using for example succinimide esters of these compounds. Biotinylation can be accomplished by, for example, the use of D-biotinyl-N-hydroxysuccinimide. Succinimide groups react effectively with amino groups at pH values above 7, and preferentially between about pH 8.0 and about pH 8.5. Biotinylation can be accomplished by, for example, treating the cells with dithiothreitol followed by the addition of biotin maleimide.

Flow Cytometry

[0092] Isolation of vesicles using particles such as beads or microspheres can also be performed using flow cytometry. Flow cytometry can be used for sorting microscopic particles suspended in a stream of fluid. As particles pass through they can be selectively charged and on their exit can be deflected into separate paths of flow. It is therefore possible to separate populations from an original mix, such as a biological sample, with a high degree of accuracy and speed. Flow cytometry allows simultaneous multiparametric analysis of the physical and/or chemical characteristics of single cells flowing through an optical/electronic detection apparatus. A beam of light, usually laser light, of a single frequency (color) is directed onto a hydrodynamically focused stream of fluid. A number of detectors are aimed at the point where the stream passes through the light beam; one in line with the light beam (Forward Scatter or FSC) and several perpendicular to it (Side Scatter or SSC) and one or more fluorescent detectors.

[0093] Each suspended particle passing through the beam scatters the light in some way, and fluorescent chemicals in the particle may be excited into emitting light at a lower frequency than the light source. This combination of scattered and fluorescent light is picked up by the detectors, and by analyzing fluctuations in brightness at each detector (one for each fluorescent emission peak), it is possible to deduce various facts about the physical and chemical structure of each individual particle. FSC correlates with the cell size and SSC depends on the inner complexity of the particle, such as shape of the nucleus, the amount and type of cytoplasmic granules or the membrane roughness. Some flow cytometers have eliminated the need for fluorescence and use only light scatter for measurement.

[0094] Flow cytometers can analyze several thousand particles every second in "real time" and can actively separate out and isolate particles having specified properties. They offer high-throughput automated quantification, and separation, of the set parameters for a high number of single cells during each analysis session. Modern instruments have multiple lasers and fluorescence detectors, for example up to 4 lasers and 18 fluorescence detectors, allowing multiple labels to be used to more precisely specify a target population by their phenotype.

[0095] The data resulting from flow-cytometers can be plotted in 1 dimension to produce histograms or seen in 2 dimensions as dot plots or in 3 dimensions with newer software. The regions on these plots can be sequentially separated by a series of subset extractions which are termed gates. Specific gating protocols exist for diagnostic and clinical purposes especially in relation to hematology. The plots are often made on logarithmic scales. Because different fluorescent dye's emission spectra overlap, signals at the detectors have to be compensated electronically as well as computationally. Fluorophores for labeling biomarkers may include those described in *Ormerod, Flow Cytometry 2nd ed., Springer-Verlag, New York (1999),* and in *Nida et al., Gynecologic Oncology 2005;4 889-894* which is incorporated herein by reference.

Multiplexing

[0096] Different binding agents can be used for multiplexing different vesicle populations. Different vesicle populations can be isolated or detected using different binding agents. Each vesicle population in a biological sample can be labeled with a different signaling label, such as fluorophores, quantum dots, or radioactive labels, such as described above. The label can be directly conjugated to a binding agent or indirectly used to detect a binding agent. The number of populations detected in a multiplexing assay is dependent on the resolution capability of the labels and the summation of signals, as more than two differentially labeled vesicle populations that bind two or more affinity elements can produce summed signals.

[0097] Multiplexing of at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 50, 75 or 100 different vesicle populations may be performed. For example, one population of vesicles specific to a cell-of-origin can be assayed along with a second population of vesicles specific to a different cell-of-origin, where each population is labeled with a different label. Alternatively, a population of vesicles with a particular biomarker or biosignature can be assayed along with a second population of vesicles with a different biomarker or biosignature.

[0098] In one embodiment, multiplex analysis is performed by applying a plurality of vesicles comprising more than one population of vesicles to a plurality of substrates, such as beads. Each bead is coupled to one or more capture agents. The plurality of beads is divided into subsets, where beads with the same capture agent or combination of capture agents form a subset of beads, such that each subset of beads has a different capture agent or combination of capture agents than another subset of beads. The beads can then be used to capture vesicles that comprise a component that binds to the capture agent. The different subsets can be used to capture different populations of vesicles. The captured vesicles can then be analyzed by detecting one or more biomarkers of the vesicles.

[0099] Flow cytometry can be used in combination with a particle-based or bead based assay. Multiparametric immunoassays or other high throughput detection assays using bead coatings with cognate ligands and reporter molecules with specific activities consistent with high sensitivity automation can be used. For example, beads in each subset can be differentially labeled from another subset. For example, in a particle based assay system, a binding agent or capture agent for a vesicle, such as a capture antibody, can be immobilized on addressable beads or microspheres. Each binding agent for each individual binding assay (such as an immunoassay when the binding agent is an antibody) can be coupled to a distinct type of microsphere (i.e., microbead) and the binding assay reaction takes place on the surface of the microspheres. Microspheres can be distinguished by different labels, for example, a microsphere with a specific capture agent would have a different signaling label as compared to another microsphere with a different capture agent. For example,

microspheres can be dyed with discrete fluorescence intensities such that the fluorescence intensity of a microsphere with a specific binding agent is different than that of another microsphere with a different binding agent.

[00100] The microsphere can be labeled or dyed with at least 2 different labels or dyes. In some embodiments, the microsphere is labeled with at least 3, 4, 5, 6, 7, 8, 9, or 10 different labels. Different microspheres in a plurality of microspheres can have more than one label or dye, wherein various subsets of the microspheres have various ratios and combinations of the labels or dyes permitting detection of different microspheres with different binding agents. For example, the various ratios and combinations of labels and dyes can permit different fluorescent intensities. Alternatively, the various ratios and combinations maybe used to generate different detection patters to identify the binding agent. The microspheres can be labeled or dyed externally or may have intrinsic fluorescence or signaling labels. Beads can be loaded separately with their appropriate binding agents and thus, different vesicle populations can be isolated based on the different binding agents on the differentially labeled microspheres to which the different binding agents are coupled.

[00101] In another embodiment, multiplex analysis can be performed using a planar substrate, wherein the said substrate comprises a plurality of capture agents. The plurality of capture agents can capture one or more populations of vesicles, and one or more biomarkers of the captured vesicles detected. The planar substrate can be a microarray or other substrate as further described herein.

Microfluidics

[00102] The methods for isolating or identifying vesicles can be used in combination with microfluidic devices. The methods of isolating vesicles disclosed herein can be performed using microfluidic devices. Microfluidic devices, which may also be referred to as "lab-on-a-chip" systems, biomedical micro-electro-mechanical systems (bioMEMs), or multicomponent integrated systems, can be used for isolating, and analyzing, vesicles. Such systems miniaturize and compartmentalize processes that allow for binding of vesicles, detection of exosomal biomarkers, and other processes.

[00103] A microfluidic device can also be used for isolation of a vesicle through size differential or affinity selection. For example, a microfluidic device can use one more channels for isolating a vesicle from a biological sample based on size, or by using one or more binding agents for isolating a vesicle, from a biological sample. A biological sample can be introduced into one or more microfluidic channels, which selectively allows the passage of vesicles. The selection can be based on a property of the vesicles, for example, size, shape, deformability, biomarker profile, or biosignature.

[00104] Alternatively, a heterogeneous population of vesicles can be introduced into a microfluidic device, and one or more different homogeneous populations of vesicles can be obtained. For example, different channels can have different size selections or binding agents to select for different vesicle populations. Thus, a microfluidic device can isolate a plurality of vesicles, wherein at least a subset of the plurality of vesicles comprises a different biosignature from another subset of said plurality of vesicles. For example, the microfluidic device can isolate at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 50, 60, 70, 80, 90, or 100 different subsets of vesicles, wherein each subset of vesicles comprises a different biosignature.

[00105] In some embodiments, the microfluidic device can comprise one or more channels that permit further enrichment or selection of vesicles. A population of vesicles that has been enriched after passage through a first

channel can be introduced into a second channel, which allows the passage of the desired vesicle population to be further enriched, such as through binding agents present in the second channel.

[00106] Array-based assays and bead-based assays can be used with microfluidic device. For example, the binding agent can be coupled to beads and the binding reaction between the beads and vesicles can be performed in a microfluidic device. Multiplexing can also be performed using a microfluidic device. Different compartments can comprise different binding agents for different populations of vesicles, where each population is of a different cell-of-origin specific vesicle population or each population has a different biosignature. The hybridization reaction between the microspheres and vesicles can be performed in a microfluidic device and the reaction mixture can be delivered to a detection device. The detection device, such as a dual or multiple laser detection system can be part of the microfluidic system and can use a laser to identify each bead or microsphere by its color-coding, and another laser can detect the hybridization signal associated with each bead.

[00107] Examples of microfluidic devices that may be used, or adapted for use with vesicles, include but are not limited to those described in U.S. Pat. Nos. 7,591,936, 7,581,429, 7,579,136, 7,575,722, 7,568,399, 7,552,741, 7,544,506, 7,541,578, 7,518,726, 7,488,596, 7,485,214, 7,467,928, 7,452,713, 7,452,509, 7,449,096, 7,431,887, 7,422,725, 7,422,669, 7,419,822, 7,419,639, 7,413,709, 7,411,184, 7,402,229, 7,390,463, 7,381,471, 7,357,864, 7,351,592, 7,351,380, 7,338,637, 7,329,391, 7,323,140, 7,261,824, 7,258,837, 7,253,003, 7,238,324, 7,238,255, 7,233,865, 7,229,538, 7,201,881, 7,195,986, 7,189,581, 7,189,580, 7,189,368, 7,141,978, 7,138,062, 7,135,147, 7,125,711, 7,118,910, and 7,118,661.

[00108] With methods of detecting isolated vesicles as described here, e.g., antibody affinity isolation, the consistency of the results can be optimized as necessary using various concentration or isolation procedures. Such steps can include agitation such as shaking or vortexing, different isolation techniques such as polymer based isolation, e.g., with PEG, and concentration to different levels during filtration or other steps. It will be understood by those in the art that such treatments can be applied at various stages of testing the vesicle containing sample. In one embodiment, the sample itself, e.g., a bodily fluid such as plasma or serum, is vortexed. In some embodiments, the sample is vortexed after one or more sample treatment step, e.g., vesicle isolation, has occurred. Agitation can occur at some or all appropriate sample treatment steps as desired.

[00109] The results can also be optimized as desireable by treating the vesicle-containing sample with various agents. Such agents include additives to control aggregation and/or additives to adjust pH or ionic strength. Additives that control aggregation include blocking agents such as bovine serum albumen (BSA) and milk, chaotropic agents such as guanidium hydro chloride, and detergents or surfactants. Useful ionic detergents include sodium dodecyl sulfate (SDS, sodium lauryl sulfate (SLS)), sodium laureth sulfate (SLS, sodium lauryl ether sulfate (SLES)), ammonium lauryl sulfate (ALS), cetrimonium bromide, cetrimonium chloride, cetrimonium stearate, and the like. Useful non-ionic (zwitterionic) detergents include polyoxyethylene glycols, polysorbate 20 (also known as Tween 20), other polysorbates (e.g., 40, 60, 65, 80, etc), Triton-X (e.g., X100, X114), 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS), CHAPSO, deoxycholic acid, sodium deoxycholate, NP-40, glycosides, octyl-thio-glucosides, maltosides, and the like. In some embodiments, Pluronic F-68, a surfactant shown to reduce platelet aggregation, is used to treat samples containing vesicles during isolation and/or detection. F68 can be used from a 0.1% to 10% concentration, e.g., a 1%, 2.5% or 5% concentration. The pH and/or ionic strength of the solution can be adjusted with various acids, bases, buffers or salts, including without limitation sodium chloride (NaCl), phosphate-buffered saline (PBS), tris-buffered saline

(TBS), sodium phosphate, potassium chloride, potassium phosphate, sodium citrate and saline-sodium citrate (SSC) buffer. In some embodiments, NaCl is added at a concentration of 0.1% to 10%, e.g., 1%, 2.5% or 5% final concentration. In some embodiments, Tween 20 is added to 0.005 to 2% concentration, e.g., 0.05%, 0.25% or 0.5 % final concentration. In some embodiments, BSA is added to 0.1% to 10% concentration, e.g., 3%, 3.5% or 7% concentration. In some embodiments, SSC/detergent (e.g., 20X SSC with 0.5% Tween 20 or 0.1% Triton-X 100) is added to 0.1% to 10% concentration, e.g., at 1.0% or 5.0% concentration.

[00110] It will be understood that the methods of detecting vesicles can be optimized as desired with various combinations of protocols and treatments as described herein. A detection protocol can be optimized by various combinations of agitation, isolation methods, and additives. In some embodiments, the patient sample is vortexed before and after isolation steps, and the sample is treated with blocking agents including BSA and F68. Such treatments may reduce the formation of large aggregates or protein or other biological debris and thus provide a more consistent detection reading.

Substrates

[00111] A binding agent can be linked directly or indirectly to a solid surface or substrate. A solid surface or substrate includes physically separable solids to which a binding agent can be directly or indirectly attached. These surfaces or substrates include without limitation surfaces provided by microarrays, wells, particles such as beads, columns, optical fibers, wipes, glass and modified or functionalized glass, quartz, mica, diazotized membranes (paper or nylon), polyformaldehyde, cellulose, cellulose acetate, paper, ceramics, metals, metalloids, semiconductive materials, quantum dots, coated beads or particles, other chromatographic materials, magnetic particles; plastics (including acrylics, polystyrene, copolymers of styrene or other materials, polypropylene, polyethylene, polybutylene, polyurethanes, TEFLON™, etc.), polysaccharides, nylon or nitrocellulose, resins, silica or silica-based materials including silicon and modified silicon, carbon, metals, inorganic glasses, plastics, ceramics, conducting polymers (including polymers such as polypyrole and polyindole); micro or nanostructured surfaces such as nucleic acid tiling arrays, nanotube, nanowire, or nanoparticulate decorated surfaces; or porous surfaces or gels such as methacrylates, acrylamides, sugar polymers, cellulose, silicates, or other fibrous or stranded polymers. In addition, as is known the art, the substrate may be coated using passive or chemically-derivatized coatings with any number of materials, including polymers, such as dextrans, acrylamides, gelatins or agarose. Such coatings can facilitate the use of the substrate with a biological sample.

[00112] In some embodiments, an antibody used to isolate a vesicle is bound to a solid substrate of a well, such as a well of a commercially available plate (e.g. from Nunc, Milan Italy). Such plates are known in the art, e.g., 96 and 384 well plates. Each well can be coated with an antibody. In some embodiments, the antibody used to isolate a vesicle is bound to a solid substrate in an array. The array can have a predetermined spatial arrangement of molecule interactions, binding islands, biomolecules, zones, domains or spatial arrangements of binding islands or binding agents deposited within discrete boundaries. The term array may be used herein to refer to multiple arrays arranged on a surface, such as would be the case with a surface bearing multiple copies of an array. Such surfaces bearing multiple arrays may also be referred to as multiple arrays or repeating arrays.

[00113] Arrays typically contain addressable moieties that can detect the presense of an entity, e.g., a vesicle in the sample via a binding event. An array may be referred to as a microarray. Arrays or microarrays include without limitation DNA microarrays, such as cDNA microarrays, oligonucleotide microarrays and SNP microarrays, microRNA arrays, protein microarrays, antibody microarrays, tissue microarrays, cellular

microarrays (also called transfection microarrays), chemical compound microarrays, and carbohydrate arrays (glycoarrays). DNA arrays typically comprise addressable nucleotide sequences that can bind to sequences present in a sample. MicroRNA arrays, e.g., the MMChips array from the University of Louisville or commercial systems from Agilent, can be used to detect microRNAs. Protein microarrays can be used to identify protein–protein interactions, including without limitation identifying substrates of protein kinases, transcription factor protein-activation, or to identify the targets of biologically active small molecules. Protein arrays may comprise an array of different protein molecules, commonly antibodies, or nucleotide sequences that bind to proteins of interest. In a non-limiting example, a protein array can be used to detect vesicles having certain proteins on their surface. Antibody arrays comprise antibodies spotted onto the protein chip that are used as capture molecules to detect proteins or other biological materials from a sample, e.g., from cell or tissue lysate solutions. For example, antibody arrays can be used to detect vesicle-associated biomarkers from bodily fluids, e.g., serum or urine. Tissue microarrays comprise separate tissue cores assembled in array fashion to allow multiplex histological analysis. Cellular microarrays, also called transfection microarrays, comprise various capture agents, such as antibodies, proteins, or lipids, which can interact with cells to facilitate their capture on addressable locations. Cellular arrays can also be used to capture vesicles due to the similarity between a vesicle and cellular membrane. Chemical compound microarrays comprise arrays of chemical compounds and can be used to detect protein or other biological materials that bind the compounds. Carbohydrate arrays (glycoarrays) comprise arrays of carbohydrates and can detect, e.g., protein that bind sugar moieties. One of skill will appreciate that similar technologies or improvements can be used according to the methods of the invention.

[00114] A binding agent can also be bound to particles such as beads or microspheres. For example, an antibody specific for a component of a vesicle can be bound to a particle, and the antibody-bound particle is used to isolate a vesicle from a biological sample. In some embodiments, the microspheres may be magnetic or fluorescently labeled. In addition, a binding agent for isolating vesicles can be a solid substrate itself. In some embodiments, latex beads, such as aldehyde/sulfate beads (Interfacial Dynamics, Portland, OR) are used.

[00115] Binding agents bound to magnetic beads can be used to isolate a vesicle. In a non-limiting example, consider that a biological sample such as serum from a patient is collected for prostate cancer screening. The sample can be incubated with anti-PSMA or anti-PCSA coupled to magnetic microbeads and isolated, thereby capturing vesicles of prostate epithelial cell origin. In an embodiment, a low-density microcolumn can be placed in the magnetic field of a MACS Separator and the column is then washed with a buffer solution such as Tris-buffered saline. The magnetic immune complexes can then be applied to the column and unbound, non-specific material discarded. The PSMA or PCSA selected vesicle can be recovered by removing the column from the separator and placing it on a collection tube. A buffer can be added to the column and the magnetically labeled vesicle can be released by applying the plunger supplied with the column. The isolated vesicle can be diluted in IgG elution buffer and the complex can then be centrifuged to separate the microbeads from the vesicle. The pelleted isolated cell-of-origin specific vesicle can be resuspended in buffer such as phosphate-buffered saline and quantitated. Alternatively, due to the strong adhesion force between the antibody captured cell-of-origin specific vesicle and the magnetic microbeads, a proteolytic enzyme such as trypsin can be used for the release of captured vesicles without the need for centrifugation. The proteolytic enzyme can be incubated with the antibody captured cell-of-origin specific vesicles for at least a time sufficient to release the vesicles.

One of skill will appreciate that this approach can be applied to isolating other specific vesicles by using binding agents that recognize different biomarkers.

[00116] A binding agent attached directly or indirectly to a solid surface or substrate can be used to capture a vesicle. The capture vesicle can be released from the substrate and analyzed or subjected to further isolation or concentration methods. Alternatively, the captured vesicle can be analyzed while still attached to the substrate.

[00117] A binding agent, such as an antibody specific to an antigen listed in **Table 2**, or a binding agent listed in **Table 3**, can be labeled to allow for its detection. Appropriate labels include without limitation a magnetic label, a fluorescent moiety, an enzyme, a chemiluminescent probe, a metal particle, a non-metal colloidal particle, a polymeric dye particle, a pigment molecule, a pigment particle, an electrochemically active species, semiconductor nanocrystal or other nanoparticles including quantum dots or gold particles, fluorophores, quantum dots, or radioactive labels. Protein labels include green fluorescent protein (GFP) and variants thereof (e.g., cyan fluorescent protein and yellow fluorescent protein); and luminescent proteins such as luciferase, as described below. Radioactive labels include without limitation radioisotopes (radionuclides), such as ^{3}H, ^{11}C, ^{14}C, ^{18}F, ^{32}P, ^{35}S, ^{64}Cu, ^{68}Ga, ^{86}Y, ^{99}Tc, ^{111}In, ^{123}I, ^{124}I, ^{125}I, ^{131}I, ^{133}Xe, ^{177}Lu, ^{211}At, or ^{213}Bi. Fluorescent labels include without limitation a rare earth chelate (e.g., europium chelate), rhodamine; fluorescein types including without limitation FITC, 5-carboxyfluorescein, 6-carboxy fluorescein; a rhodamine type including without limitation TAMRA; dansyl; Lissamine; cyanines; phycoerythrins; Texas Red; Cy3, Cy5, dapoxyl, NBD, Cascade Yellow, dansyl, PyMPO, pyrene, 7-diethylaminocoumarin-3-carboxylic acid and other coumarin derivatives, Marina Blue™, Pacific Blue™, Cascade Blue™, 2-anthracenesulfonyl, PyMPO, 3,4,9,10-perylene-tetracarboxylic acid, 2,7-difluorofluorescein (Oregon Green™ 488-X), 5-carboxyfluorescein, Texas Red™-X, Alexa Fluor 430, 5-carboxytetramethylrhodamine (5-TAMRA), 6-carboxytetramethylrhodamine (6-TAMRA), BODIPY FL, bimane, and Alexa Fluor 350, 405, 488, 500, 514, 532, 546, 555, 568, 594, 610, 633, 647, 660, 680, 700, and 750, and derivatives thereof, among many others. See, e.g., "The Handbook--A Guide to Fluorescent Probes and Labeling Technologies," Tenth Edition, available at probes.invitrogen.com/handbook.

[00118] A binding agent can be labeled directly, e.g., via a covalent bond. Binding agents can also be indirectly labeled, such as when a label is attached to the binding agent through a binding system. In a non-limiting example, consider an antibody labeled through biotin-streptavidin. Alternatively, an antibody is not labeled, but is later contacted with a second antibody that is labeled after the first antibody is bound to an antigen of interest. For example, various enzyme-substrate labels are available or disclosed (see for example, U.S. Pat. No. 4,275,149). The enzyme generally catalyzes a chemical alteration of a chromogenic substrate that can be measured using various techniques. For example, the enzyme may catalyze a color change in a substrate, which can be measured spectrophotometrically. Alternatively, the enzyme may alter the fluorescence or chemiluminescence of the substrate. Examples of enzymatic labels include luciferases (e.g., firefly luciferase and bacterial luciferase; U.S. Pat. No. 4,737,456), luciferin, 2,3-dihydrophthalazinediones, malate dehydrogenase, urease, peroxidase such as horseradish peroxidase (HRP), alkaline phosphatase (AP), β-galactosidase, glucoamylase, lysozyme, saccharide oxidases (e.g., glucose oxidase, galactose oxidase, and glucose-6-phosphate dehydrogenase), heterocyclic oxidases (such as uricase and xanthine oxidase), lactoperoxidase, microperoxidase, and the like. Examples of enzyme-substrate combinations include without limitation horseradish peroxidase (HRP) with hydrogen peroxidase as a substrate, wherein the hydrogen peroxidase oxidizes a dye precursor (e.g., orthophenylene diamine (OPD) or 3,3',5,5'-tetramethylbenzidine

hydrochloride (TMB)); alkaline phosphatase (AP) with para-nitrophenyl phosphate as chromogenic substrate; and β-D-galactosidase (β-D-Gal) with a chromogenic substrate (e.g., p-nitrophenyl- β-D-galactosidase) or fluorogenic substrate 4-methylumbelliferyl-β-D-galactosidase.

[00119] A binding agent can be bound to particles such as beads or microspheres. For example, an antibody specific for a vesicle component can be bound to a particle, and the antibody-bound particle is used to isolate vesicles from a biological sample. In some embodiments, the microspheres may be magnetic or fluorescently labeled, such as described herein. The microspheres may be magnetic or fluorescently labeled. The microspheres or nanospheres may comprise plastics (including acrylics, polystyrene, copolymers of styrene or other materials, polypropylene, polyethylene, polybutylene, polyurethanes, TEFLON™, etc.), polysaccharides, nylon or nitrocellulose, resins, silica or silica-based materials including silicon and modified silicon, carbon, metals, inorganic glasses, plastics, ceramics, conducting polymers (including polymers such as polypyrole and polyindole). The particle may be intrinsically or extrinsically labeled. For example, the particle may be intrinsically dyed or contain a metal core, such as gold or silver core, such as commercially available from Luminex (Austin, TX) or Oxonica, Inc. (Mountain View, CA). Other labels are described herein.

[00120] The binding agent may be linked to a solid surface or substrate, such as arrays, particles, wells and other substrates described above. Methods for direct chemical coupling of antibodies, to the cell surface are known in the art, and may include, for example, coupling using glutaraldehyde or maleimide activated antibodies. Methods for chemical coupling using multiple step procedures include biotinylation, coupling of trinitrophenol (TNP) or digoxigenin using for example succinimide esters of these compounds. Biotinylation can be accomplished by, for example, the use of D-biotinyl-N-hydroxysuccinimide. Succinimide groups react effectively with amino groups at pH values above 7, and preferentially between about pH 8.0 and about pH 8.5. Biotinylation can be accomplished by, for example, treating the cells with dithiothreitol followed by the addition of biotin maleimide.

Characterizing a Phenotype

[00121] In an aspect of the invention, a phenotype of a subject is characterized by analyzing a biological sample and determining the presence, level, amount, or concentration of one or more populations of microRNA in the sample. In embodiments, characterization includes determining an absolute presence or absence, a quantitative level, or a relative level compared to a reference, e.g., the level of all RNA or microRNA present, the level of a housekeeping marker, and/or the level of a spiked-in marker.

[00122] In some embodiments, vesicles are purified or concentrated from a sample prior to determining the amount of microRNA in the vesicles. In other embodiments, vesicles are directly assessed from a sample, without prior purification or concentration. The vesicles can be cell-of-origin specific vesicles or vesicles with a specific biosignature. Biosignature include specific pattern of biomarkers, e.g., patterns of biomarkers indicative of a phenotype that is desirable to detect, such as a disease phenotype. The detected amount of microRNA and/or vesicles can be used when characterizing a phenotype, such as a diagnosis, prognosis, theranosis, or prediction of responder / non-responder status. In some embodiments, the detected amount is used to determine a physiological or biological state, such as pregnancy or the stage of pregnancy. The detected amount of microRNA, including that in vesicle payload, can also be used to determine treatment efficacy, stage of a disease or condition, or progression of a disease or condition. For example, the amount of one or more microRNA and/or vesicles can be proportional or inversely proportional to an increase in disease stage or

progression. The detected amount of microRNA and/or vesicles can also be used to monitor progression of a disease or condition or to monitor a subject's response to a treatment.

[00123] The microRNA can be evaluated by comparing the level of microRNA with a reference level or value of microRNA. The reference value can be particular to physical or temporal endpoint. For example, the reference value can be from the same subject from whom a sample is assessed, or the reference value can be from a representative population of samples, e.g., samples from normal subjects without a disease. Therefore, a reference value provides a threshold measurement that can be compared to the readout for a microRNA population assayed in a given sample. Such reference values may be set according to data pooled from groups of sample corresponding to a particular cohort, including but not limited to age (e.g., newborns, infants, adolescents, young, middle-aged adults, seniors and adults of varied ages), racial/ethnic groups, normal versus diseased subjects, smoker v. non-smoker, subjects receiving therapy versus untreated subjects, different time points of treatment for a particular individual or group of subjects similarly diagnosed or treated or combinations thereof. Determining microRNA levels at different time points of treatment for a particular individual can provide a method for monitoring the individual's response to the treatment or progression of a disease or condition for which the individual is being treated.

[00124] A reference value may be based on samples assessed from the same subject so to provide individualized tracking. In some embodiments, frequent testing of microRNA in samples from a subject provides better comparisons to the reference values previously established for that subject. Such time course measurements are used to allow a physician to more accurately assess the subject's disease stage or progression and therefore inform a better decision for treatment. In some cases, the variance of microRNA levels is reduced when comparing a subject's own microRNA levels over time, thus allowing an individualized threshold to be defined for the subject, e.g., a threshold at which a diagnosis is made. Temporal intrasubject variation allows each individual to serve as their own longitudinal control for optimum analysis of disease or physiological state. As an illustrative example, consider that the level of microRNA derived from colorectal cells is measured in a subject's blood over time. A spike in the level of colon-derived microRNA in the subject's blood can indicate hyperproliferation of colon cells, e.g., due to colorectal cancer.

[00125] In some embodiments, reference values are established for unaffected individuals of varying ages, ethnic backgrounds and sexes by determining the amount of microRNA of interest in the unaffected individuals. The reference value for a reference population can be used as a baseline for detection of one or more microRNA populations in a test subject. If a sample from a subject has a level or value that is similar to the reference, the subject might be determined to not have the disease, or of having a low risk of developing a disease.

[00126] In other embodiments, reference values or levels are established for individuals with a particular phenotype by determining the amount of one or more populations of microRNA in an individual with the phenotype, e.g., a disease or a response to therapy. In an embodiment, an index of values is generated for a particular phenotype. Different disease stages can have different values, determined from individuals with the different disease stages. A subject's value can be compared to the index and a diagnosis, prognosis or theranosis of the disease can be determined, e.g., the disease stage or progression wherein the subject's levels most closely correlate with the index. In other embodiments, an index of values is generated for therapeutic efficacies. For example, the level of microRNA of individuals with a particular disease can be generated and correlated with treatments that were effective for the individual. The levels can be used to generate values of

which is a subject's value is compared, and a treatment or therapy can be selected for the individual, e.g., by predicting from the levels whether the subject is likely to be a responder or non-responder for a treatment.

[00127] In some embodiments, a reference value is determined for individuals without a phenotype, by detecting microRNA linked to the phenotype. As a non-limiting example, individuals with varying stages of colorectal cancer and noncancerous polyps can be surveyed using the same techniques described for unaffected individuals and the levels of circulating microRNA for each group can be determined. In some embodiments, the levels are defined as means ± standard deviations from at least two separate experiments performed in at least triplicate. Comparisons between these groups can be made using statistical tests to determine statistical significance of distinguishing microRNA biosignatures. In some embodiments, statistical significance is determined using a parametric statistical test. The parametric statistical test can comprise, without limitation, a fractional factorial design, analysis of variance (ANOVA), a t-test, least squares, a Pearson correlation, simple linear regression, nonlinear regression, multiple linear regression, or multiple nonlinear regression. Alternatively, the parametric statistical test can comprise a one-way analysis of variance, two-way analysis of variance, or repeated measures analysis of variance. In other embodiments, statistical significance is determined using a nonparametric statistical test. Examples include, but are not limited to, a Wilcoxon signed-rank test, a Mann-Whitney test, a Kruskal-Wallis test, a Friedman test, a Spearman ranked order correlation coefficient, a Kendall Tau analysis, and a nonparametric regression test. In some embodiments, statistical significance is determined at a p-value of less than 0.05, 0.01, 0.005, 0.001, 0.0005, or 0.0001. The p-values can also be corrected for multiple comparisons, e.g., using a Bonferroni correction, a modification thereof, or other technique known to those in the art, e.g., the Hochberg correction, Holm-Bonferroni correction, Šidák correction, Dunnett's correction or Tukey's multiple comparisons. In some embodiments, an ANOVA is followed by Tukey's correction for post-test comparing of the biomarkers from each population.

[00128] Reference values can also be established for disease recurrence monitoring (or exacerbation phase in MS), for therapeutic response monitoring, or for predicting responder / non-responder status.

[00129] In some embodiments, a reference value for microRNA obtained from vesicles is determined using an artificial vesicle, also referred to herein as a synthetic vesicle. Methods for manufacturing artificial vesicles are known to those of skill in the art, e.g., using liposomes. Artificial exosomes can be manufactured using methods disclosed in US20060222654 and US4448765, which are incorporated herein by reference in its entirety. Artificial vesicles can be constructed with known markers to facilitate capture and/or detection. In some embodiments, artificial vesicles are spiked into a bodily sample prior to processing. The level of intact synthetic vesicle can be tracked during processing, e.g., using filtration or other isolation methods disclosed herein, to provide a control for the amount of vesicles in the initial versus processed sample. Similarly, artificial vesicles can be spiked into a sample before or after any processing steps. In some embodiments, artificial vesicles are used to calibrate equipment used for isolation and detection of vesicles.

[00130] Artificial vesicles can be produced and used a control to test the viability of an assay, such as a bead-based assay. The artificial vesicle can bind to both the beads and to the detection antibodies. Thus, the artificial vesicle contains the amino acid sequence/conformation that each of the antibodies binds. The artificial vesicle can comprise a purified protein or a synthetic peptide sequence to which the antibody binds. The artificial vesicle could be a bead, e.g., a polystyrene bead, that is capable of having biological molecules attached thereto.

If the bead has an available carboxyl group, then the protein or peptide could be attached to the bead via an available amine group, such as using carbodiimide coupling.

[00131] In another embodiment, the artificial vesicle can be a polystyrene bead coated with avidin and a biotin is placed on the protein or peptide of choice either at the time of synthesis or via a biotin-maleimide chemistry. The proteins/peptides to be on the bead can be mixed together in ratio specific to the application the artificial vesicle is being used for, and then conjugated to the bead. These artificial vesicles can then serve as a link between the capture beads and the detection antibodies, thereby providing a control to show that the components of the assay are working properly.

[00132] The reference value can be a quantitative or qualitative value. The value can be a direct measurement of the level of microRNA (example, copies / ml), or an indirect measure, such as the amount of a specific biomarker associated with a vesicle. The value can be a quantitative, such as a numerical value. In other embodiments, the value is qualitative, such as no microRNA, low level of microRNA, medium level, high level of microRNA, or variations thereof.

[00133] The reference value can be stored in a database and used as a reference for the diagnosis, prognosis, theranosis, disease stratification, disease monitoring, treatment monitoring or prediction of non-responder / responder status of a disease or condition based on the level or amount of microRNA, such as total amount of microRNA, or the amount of a specific population of microRNA, such as cell-of-origin specific microRNA or microRNA from vesicles with a specific biosignature. In an illustrative example, consider a method of determining a diagnosis for a cancer. MicroRNA from reference subjects with and without the cancer are assessed and stored in the database. The reference subjects provide biosignature indicative of the cancer or of another state, e.g., a healthy state. A sample from a test subject is then assayed and the microRNA biosignature is compared against those in the database. If the subject's biosignature correlates more closely with reference values indicative of cancer, a diagnosis of cancer may be made. Conversely, if the subject's biosignature correlates more closely with reference values indicative of a healthy state, the subject may be determined to not have the disease. One of skill will appreciate that this example is non-limiting and can be expanded for assessing other phenotypes, e.g., other diseases, prognosis, theranosis, disease stratification, disease monitoring, treatment monitoring or prediction of non-responder / responder status, and the like.

[00134] A biosignature for characterizing a phenotype can be determined by detecting microRNA and/or vesicles. The microRNA can be assessed within a vesicle. Alternately, the microRNA and vesicles in a sample are analyzed to characterize the phenotype without isolating the microRNA from the vesicles. Many analytical techniques are available to assess vesicles. In some embodiments, vesicle levels are characterized using mass spectrometry, flow cytometry, immunocytochemical staining, Western blotting, electrophoresis, chromatography or x-ray crystallography in accordance with procedures known in the art. For example, vesicles can be characterized and quantitatively measured using flow cytometry as described in *Clayton et al., Journal of Immunological Methods 2001;163-174,* which is herein incorporated by reference in its entirety. Vesicle levels may be determined using binding agents as described above. For example, a binding agent to vesicles can be labeled and the label detected and used to determine the amount of vesicles in a sample. The binding agent can be bound to a substrate, such as arrays or particles, such as described above. Alternatively, the vesicles may be labeled directly.

[00135] In some embodiments, electrophoretic tags or eTags are used to determine the amount of vesicles of interest. eTags are small fluorescent molecules linked to nucleic acids or antibodies and are designed to bind one specific nucleic acid sequence or protein, respectively. After the eTag binds its target, an enzyme is used to cleave the bound eTag from the target. The signal generated from the released eTag, called a "reporter," is proportional to the amount of target nucleic acid or protein in the sample. The eTag reporters can be identified by capillary electrophoresis. The unique charge-to-mass ratio of each eTag reporter--that is, its electrical charge divided by its molecular weight--makes it show up as a specific peak on the capillary electrophoresis readout. Thus, by targeting a specific biomarker of a vesicle with an eTag, the amount or level of vesicles is determined.

[00136] The vesicle level can determined from a heterogeneous population of vesicles, such as the total population of vesicles in a sample. Alternatively, the vesicles level is determined from a homogenous population, or substantially homogenous population of vesicles, such as the level of specific cell-of-origin vesicles, such as vesicles from prostate cancer cells. In yet other embodiments, the level is determined for vesicles with a particular biomarker or combination of biomarkers, such as a biomarker specific for colorectal cancer. Determining the level vesicles can be performed in conjunction with determining the biomarker or combination of biomarkers of a vesicle. Alternatively, determining the amount of vesicle may be performed prior to or subsequent to determining the biomarker or combination of biomarkers of the vesicles.

[00137] The amount of vesicles in a sample can be assayed in a multiplexed manner. Multiplex analysis can be used for determining the amount of more than one population of vesicles, such as different cell-of-origin specific vesicles with different biomarkers or combination of biomarkers.

Specificity and Sensitivity

[00138] Performance of a diagnostic or related test is typically assessed using statistical measures. The performance of the characterization can be assessed by measuring sensitivity, specificity and related measures. For example, a level of microRNAs of interest can be assayed to characterize a phenotype, such as detecting a disease. The sensitivity and specificity of the assay to detect the disease is determined.

[00139] A true positive is a subject with a characteristic, e.g., a disease or disorder, correctly identified as having the characteristic. A false positive is a subject without the characteristic that the test improperly identifies as having the characteristic. A true negative is a subject without the characteristic that the test correctly identifies as not having the characteristic. A false negative is a person with the characteristic that the test improperly identifies as not having the characteristic. The ability of the test to distinguish between these classes provides a measure of test performance.

[00140] The specificity of a test is defined as the number of true negatives divided by the number of actual negatives (i.e., sum of true negatives and false positives). Specificity is a measure of how many subjects are correctly identified as negatives. A specificity of 100% means that the test recognizes all actual negatives - for example, all healthy people will be recognized as healthy. A lower specificity indicates that more negatives will be determined as positive.

[00141] The sensitivity of a test is defined as the number of true positives divided by the number of actual positives (i.e., sum of true positives and false negatives). Specificity is a measure of how many subjects are correctly identified as positives. A sensitivity of 100% means that the test recognizes all actual positives - for example, all sick people will be recognized as sick. A lower sensitivity indicates that more positives will be missed by being determined as negative.

[00142] The accuracy of a test is defined as the number of true positives and true negatives divided by the sum of all true and false positives and all true and false negatives. It provides one number that combines sensitivity and specificity measurements.

[00143] Sensitivity, specificity and accuracy are determined at a particular discrimination threshold value. For example, a common threshold for prostate cancer (PCa) detection is 4 ng/mL of prostate specific antigen (PSA) in serum. A level of PSA equal to or above the threshold is considered positive for PCa and any level below is considered negative. As the threshold is varied, the sensitivity and specificity will also vary. For example, as the threshold for detecting cancer is increased, the specificity will increase because it is harder to call a subject positive, resulting in fewer false positives. At the same time, the sensitivity will decrease. A receiver operating characteristic curve (ROC curve) is a graphical plot of the true positive rate (i.e., sensitivity) versus the false positive rate (i.e., 1 − specificity) for a binary classifier system as its discrimination threshold is varied. The ROC curve shows how sensitivity and specificity change as the threshold is varied. The Area Under the Curve (AUC) of an ROC curve provides a summary value indicative of a test's performance over the entire range of thresholds. The AUC is equal to the probability that a classifier will rank a randomly chosen positive sample higher than a randomly chosen negative sample. An AUC of 0.5 indicates that the test has a 50% chance of proper ranking, which is equivalent to no discriminatory power (a coin flip also has a 50% chance of proper ranking). An AUC of 1.0 means that the test properly ranks (classifies) all subjects. The AUC is equivalent to the Wilcoxon test of ranks.

[00144] A biosignature according to the invention can be used to characterize a phenotype with at least 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, or 70% sensitivity, such as with at least 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, or 87% sensitivity. In some embodiments, the phenotype is characterized with at least 87.1, 87.2, 87.3, 87.4, 87.5, 87.6, 87.7, 87.8, 87.9, 88.0, or 89% sensitivity, such as at least 90% sensitivity. The phenotype can be characterized with at least 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sensitivity.

[00145] A biosignature according to the invention can be used to characterize a phenotype of a subject with at least 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, or 97% specificity, such as with at least 97.1, 97.2, 97.3, 97.4, 97.5, 97.6, 97.7, 97.8, 97.8, 97.9, 98.0, 98.1, 98.2, 98.3, 98.4, 98.5, 98.6, 98.7, 98.8, 98.9, 99.0, 99.1, 99.2, 99.3, 99.4, 99.5, 99.6, 99.7, 99.8, 99.9 or 100% specificity.

[00146] A biosignature according to the invention can be used to characterize a phenotype of a subject, e.g., based on microRNA level or other characteristic, with at least 50% sensitivity and at least 60, 65, 70, 75, 80, 85, 90, 95, 99, or 100% specificity; at least 55% sensitivity and at least 60, 65, 70, 75, 80, 85, 90, 95, 99, or 100% specificity; at least 60% sensitivity and at least 60, 65, 70, 75, 80, 85, 90, 95, 99, or 100% specificity; at least 65% sensitivity and at least 60, 65, 70, 75, 80, 85, 90, 95, 99, or 100% specificity; at least 70% sensitivity and at least 60, 65, 70, 75, 80, 85, 90, 95, 99, or 100% specificity; at least 75% sensitivity and at least 60, 65, 70, 75, 80, 85, 90, 95, 99, or 100% specificity; at least 80% sensitivity and at least 60, 65, 70, 75, 80, 85, 90, 95, 99, or 100% specificity; at least 85% sensitivity and at least 60, 65, 70, 75, 80, 85, 90, 95, 99, or 100% specificity; at least 86% sensitivity and at least 60, 65, 70, 75, 80, 85, 90, 95, 99, or 100% specificity; at least 87% sensitivity and at least 60, 65, 70, 75, 80, 85, 90, 95, 99, or 100% specificity; at least 88% sensitivity and at least 60, 65, 70, 75, 80, 85, 90, 95, 99, or 100% specificity; at least 89% sensitivity and at least 60, 65, 70, 75, 80, 85, 90, 95, 99,

or 100% specificity; at least 90% sensitivity and at least 60, 65, 70, 75, 80, 85, 90, 95, 99, or 100% specificity; at least 91% sensitivity and at least 60, 65, 70, 75, 80, 85, 90, 95, 99, or 100% specificity; at least 92% sensitivity and at least 60, 65, 70, 75, 80, 85, 90, 95, 99, or 100% specificity; at least 93% sensitivity and at least 60, 65, 70, 75, 80, 85, 90, 95, 99, or 100% specificity; at least 94% sensitivity and at least 60, 65, 70, 75, 80, 85, 90, 95, 99, or 100% specificity; at least 95% sensitivity and at least 60, 65, 70, 75, 80, 85, 90, 95, 99, or 100% specificity; at least 96% sensitivity and at least 60, 65, 70, 75, 80, 85, 90, 95, 99, or 100% specificity; at least 97% sensitivity and at least 60, 65, 70, 75, 80, 85, 90, 95, 99, or 100% specificity; at least 98% sensitivity and at least 60, 65, 70, 75, 80, 85, 90, 95, 99, or 100% specificity; at least 99% sensitivity and at least 60, 65, 70, 75, 80, 85, 90, 95, 99, or 100% specificity; or substantially 100% sensitivity and at least 60, 65, 70, 75, 80, 85, 90, 95, 99, or 100% specificity.

[00147] A biosignature according to the invention can be used to characterize a phenotype of a subject with at least 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, or 97% accuracy, such as with at least 97.1, 97.2, 97.3, 97.4, 97.5, 97.6, 97.7, 97.8, 97.8, 97.9, 98.0, 98.1, 98.2, 98.3, 98.4, 98.5, 98.6, 98.7, 98.8, 98.9, 99.0, 99.1, 99.2, 99.3, 99.4, 99.5, 99.6, 99.7, 99.8, 99.9 or 100% accuracy.

[00148] In some embodiments, a biosignature according to the invention is used to characterize a phenotype of a subject with an AUC of at least 0.60, 0.61, 0.62, 0.63, 0.64, 0.65, 0.66, 0.67, 0.68, 0.69, 0.70, 0.71, 0.72, 0.73, 0.74, 0.75, 0.76, 0.77, 0.78, 0.79, 0.80, 0.81, 0.82, 0.83, 0.84, 0.85, 0.86, 0.87, 0.88, 0.89, 0.90, 0.91, 0.92, 0.93, 0.94, 0.95, 0.96, or 0.97, such as with at least 0.971, 0.972, 0.973, 0.974, 0.975, 0.976, 0.977, 0.978, 0.978, 0.979, 0.980, 0.981, 0.982, 0.983, 0.984, 0.985, 0.986, 0.987, 0.988, 0.989, 0.99, 0.991, 0.992, 0.993, 0.994, 0.995, 0.996, 0.997, 0.998, 0.999 or 1.00.

[00149] Furthermore, the confidence level for determining the specificity, sensitivity, accuracy or AUC, may be determined with at least 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, or 99% confidence.

[00150] Other related performance measures include positive and negative likelihood ratios [positive LR = sensitivity/(1-specificity); negative LR = (1-sensitivity)/specificity]. Such measures can also be used to gauge test performance according to the methods of the invention.

Classification

[00151] Biosignature according to the invention can be used to classify a sample. For example, a sample can be classified as, or predicted to be, a responder or non-responder to a given treatment for a given disease or disorder. Many statistical classification techniques are known to those of skill in the art. In supervised learning approaches, a group of samples from two or more groups are analyzed with a statistical classification method. Biomarkers can be discovered that can be used to build a classifier that differentiates between the two or more groups. A new sample can then be analyzed so that the classifier can associate the new with one of the two or more groups. Commonly used supervised classifiers include without limitation the neural network (multi-layer perceptron), support vector machines, k-nearest neighbors, Gaussian mixture model, Gaussian, naive Bayes, decision tree and radial basis function (RBF) classifiers. Linear classification methods include Fisher's linear discriminant, logistic regression, naive Bayes classifier, perceptron, and support vector machines (SVMs). Other classifiers for use with the invention include quadratic classifiers, k-nearest neighbor, boosting, decision

trees, random forests, neural networks, pattern recognition, Bayesian networks and Hidden Markov models. One of skill will appreciate that these or other classifiers, including improvements of any of these, are contemplated within the scope of the invention.

[00152] Classification using supervised methods is generally performed by the following methodology:

[00153] In order to solve a given problem of supervised learning (e.g. learning to recognize handwriting) one has to consider various steps:

[00154] 1. Gather a training set. These can include, for example, samples that are from a subject with or without a disease or disorder, subjects that are known to respond or not respond to a treatment, subjects whose disease progresses or does not progress, etc. The training samples are used to "train" the classifier.

[00155] 2. Determine the input "feature" representation of the learned function. The accuracy of the learned function depends on how the input object is represented. Typically, the input object is transformed into a feature vector, which contains a number of features that are descriptive of the object. The number of features should not be too large, because of the curse of dimensionality; but should be large enough to accurately predict the output. The features might include a set of biomarkers such as those derived from vesicles as described herein.

[00156] 3. Determine the structure of the learned function and corresponding learning algorithm. A learning algorithm is chosen, e.g., artificial neural networks, decision trees, Bayes classifiers or support vector machines. The learning algorithm is used to build the classifier.

[00157] 4. Build the classifier. The learning algorithm is run the gathered training set. Parameters of the learning algorithm may be adjusted by optimizing performance on a subset (called a validation set) of the training set, or via cross-validation. After parameter adjustment and learning, the performance of the algorithm may be measured on a test set of naive samples that is separate from the training set.

[00158] Once the classifier is determined as described above, it can be used to classify a sample, e.g., that of a subject who is being analyzed by the methods of the invention. As an example, a classifier can be built using data for levels of microRNA of interest in reference subjects with and without a disease as the training and test sets. MicroRNA levels found in a sample from a test subject are assessed and the classifier is used to classify the subject as with or without the disease.

[00159] Unsupervised learning approaches can also be used with the invention. Clustering is an unsupervised learning approach wherein a clustering algorithm correlates a series of samples without the use the labels. The most similar samples are sorted into "clusters." A new sample could be sorted into a cluster and thereby classified with other members that it most closely associates. Many clustering algorithms are known to those of skill in the art.

Cell-of-Origin and Disease-Specific Vesicles

[00160] The microRNA of the invention can be isolated from vesicles. Thus, a biosignature can include the level of the microRNA along with characteristics of the vesicles, such as vesicle surface antigen, or mRNA payload. The microRNA can be isolated from a vesicle with a biosignature of interest, e.g., a vesicle from a certain cellular origin or from a cell having a particular disorder or disease. The vesicle binding agents disclosed herein can be used to isolate or detect a vesicle, such as a cell-of-origin vesicle or vesicle with a specific biosignature. In one embodiment, binding agents are used to isolate or detect a heterogeneous population of vesicles from a sample. In one embodiment, the binding agents are used to isolate or detect a

homogeneous population of vesicles from a heterogeneous population of vesicles. The homogeneous population can be cell-of-origin specific vesicles or other populations of vesicles with specific biosignatures.

[00161] A homogeneous population of vesicles, such as cell-of-origin specific vesicles, can be analyzed to characterize a phenotype for a subject. Cell-of-origin specific vesicles are vesicles derived from specific cell types, which include without limitation cells of a defined tissue, defined organ, tumor of interest or other diseased tissue of interest, circulating tumor or diseased cells, or cells of maternal or fetal origin. In some embodiments, the vesicles are derived from tumor cells or lung, pancreas, stomach, intestine, bladder, kidney, ovary, testis, skin, colorectal, breast, prostate, brain, esophagus, liver, placenta, or fetal cells. The isolated vesicle can also be from a particular sample type, such as vesicle from urine, blood, semen, feces, saliva, other bodily fluids, or solid tissue.

[00162] A cell-of-origin specific vesicle from a biological sample can be isolated using one or more binding agents that are specific for vesicles for that cell-of-origin. In one embodiment, the binding agents recognize surface antigens on the surface of the vesicles, e.g., surface proteins. In an embodiment, vesicles for analysis of a disease or condition are isolated using one or more binding agents specific for biomarkers for that disease or condition. The disease include cancers, neurological disorders, cardiovascular disorders, immune disorders (e.g., autoimmune diseases), infectious disorders (e.g., microbial or viral diseases).

[00163] A vesicle can be concentrated prior to isolation or detection of a cell-of-origin specific vesicle, such as through centrifugation, chromatography, or filtration, as described above. This step or steps can produce a heterogeneous population of vesicles prior to isolation of cell-of-origin specific vesicles. Alternatively, the vesicle is not concentrated, or the biological sample is not enriched for a vesicle, prior to isolation of a cell-of-origin vesicle. An example of the later case includes direct capture from a bodily fluid such as blood.

[00164] **FIG. 1B** illustrates a flowchart which depicts one method **100B** for isolating or identifying a cell-of-origin specific vesicle. First, a biological sample is obtained from a subject in step **102**. The sample can be obtained from a third party or from the same party performing the vesicle analysis. Next, cell-of-origin specific vesicles are isolated from the biological sample in step **104**. The isolated cell-of-origin specific vesicles are then analyzed in step **106** and a biomarker or biosignature for a particular phenotype is identified in step **208**. The method may be applied to measure any appropriate phenotype. The biosignature can include vesicle payload, e.g., the level of microRNAs of interest within the vesicle. In some embodiments, prior to step **104,** vesicles are concentrated or isolated from a biological sample to produce a homogeneous population of vesicles. For example, a heterogeneous population of vesicles may be isolated using centrifugation, chromatography, filtration, or other methods as described above, prior to use of one or more binding agents specific for isolating or identifying vesicles derived from specific cell types.

[00165] A cell-of-origin specific vesicle can be isolated from a biological sample of a subject using one or more binding agents that bind with high specificity to the cell-of-origin specific vesicle. In some embodiments, a single binding agent is used to isolate a cell-of-origin specific vesicle. In other embodiments, a combination of binding agents is used to isolate a cell-of-origin specific vesicle. In some embodiments, at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 50, 75, or 100 different binding agents are used to isolate a cell-of-origin vesicle. A population of vesicles having the same binding agent profile can be identified by using a single or a plurality of binding agents.

[00166] One or more binding agents can be selected based on their specificity for a target antigen(s) that is specific to a cell-of-origin, e.g., a cell-of-origin that is related to a tumor, autoimmune disease, cardiovascular disease, neurological disease, infection or other disease or disorder. The cell-of-origin can be from a cell that is informative for a diagnosis, prognosis, disease stratification, theranosis, prediction of responder / non-responder status, disease monitoring, treatment monitoring and the like as related to such diseases and disorders. The cell-of-origin can also be from a cell useful to discover biomarkers for use thereto. Non-limiting examples of antigens which may be used singularly, or in combination, to isolate a cell-of-origin specific vesicle, disease specific vesicle, or tumor specific vesicle, as listed in **Table 2** and are also described herein. The antigen can comprise membrane bound antigens which are accessible to binding agents, e.g., surface proteins or fragments thereof. In some embodiments, the antigen is a biomarker related to characterizing a phenotype, e.g., a disease marker. In some embodiments, the antigen is a biomarker specific to a cell-of-origin, e.g., a cell derived from the prostate, lung, breast, or GI tract. In some embodiments, the antigen is a biomarker specific to a class of vesicles, e.g., exosomes.

[00167] A number of exemplary binding agents useful for binding to vesicles associated with cancer, autoimmune diseases, cardiovascular diseases, neurological diseases, infection or other disease or disorders are presented in U.S. Patent Application No. 12/591,226, filed November 12, 2009 and entitled "METHODS AND SYSTEMS OF USING EXOSOMES FOR DETERMINING PHENOTYPES," which application is hereby incorporated by reference in its entirety.

[00168] One of skill will appreciate that any applicable antigen that can be used to isolate an informative vesicle is contemplated by the invention. Binding agents, e.g., antibodies, aptamers and lectins, can be chosen that recognize surface antigens and/or fragments thereof, as outlined herein. The binding agents can recognize antigens specific to the desired cell type or location and/or recognize biomarkers associated with the desired cells. The cells can be, e.g., tumor cells, other diseased cells, cells that serve as markers of disease such as activated immune cells, etc. One of skill will appreciate that binding agents for any cells of interest can be useful for isolating vesicles associated with those cells. One of skill will further appreciate that the binding agents disclosed herein can be used for detecting vesicles of interest. As a non-limiting example, a binding agent to a vesicle biomarker can be labeled directly or indirectly in order to detect vesicles bound by one of more of the same or different binding agents.

[00169] The binding agents are chosen to characterize the phenotype of interest. For example, a vesicle derived from a prostate cancer cell can be isolated using a binding agent, e.g., an antibody or aptamer, that is specific for an antigen associated with vesicles from a cell of prostate cancer origin, including without limitation PSA, TMPRSS2, FASLG, TNFSF10, PSMA, PCSA, NGEP, Il-7RI, CSCR4, CysLT1R, TRPM8, Kv1.3, TRPV6, TRPM8, PSGR, MISIIR, galectin-3, PCA3, TMPRSS2:ERG, or a combination thereof. Any appropriate antigens that are specific for vesicles derived from prostate cancer cells can be used for isolation thereof. Similarly, a vesicle derived from a benign prostatic hyperplasia (BPH) cell can be isolated using a binding agent, e.g., an antibody or aptamer, which is specific for an antigen associated with vesicles from a cell associated with BPH including, but not limited to, KIA1, intact fibronectin, or a combination thereof. Any appropriate antigens that are specific for vesicles derived from cells associated with BPH can be used for isolation thereof.

[00170] One of skill will appreciate that binding agents for biomarkers of vesicles associated with other cells of interest can be used similarly, including those disclosed in U.S. Patent Application No. 12/591,226, filed November 12, 2009 and entitled "METHODS AND SYSTEMS OF USING EXOSOMES FOR DETERMINING PHENOTYPES," which application is hereby incorporated by reference in its entirety. Likewise, additional markers for the cell types can be useful for isolating those vesicles, either individually, in combination with one or more markers listed above, or in combination with other markers. Cell-specific binding agents can be used in combination with vesicle specific binding agents to isolate vesicles from a given origin. As a non-limiting illustrative example, vesicle binding agents can be used in combination with breast cancer-specific binding agents to detect or isolate vesicles of breast cancer origin.

[00171] A cell-of-origin specific vesicle can be isolated using novel binding agents, e.g., using methods such as described herein. A cell-of-origin specific vesicle can also be isolated from a biological sample using isolation methods based on cellular binding partners or binding agents of such vesicles. Such cellular binding partners include without limitation peptides, proteins, RNA, DNA, apatmers, lectins, cells or serum-associated proteins. Useful binding partners bind in a recognizable manner to desired vesicles when one or more specific biomarkers are present. Isolation or detection of a cell-of-origin specific vesicle can be carried out with a single binding partner or binding agent, or a combination of binding partners or binding agents whose singular application or combined application results in cell-of-origin specific isolation or detection. Non-limiting examples of such binding agents are provided in **Table 3**. As a non-limiting illustrative example, a vesicle for characterizing breast cancer can be isolated with one or more binding agents including estrogen, progesterone, trastuzumab, CCND1, MYC PNA, IGF-1 PNA, MYC PNA, SC4 aptamer (Ku), AII-7 aptamer (ERB2), Galectin-3, mucin-type O-glycans, L-PHA, and/or Galectin-9. In some embodiments, one or more of these are used along with antibodies that recognize breast cancer markers as described above.

[00172] In various embodiments, binding agents are used for isolating or detecting cell-of-origin specific vesicles based on: i) detection of binding to antigens specific for cell-of-origin specific vesicles; ii) the absence of detection of markers specific for cell-of-origin specific vesicles; or iii) detection of expression levels of biomarkers specific for cell-of-origin specific vesicles. In an embodiment, a heterogeneous population of vesicles is applied to a surface coated with specific binding agents designed to identify the cell-of-origin characteristics of the vesicles. Various binding agents, e.g., antibodies or aptamers, can be arrayed on a solid surface or substrate wherein the heterogeneous population of vesicles is allowed to contact the solid surface or substrate for a sufficient time to allow binding events to take place. The presence or absence of binding events at given locations on the array surface or substrate can identify the presence or absence of vesicle populations that are specific to a given cell-of-origin. That is, binding events signal the presence of a vesicle having an antigen recognized by the bound antibody or aptamer. Conversely, lack of binding events signal that the absence of vesicles having an antigen recognized by the bound antibody or aptamer.

[00173] A cell-of-origin specific vesicle can be enriched or isolated using one or more binding agents using a magnetic capture method, fluorescence activated cell sorting (FACS) or laser cytometry as described herein. Magnetic capture methods include, but are not limited to, the use of magnetically activated cell sorter (MACS) microbeads or magnetic columns. Examples of immunoaffinity and magnetic particle methods that can be used are found in U.S. Patent Nos. 4,551,435, 4,795,698, 4,925,788, 5,108,933, 5,186,827, 5,200,084 or 5,158,871.

A cell-of-origin specific vesicle can also be isolated following the general methods described in U.S. Patent No. 7,399,632, by using combination of antigens specific to a vesicle.

[00174] Any other appropriate method for isolating or otherwise enriching the cell-of-origin specific vesicles with respect to a biological sample can be used according to the present invention. As described herein, size exclusion chromatography such as gel permeation columns, centrifugation or density gradient centrifugation, and filtration methods can be used in combination with the other antigen selection methods described herein. The cell-of-origin specific vesicles may also be isolated following the methods described in *Koga et al., Anticancer Research, 25:3703-3708 (2005), Taylor et al., Gynecologic Oncology, 110:13-21 (2008), Nanjee et al., Clin Chem, 2000;46:207-223* or U.S Patent No. 7,232,653.

[00175] Vesicles can be isolated and/or detected to provide diagnosis, prognosis, disease stratification, theranosis, prediction of responder or non-responder status, disease monitoring, treatment monitoring and the like. This analysis can include assessing the payload within the vesicles, e.g., microRNA species of interest. In one embodiment, vesicles are isolated from cells having a disease or disorder, e.g., cells derived from a malignant cell, a site of autoimmune disease, cardiovascular disease, neurological disease, or infection. In some embodiments, the isolated vesicles are derived from cells related to such diseases and disorders. The isolated vesicles are also useful to discover novel biomarkers. By identifying biomarkers associated with vesicles, isolated vesicles can be assessed for characterizing a phenotype as described herein.

Biosignature

[00176] A microRNA biosignature from a subject can be used to characterize a phenotype of the subject. A biosignature can further include the level of one or more additional biomarkers, e.g., circulating biomarkers or biomarkers associated with a vesicle of interest. A biosignature of a vesicle of interest can include particular antigens or biomarkers that are present on the vesicle. The biosignature can also include one or more antigens or biomarkers that are carried as payload within the vesicle, including the microRNA under examination. The biosignature can comprise a combination of one or more antigens or biomarkers that are present on the vesicle with one or more biomarkers that are detected in the vesicle. The biosignature can further comprise other information about a vesicle aside from its biomarkers. Such information can include vesicle size, circulating half-life, metabolic half-life, and specific activity in vivo or in vitro. The biosignature can comprise the biomarkers or other characteristics used to build a classifier.

[00177] In some embodiments, the microRNA is detected directly in a biological sample. For example, RNA in a bodily fluid can be isolated using commercially available kits such as *mir*Vana kits (Applied Biosystems/ Ambion, Austin, TX), MagMAX™ RNA Isolation Kit (Applied Biosystems/ Ambion, Austin, TX), and QIAzol Lysis Reagent and RNeasy Midi Kit (Qiagen Inc., Valencia CA). Particular species of microRNAs can be determined using array or PCR techniques as described below.

[00178] In some embodiments, the microRNA payload with vesicles is assessed in order to characterize a phenotype. The vesicles can be purified or concentrated prior to determining the biosignature. For example, a cell-of-origin specific vesicle can be isolated and its biosignature determined. Alternatively, the biosignature of the vesicle can be directly assayed from a sample, without prior purification or concentration. The biosignature of the invention can be used to determine a diagnosis, prognosis, or theranosis of a disease or condition or similar measures described herein. A biosignature can also be used to determine treatment efficacy, stage of a

disease or condition, or progression of a disease or condition, or responder / non-responder status. Furthermore, a biosignature may be used to determine a physiological state, such as pregnancy.

[00179] A characteristic of a vesicle in and of itself can be assessed to determine a biosignature. The characteristic can be used to diagnose, detect or determine a disease stage or progression, the therapeutic implications of a disease or condition, or characterize a physiological state. Such characteristics include without limitation the level or amount of vesicles, vesicle size, temporal evaluation of the variation in vesicle half-life, circulating vesicle half-life, metabolic half-life of a vesicle, or activity of a vesicle.

[00180] Biomarkers that can be included in a biosignature include one or more proteins or peptides (e.g., providing a protein signature), nucleic acids (e.g. RNA signature as described, or a DNA signature), lipids (e.g. lipid signature), or combinations thereof. In some embodiments, the biosignature can also comprise the type or amount of drug or drug metabolite present in a vesicle, (e.g., providing a drug signature), as such drug may be taken by a subject from which the biological sample is obtained, resulting in a vesicle carrying the drug or metabolites of the drug.

[00181] A biosignature can also include an expression level, presence, absence, mutation, variant, copy number variation, truncation, duplication, modification, or molecular association of one or more biomarkers. A genetic variant, or nucleotide variant, refers to changes or alterations to a gene or cDNA sequence at a particular locus, including, but not limited to, nucleotide base deletions, insertions, inversions, and substitutions in the coding and non-coding regions. Deletions may be of a single nucleotide base, a portion or a region of the nucleotide sequence of the gene, or of the entire gene sequence. Insertions may be of one or more nucleotide bases. The genetic variant may occur in transcriptional regulatory regions, untranslated regions of mRNA, exons, introns, or exon/intron junctions. The genetic variant may or may not result in stop codons, frame shifts, deletions of amino acids, altered gene transcript splice forms or altered amino acid sequence.

[00182] In an embodiment, nucleic acid biomarkers, including nucleic acid payload within a vesicle, is assessed for nucleotide variants. The nucleic acid biomarker may comprise one or more RNA species, e.g., mRNA, miRNA, snoRNA, snRNA, rRNAs, tRNAs, siRNA, hnRNA, shRNA, or a combination thereof. Similarly, DNA payload can be assessed to form a DNA signature.

[00183] An RNA signature or DNA signature can also include a mutational, epigenetic modification, or genetic variant analysis of the RNA or DNA present in the vesicle. Epigenetic modifications include patterns of DNA methylation. See, e.g., Lesche R. and Eckhardt F., DNA methylation markers: a versatile diagnostic tool for routine clinical use. Curr Opin Mol Ther. 2007 Jun;9(3):222-30, which is incorporated herein by reference in its entirety. Thus, a biomarker can be the methylation status of a segment of DNA.

[00184] A biosignature can comprise one or more miRNA signatures combined with one or more additional signatures including, but not limited to, an mRNA signature, DNA signature, protein signature, peptide signature, antigen signature, or any combination thereof. For example, the biosignature can comprise one or more miRNA biomarkers with one or more DNA biomarkers, one or more mRNA biomarkers, one or more snoRNA biomarkers, one or more protein biomarkers, one or more peptide biomarkers, one or more antigen biomarkers, one or more antigen biomarkers, one or more lipid biomarkers, or any combination thereof.

[00185] A biosignature can comprise a combination of one or more antigens or binding events with more or more binding agents, such as listed in **Tables 2** and **3**, or those described in U.S. Patent Application No. 12/591,226, filed November 12, 2009 and entitled "METHODS AND SYSTEMS OF USING EXOSOMES

FOR DETERMINING PHENOTYPES," which application is hereby incorporated by reference in its entirety. The biosignature can further comprise one or more other biomarkers, such as, but not limited to, miRNA, DNA (e.g. single stranded DNA, complementary DNA, or noncoding DNA), or mRNA. For example, the biosignature of a vesicle can comprise a combination of one or more antigens, such as shown in **Table 2**, one or more binding agents, such as shown in **Table 3**, and one or more biomarkers for a condition or disease of interest such as those described in U.S. Patent Application No. 12/591,226. The biosignature can comprise one or more biomarkers, for example miRNA, with one or more antigens specific for a cancer cell (for example, as shown in **Table 2**). The biosignature can be derived from surface markers on the vesicle and/or payload markers from within the vesicle (e.g., miRNA payload).

[00186] In some embodiments, a vesicle used in the subject methods has a biosignature that is specific to the cell-of-origin and is used to derive disease-specific or biological state specific diagnostic, prognostic or therapy-related biosignatures representative of the cell-of-origin. In other embodiments, a vesicle has a biosignature that is specific to a given disease or physiological condition that is different from the biosignature of the cell-of-origin for use in the diagnosis, prognosis, staging, therapy-related determinations or physiological state characterization. Biosignatures can also comprise a combination of cell-of-origin specific and non-specific vesicles.

[00187] Biosignatures can be used to evaluate diagnostic criteria such as presence of disease, disease staging, disease monitoring, disease stratification, or surveillance for detection, metastasis or recurrence or progression of disease. A biosignature can also be used clinically in making decisions concerning treatment modalities including therapeutic intervention. A biosignature can further be used clinically to make treatment decisions, including whether to perform surgery or what treatment standards should be utilized along with surgery (e.g., either pre-surgery or post-surgery). As an illustrative example, a microRNA (miRNA) biosignature that indicates an aggressive form of cancer may call for a more aggressive surgical procedure and/or more aggressive therapeutic regimen to treat the patient.

[00188] A biosignature can be used in therapy related diagnostics to provide tests useful to diagnose a disease or choose the correct treatment regimen, such as provide a theranosis. Theranostics includes diagnostic testing that provides the ability to affect therapy or treatment of a diseased state. Theranostics testing provides a theranosis in a similar manner that diagnostics or prognostic testing provides a diagnosis or prognosis, respectively. As used herein, theranostics encompasses any desired form of therapy related testing. Therapy related tests can be used to predict and assess drug response in individual subjects, i.e., to provide personalized medicine. Therapy related tests are also useful to select a subject for treatment who is particularly likely to benefit from the treatment or to provide an early and objective indication of treatment efficacy in an individual subject. Thus, a biosignature as disclosed herein may indicate that treatment should be altered to select a more promising treatment, thereby avoiding the great expense of delaying beneficial treatment and avoiding the financial and morbidity costs of administering an ineffective drug(s).

[00189] Therapy related diagnostics are also useful in clinical diagnosis and management of a variety of diseases and disorders, which include, but are not limited to cardiovascular disease, cancer, infectious diseases, sepsis, neurological diseases, central nervous system related diseases, endovascular related diseases, and autoimmune related diseases. Therapy related diagnostics also aid in the prediction of drug toxicity, drug resistance or drug response. Therapy related tests may be developed in any suitable diagnostic testing format,

which include, but are not limited to, e.g., immunohistochemical tests, clinical chemistry, immunoassay, cell-based technologies, nucleic acid tests or body imaging methods. Therapy related tests can further include but are not limited to, testing that aids in the determination of therapy, testing that monitors for therapeutic toxicity, or response to therapy testing. Thus, a biosignature can be used to predict or monitor a subject's response to a treatment. A biosignature can be determined at different time points for a subject after initiating, removing, or altering a particular treatment.

[00190] In some embodiments, a determination or prediction as to whether a subject is responding to a treatment is made based on a change in the amount of one or more components of a biosignature (i.e., the microRNA, vesicles and/or biomarkers of interest), an amount of one or more components of a particular biosignature, or the biosignature detected for the components. In another embodiment, a subject's condition is monitored by determining a biosignature at different time points. The progression, regression, or recurrence of a condition is determined. Response to therapy can also be measured over a time course. Thus, the invention provides a method of monitoring a status of a disease or other medical condition in a subject, comprising isolating or detecting a biosignature from a biological sample from the subject, detecting the overall amount of the components of a particular biosignature, or detecting the biosignature of one or more components (such as the presence, absence, or expression level of a biomarker). The biosignatures are used to monitor the status of the disease or condition.

[00191] In some embodiments, a biosignature is used to determine whether a particular disease or condition is resistant to a drug. If a subject is drug resistant, a physician need not waste valuable time with such drug treatment. To obtain early validation of a drug choice or treatment regimen, a biosignature is determined for a sample obtained from a subject. The biosignature is used to assess whether the particular subject's disease has the biomarker associated with drug resistance. Such a determination enables doctors to devote critical time as well as the patient's financial resources to effective treatments.

[00192] In some embodiments, a biosignature is used to assess whether a subject is afflicted with disease, is at risk for developing disease or to assess the stage or progression of the disease. In illustrative examples, a biosignature is used to assess whether a subject has prostate cancer by detecting one or more of the general vesicle markers CD9, CD63 and CD81; one or more prostate epithelial markers including PCSA or PSMA; and one or more cancer markers such as B7H3 and/or EpCam. Higher levels of the markers in a sample from a subject than in a control individual without prostate cancer can indicate the presence of the prostate cancer in the subject. In another illustrative example, a biosignature is used to determine a stage of a disease or condition as described in U.S. Patent Application No. 12/591,226.

[00193] In some embodiments, characterizing a phenotype comprises determining the amount of a heterogeneous population of components of a biosignature and the amount of one or more homogeneous components of the same biosignature. In an embodiment, characterizing a phenotype comprises determining of the total amount of microRNA in a sample and determining the presence of one or more particular species of microRNAs. A ratio of the particular microRNA to the total amount of microRNA can be compared in the sample and the reference. Threshold values, or reference values or amounts can be determined based on comparisons of normal subjects and subjects with the phenotype of interest, as further described herein, and criteria based on the threshold or reference values determined. The different criteria can be used to characterize a phenotype.

[00194] One of skill will understand that the same method can be applied to other components of a biosignature. As a non-limiting example, the level of a vesicle with a certain surface antigen in a sample can be compared to the total vesicle level in the sample. The ratio can be compared to a reference, thereby characterizing a phenotype of interest. In one embodiment, general vesicle markers, such as tetraspanins such as CD9, CD81, and CD63, are used to determine the amount of vesicles in a sample. The expression level of CD9, CD81, CD63, or a combination thereof can be detected and if the level is greater than a threshold level, the criterion is met. In another embodiment, the criterion is met if a level of CD9, CD81 and/or CD63, is lower than a threshold value or reference value. In another embodiment, the criterion is based on whether the amount of vesicles is higher than a threshold or reference value. Another criterion is based on the amount of vesicles with a specific biosignature. If the amount of vesicles with the specific biosignature is lower than a threshold or reference value, the criterion is met. In another embodiment, if the amount of vesicles with the specific biosignature is higher than a threshold or reference value, the criterion is met. A criterion can also be based on the amount of vesicles derived from a particular cell type. If the amount is lower than a threshold or reference value, the criterion is met. In another embodiment, if the amount is higher than a threshold value, the criterion is met. Similar rules can be used to compare the level of a microRNA to a reference.

[00195] In a non-limiting example, consider that vesicles from prostate cells are determined by detecting the biomarker PCSA or PSCA, and that a criterion is met if the level of detected PCSA or PSCA is greater than a threshold level. The threshold can be the level of the same markers in a sample from a control cell line or control subject. Another criterion can be based on whether the amount of vesicles derived from a cancer cell or comprising one or more cancer specific biomarkers. For example, the biomarkers B7H3, EpCam, or both, can be determined and a criterion met if the level of detected B7H3 and/or EpCam is greater than a threshold level or within a pre-determined range. If the amount is lower, or higher, than a threshold or reference value, the criterion is met. A criterion can also be the reliability of the result, such as meeting a quality control measure or value. A detected amount of B7H3 and/or EpCam in a test sample that is above the amount of these markers in a control sample may indicate the presence of a cancer in the test sample.

[00196] A phenotype for a subject can be characterized based on meeting any number of useful criteria. In some embodiments, at least 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 criteria are used. For example, for the characterizing of a cancer, a number of different criteria can be used when the subject is diagnosed with a cancer: 1) if the amount of microRNA in a sample from a subject is higher than a reference value; 2) if the amount of a microRNA within cell type specific vesicles (i.e. vesicles derived from a specific tissue or organ) is higher than a reference value; or 3) if the amount of microRNA within vesicles with one or more cancer specific biomarkers is higher than a reference value. Similar rules can apply if the amount of microRNA is less than or the same as the reference. The method can further include a quality control measure, such that the results are provided for the subject if the samples meet the quality control measure. In some embodiments, if the criteria are met but the quality control is questionable, the subject is reassessed. The same criteria can be applied to other biomarkers. For example, the levels of one or more circulating biomarkers (e.g., RNA, DNA, peptides), vesicles, mutations, etc, can be compared to a reference. Different components of a biosignature can have different criteria. As a non-limiting example, a biosignature used to diagnose a cancer can include overexpression of one miR species as compared to a reference and underexpression of a vesicle surface antigen as compared to a reference.

[00197] A biosignature can be determined by comparing the amount of vesicles, the structure of a vesicle, or any other informative characteristic of a vesicle. Vesicle structure can be assessed using transmission electron microscopy, see for example, *Hansen et al., Journal of Biomechanics 31, Supplement 1: 134-134(1) (1998),* or scanning electron microscopy. Various combinations of methods and techniques or analyzing one or more vesicles can be used to determine a phenotype for a subject.

[00198] A biosignature can include without limitation the presence or absence, copy number, expression level, or activity level of a biomarker. Other useful components of a biosignature include the presence of a mutation (e.g., mutations which affect activity of a transcription or translation product, such as substitution, deletion, or insertion mutations), variant, or post-translation modification of a biomarker. Post-translational modification of a protein biomarker include without limitation acylation, acetylation, phosphorylation, ubiquitination, deacetylation, alkylation, methylation, amidation, biotinylation, gamma-carboxylation, glutamylation, glycosylation, glycyation, hydroxylation, covalent attachment of heme moiety, iodination, isoprenylation, lipoylation, prenylation, GPI anchor formation, myristoylation, farnesylation, geranylgeranylation, covalent attachment of nucleotides or derivatives thereof, ADP-ribosylation, flavin attachment, oxidation, palmitoylation, pegylation, covalent attachment of phosphatidylinositol, phosphopantetheinylation, polysialylation, pyroglutamate formation, racemization of proline by prolyl isomerase, tRNA-mediation addition of amino acids such as arginylation, sulfation, the addition of a sulfate group to a tyrosine, or selenoylation of the biomarker.

[00199] The methods described herein can be used to identify a biosignature that is associated with a disease, condition or physiological state. The biosignature can also be used to determine if a subject is afflicted with cancer or is at risk for developing cancer. A subject at risk of developing cancer can include those who may be predisposed or who have pre-symptomatic early stage disease. A biosignature can also be used to provide a diagnostic or theranostic determination for other diseases and conditions such as those disclosed in U.S. Patent Application No. 12/591,226.

[00200] A biosignature can be used for pre-symptomatic diagnosis. Furthermore, the biosignature can be utilized to detect disease, determine disease stage or progression, determine the recurrence of disease, identify treatment protocols, determine efficacy of treatment protocols or evaluate the physiological status of individuals related to age and environmental exposure.

[00201] In some embodiments, a biosignature is used to identify conditions or diseases, including cancers of unknown origin, also known as cancers of unknown primary (CUP). In one embodiment, microRNA species that are indicative of a particular cellular origin are detected in the sample. In another embodiment, a vesicle is isolated from the biological sample to arrive at a heterogeneous population of vesicles. The heterogeneous population of vesicles is contacted with specific binding agents designed to identify antigen specific characteristics of the vesicle population that are specific to a given cell-of-origin. In still another embodiment, vesicles are isolated from the sample that are specific to a given cell-of-origin, and the microRNA payload in the isolated vesicles is assessed.

[00202] Further, as described above, the biosignature can correlate with the cancerous state of cells. Compounds that inhibit cancer in a subject may cause a change, e.g., a change in biosignature, which can be monitored by serial measurement of the biosignature over time and course of treatment. Any change in the biosignature can be determined to concomitantly monitor treatment efficacy.

[00203] In an aspect, characterizing a phenotype of a subject comprises a method of determining whether the subject is likely to respond or not respond to a therapy. The methods of the invention also include biosignatures useful in predicting whether the subject is likely to respond or not. One or more subjects that respond to a therapy (responders) and one or more subjects that do not respond to the same therapy (non-responders) can have their biosignature interrogated. Interrogation can be performed to identify biosignatures that classify a subject as a responder or non-responder to the treatment of interest. The biosignature can comprise microRNA and optionally vesicles. In some aspects, the presence, quantity, and payload of a vesicle are assayed. The payload of a vesicle includes, for example, internal proteins, nucleic acids such as miRNA, lipids or carbohydrates.

[00204] A biosignature indicative of responder / non-responder status can be used for theranosis. A sample from subjects with known or determinable responder / non-responder status may be analyzed for one or more of the following: amount of microRNA, amount of a unique subset or species of microRNA, amount of vesicles, amount of vesicles with certain antigens, biomarkers in such vesicles, etc. In one instance, vesicles such as microvesicles or exosomes from responders and non-responders are analyzed for the presence and/or quantity of one or more miRNAs, such as miR-122 or miR-141. A difference in biosignatures between responders and non-responders can be used for theranosis. In another embodiment, microRNA are obtained from subjects having a disease or condition. MicroRNAs are also obtained from subjects free of such disease or condition. The microRNAs from both groups of subjects are assayed for unique biosignatures that are associated with all subjects in that group but not in subjects from the other group. Such biosignatures or biomarkers can then used as a diagnostic for the presence or absence of the condition or disease, or to classify the subject as belonging on one of the groups (those with/without disease, aggressive/non-aggressive disease, responder/non-responder, etc).

[00205] In an aspect, characterizing a phenotype of a subject comprises a method of staging a disease. The methods of the invention also include determining new biosignatures useful in staging. In an illustrative example, microRNAs are assayed from patients having a stage I cancer and patients having stage II or stage III of the same cancer. In some embodiments, microRNAs are assayed in patients with metastatic disease. A difference in biosignatures or biomarkers from each group of patient is identified (e.g., certain microRNAs from stage III cancer may have an increased expression as compared to stage II cancer), thereby identifying a biosignature or biomarker that distinguishes different stages of a disease. Such biosignature can then be used to stage patients having the disease.

[00206] In some instances, a biosignature is determined by assaying a sample from a subject over a period of time, e.g., daily, semiweekly, weekly, biweekly, semimonthly, monthly, bimonthly, semiquarterly, quarterly, semiyearly, biyearly or yearly. For example, the biosignatures in patients on a given therapy can be monitored over time to detect signatures indicative of responders or non-responders for the therapy. Similarly, patients with differing stages of disease have their biosignatures interrogated over time. A temporal pattern can thus form a biosignature that can then be used for theranosis, diagnosis, prognosis, disease stratification, treatment monitoring, disease monitoring or making a prediction of responder / non-responder status. As an illustrative example only, an increasing amount of a biomarker (e.g., miR 122) in vesicles over a time course is associated with metastatic cancer, as opposed to a stagnant amounts of the biomarker in vesicles over the time course that are associated with non-metastatic cancer. A time course may last over at least 1 week, 2 weeks, 3 weeks, 4 weeks, 1 month, 6 weeks, 8 weeks, 2 months, 10 weeks, 12 weeks, 3 months, 4 months, 5 months, 6 months, 7

months, 8 months, 9 months, 10 months, 11 months, 12 months, one year, 18 months, 2 years, or at least 3 years.

[00207] In an embodiment, characterizing a phenotype comprises assessing the efficacy of a cancer treatment, e.g., chemotherapy, radiation therapy, surgery, or any other therapeutic approach useful for treating cancer in a subject. In addition, a biosignature can be used in a screening assay to identify candidate or test compounds or agents (e.g., proteins, peptides, peptidomimetics, peptoids, small molecules or other drugs) that have a modulatory effect on the biosignature of a vesicle. Compounds identified via such screening assays may be useful, for example, for modulating, e.g., inhibiting, ameliorating, treating, or preventing conditions or diseases.

[00208] In one embodiment, the invention provides a screening method for drug development. A biosignature comprising a level of one or more microRNA is obtained from a patient who is undergoing successful treatment for a particular disease, e.g., a cancer. Cells from a patient with the disease but not being treated with the same treatment are cultured and used for determining biosignatures. The cultured cells are treated with test compounds and the biosignature from the cultures is compared to the biosignature obtained from the patient undergoing successful treatment. Biosignatures that are similar to those of the patient undergoing successful treatment indicate a successful treatment and the corresponding test compounds can be selected for further studies.

[00209] The biosignature of the invention can be used to monitor the influence of an agent (e.g., drug compounds) on the biosignature in clinical trials. Monitoring the level of microRNA, changes in the biosignature of a vesicle, or both, can also be used in a method of assessing the efficacy of a test compound, such as a test compound for inhibiting cancer cells. A biosignature of individuals who respond to the drug can also be used as a diagnostic predict responder / non-responder status of new patients.

[00210] Within the clinical trial setting of investigational agents as defined by the FDA, MDA, EMA, USDA, and EMEA, therapy-related diagnostics as determined by a biosignature disclosed herein, can provide key information to optimize trial design, monitor efficacy, and enhance drug safety. For instance, for trial design, therapy-related diagnostics can be used for patient stratification, determination of patient eligibility (inclusion/exclusion), creation of homogeneous treatment groups, and selection of patient samples that are optimized to a matched case control cohort. Such therapy-related diagnostic can therefore provide the means for patient efficacy enrichment, thereby minimizing the number of individuals needed for trial recruitment. For efficacy, therapy-related diagnostics can be useful for monitoring therapy and assessing efficacy criteria. Alternatively, for safety, therapy-related diagnostics can be used to prevent adverse drug reactions or avoid medication error and monitor compliance with the therapeutic regimen.

[00211] In some embodiments, the invention provides a method of identifying responder and non-responders to a treatment undergoing clinical trials, comprising detecting biosignatures comprising microRNA in subjects enrolled in the clinical trial, and identifying biosignatures that distinguish between responders and non-responders. In a further embodiment, the biosignatures are measured in a drug naive subject and used to predict whether the subject will be a responder or non-responder. The prediction can be based upon whether the biosignatures of the drug naive subject correlate more closely with the clinical trial subjects identified as responders, thereby predicting that the drug naive subject will be a responder. Conversely, if the biosignatures of the drug naive subject correlate more closely with the clinical trial subjects identified as non-responders, the methods of the invention can predict that the drug naive subject will be a non-responder. The prediction can

therefore be used to stratify potential responders and non-responders to the treatment. In some embodiments, the prediction is used to guide a course of treatment, e.g., by helping treating physicians decide whether to administer the drug. In some embodiments, the prediction is used to guide selection of patients for enrollment in further clinical trials. In a non-limiting example, biosignatures that predict responder / non-responder status in Phase II trials can be used to select patients for a Phase III trial, thereby increasing the likelihood of response in the Phase III patient population. One of skill will appreciate that the method can be adapted to identify biosignatures to stratify subjects on criteria other than responder / non-responder status. In one embodiment, the criterion is treatment safety. Therefore the method is followed as above to identify subjects who are likely or not to have adverse events to the treatment. In a non-limiting example, biosignatures that predict safety profile in Phase II trials can be used to select patients for a Phase III trial, thereby increasing the treatment safety profile in the Phase III patient population.

[00212] The biosignature of the invention can comprise biomarkers in addition to microRNA. As an illustrative example, in colon cancer vesicles are typically shed from colon cancer cells and can be isolated from the peripheral blood and used to isolate one or more biomarkers, e.g., KRAS mRNA which can then be sequenced to detect KRAS mutations. In the case of mRNA biomarkers, the mRNA can be reverse transcribed into cDNA and sequenced (e.g., by Sanger sequencing or high throughput sequencing methods) to determine if there are mutations present that confer resistance to a drug (e.g., resistance to cetuximab or panitumimab). In another example, vesicles that are specifically shed from lung cancer cells are isolated from a biological sample and used to isolate a lung cancer biomarker, e.g., EGFR mRNA. The EGFR mRNA is processed to cDNA and sequenced to determine if there are EGFR mutations present that show resistance or response to specific drugs or treatments for lung cancer.

[00213] Also disclosed herein are methods of conducting retrospective analysis on samples (e.g., serum and tissue biobanks) for the purpose of correlating qualitative and quantitative properties, such as biosignatures of microRNAs, with clinical outcomes in terms of disease state, disease stage, progression, prognosis; therapeutic efficacy or selection; or physiological conditions. Furthermore, methods and compositions disclosed herein are useful for conducting prospective analysis on a sample (e.g., serum and/or tissue collected from individuals in a clinical trial) for the purpose of correlating biosignatures with clinical outcomes in terms of disease state, disease stage, progression, prognosis; therapeutic efficacy or selection; or physiological conditions can also be performed.

[00214] The biosignatures used to characterize a phenotype according to the invention can comprise multiple components (e.g., microRNA, vesicles or other biomarkers) or characteristics (e.g., vesicle size or morphology). The biosignatures can comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 40, 50, 75, or 100 components or characteristics. A biosignature with more than one component or characteristic, such as at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 40, 50, 75, or 100 components, may provide higher sensitivity and/or specificity in characterizing a phenotype. In some embodiments, assessing a plurality of components or characteristics provides increased sensitivity and/or specificity as compared to assessing fewer components or characteristics. On the other hand, it is often desirable to use the fewest number of components or characteristics sufficient to make a correct medical judgment. Fewer markers can avoid statistical overfitting of a classifier and can prevent a delay in treatment

pending further analysis as well inappropriate use of time and resources. Thus, the methods of the invention comprise determining an optimal number of components or characteristics.

[00215] The biosignatures can also be used to build a classifier to classify a sample as belonging to a group, such as belonging to a group having a disease or not, a group having an aggressive disease or not, or a group of responders or non-responders. In one embodiment, a classifier is used to determine whether a subject has an aggressive or non-aggressive cancer. In the case of prostate cancer, this can help a physician to determine whether to watch the cancer, i.e., prescribe "watchful waiting," or perform a prostatectomy. In another embodiment, a classifier is used to determine whether a breast cancer patient is likely to respond or not to tamoxifen, thereby helping the physician to determine whether or not to treat the patient with tamoxifen or another drug.

[00216] As described herein, a biosignature used to characterize a phenotype can comprise one or more biomarkers. The biomarker can be a circulating marker, a membrane associated marker, or a component present within a vesicle or on a vesicle's surface. These biomarkers include without limitation a nucleic acid (e.g. RNA (mRNA, miRNA, etc.) or DNA), protein, peptide, polypeptide, antigen, lipid, carbohydrate, or proteoglycan.

[00217] The biosignature can include the presence or absence, expression level, mutational state, genetic variant state, or any modification (such as epigenetic modification or post-translation modification) of a biomarker (e.g. any one or more biomarker listed in **Table 2**). The expression level of a biomarker can be compared to a control or reference, to determine the overexpression or underexpression (or upregulation or downregulation) of a biomarker in a sample. In some embodiments, the control or reference level comprises the amount of a same biomarker, such as a miRNA, in a control sample from a subject that does not have or exhibit the condition or disease. In another embodiment, the control of reference levels comprises that of a housekeeping marker whose level is minimally affected, if at all, in different biological settings such as diseased versus non-diseased states. In yet another embodiment, the control or reference level comprises that of the level of the same marker in the same subject but in a sample taken at a different time point. Other types of controls are described herein.

[00218] Nucleic acid biomarkers include various RNA or DNA species. For example, the biomarker can be mRNA, microRNA (miRNA), small nucleolar RNAs (snoRNA), small nuclear RNAs (snRNA), ribosomal RNAs (rRNA), heterogeneous nuclear RNA (hnRNA), ribosomal RNAS (rRNA), siRNA, transfer RNAs (tRNA), or shRNA. The DNA can be double-stranded DNA, single stranded DNA, complementary DNA, or noncoding DNA. miRNAs are short ribonucleic acid (RNA) molecules which average about 22 nucleotides long. miRNAs act as post-transcriptional regulators that bind to complementary sequences in the three prime untranslated regions (3' UTRs) of target messenger RNA transcripts (mRNAs), which can result in gene silencing. One miRNA may act upon 1000s of mRNAs. miRNAs play multiple roles in negative regulation, e.g., transcript degradation and sequestering, translational suppression, and may also have a role in positive regulation, e.g., transcriptional and translational activation. By affecting gene regulation, miRNAs can influence many biologic processes. Different sets of expressed miRNAs are found in different cell types and tissues.

[00219] Biomarkers for use with the invention further include peptides, polypeptides, or proteins, which terms are used interchangeably throughout unless otherwise noted. In some embodiments, the protein biomarker comprises its modification state, truncations, mutations, expression level (such as overexpression or

underexpression as compared to a reference level), and/or post-translational modifications, such as described above. In a non-limiting example, a biosignature for a disease can include a protein having a certain post-translational modification that is more prevalent in a sample associated with the disease than without.

[00220] A biosignature may include a number of the same type of biomarkers (e.g., two different microRNA species) or one or more of different types of biomarkers (e.g. mRNAs, miRNAs, proteins, peptides, ligands, and antigens).

[00221] The one or more biomarkers can be detected using a probe. A probe can comprise an oligonucleotide, such as DNA or RNA, an aptamer, monoclonal antibody, polyclonal antibody, Fabs, Fab', single chain antibody, synthetic antibody, peptoid, zDNA, peptide nucleic acid (PNA), locked nucleic acid (LNA), lectin, synthetic or naturally occurring chemical compound (including but not limited to a drug or labeling reagent), dendrimer, or a combination thereof. The probe can be directly detected, for example by being directly labeled, or be indirectly detected, such as through a labeling reagent. The probe can selectively recognize a biomarker. For example, a probe that is an oligonucleotide can selectively hybridize to a miRNA biomarker.

[00222] In aspects, the invention provides for the diagnosis, theranosis, prognosis, disease stratification, disease staging, treatment monitoring or predicting responder / non-responder status of a disease or disorder in a subject. The invention comprises assessing vesicles from a subject, including assessing biomarkers present on the vesicles and/or assessing payload within the vesicles, such as protein, nucleic acid or other biological molecules. Any appropriate biomarker that can be assessed using a vesicle and that relates to a disease or disorder can be used the carry out the methods of the invention. Furthermore, any appropriate technique to assess a vesicle as described herein can be used.

[00223] The methods of the invention comprise a method for characterizing a phenotype in a biological sample, comprising: i) identifying a biosignature in the biological sample, wherein the biosignature comprises a level of one or more microRNA selected from the group consisting of miR-548c-5p, miR-362-3p, miR-422a, miR-597, miR-429, miR-200a, and miR-200b; and ii) comparing the biosignature to a reference, thereby characterizing the phenotype. The phenotype can be a cancer of the gastrointestinal tract. The biosignature for GI and colorectal cancers can include any other informative biomarkers, including vesicle associated biomarkers, such as those disclosed herein. One of skill will appreciate that numerous other vesicle associated biomarkers can be used to create a biosignature for colorectal cancer and other diseases or disorders of the GI tract in addition to those specifically described here, e.g., those disclosed in U.S. Patent Application No. 12/591,226.

Biomarker Detection

[00224] A biosignature can be detected qualitatively or quantitatively by detecting a presence, level or concentration of a microRNA, vesicle or other biomarkers, as disclosed herein. These biosignature components can be detected using a number of techniques known to those of skill in the art. For example, a biomarker can be detected by microarray analysis, polymerase chain reaction (PCR) (including PCR-based methods such as real time polymerase chain reaction (RT-PCR), quantitative real time polymerase chain reaction (Q-PCR/qPCR) and the like), hybridization with allele-specific probes, enzymatic mutation detection, ligation chain reaction (LCR), oligonucleotide ligation assay (OLA), flow-cytometric heteroduplex analysis, chemical cleavage of mismatches, mass spectrometry, nucleic acid sequencing, single strand conformation polymorphism (SSCP), denaturing gradient gel electrophoresis (DGGE), temperature gradient gel electrophoresis (TGGE), restriction fragment polymorphisms, serial analysis of gene expression (SAGE), or combinations thereof. A biomarker,

such as a nucleic acid, can be amplified prior to detection. A biomarker can also be detected by immunoassay, immunoblot, immunoprecipitation, enzyme-linked immunosorbent assay (ELISA; EIA), radioimmunoassay (RIA), flow cytometry, or electron microscopy (EM).

[00225] Biosignatures can be detected using capture agents and detection agents, as described herein. A capture agent can comprise an antibody, aptamer or other entity which recognizes a biomarker and can be used for capturing the biomarker. Biomarkers that can be captured include circulating biomarkers, e.g., a protein, nucleic acid, lipid or biological complex in solution in a bodily fluid. Similarly, the capture agent can be used for capturing a vesicle. A detection agent can comprise an antibody or other entity which recognizes a biomarker and can be used for detecting the biomarker vesicle, or which recognizes a vesicle and is useful for detecting a vesicle. In some embodiments, the detection agent is labeled and the label is detected, thereby detecting the biomarker or vesicle. In many cases, the antigen or other vesicle-moiety that is recognized by the capture and detection agents are interchangeable. As a non-limiting example, consider a vesicle having a cell-of-origin specific antigen on its surface and a cancer-specific antigen on its surface. In one instance, the vesicle can be captured using an antibody to the cell-of-origin specific antigen, e.g., by tethering the capture antibody to a substrate, and then the vesicle is detected using an antibody to the cancer-specific antigen, e.g., by labeling the detection antibody with a fluorescent dye and detecting the fluorescent radiation emitted by the dye. In another instance, the vesicle can be captured using an antibody to the cancer specific antigen, e.g., by tethering the capture antibody to a substrate, and then the vesicle is detected using an antibody to the cell-of-origin specific antigen, e.g., by labeling the detection antibody with a fluorescent dye and detecting the fluorescent radiation emitted by the dye.

[00226] In some embodiments, a same biomarker is recognized by both a capture agent and a detection agent. This scheme can be used depending on the setting. In one embodiment, the biomarker is sufficient to detect a vesicle of interest, e.g., to capture cell-of-origin specific vesicles. In other embodiments, the biomarker is multifunctional, e.g., having both cell-of-origin specific and cancer specific properties. The biomarker can be used in concert with other biomarkers for capture and detection as well.

[00227] One method of detecting a biomarker comprises purifying or isolating a heterogeneous population of vesicles from a biological sample, as described above, and performing a sandwich assay. A vesicle in the population can be captured with a capture agent. The capture agent can be a capture antibody, such as a primary antibody. The capture antibody can be bound to a substrate, for example an array, well, or particle. The captured or bound vesicle can be detected with a detection agent, such as a detection antibody. For example, the detection antibody can be for an antigen of the vesicle. The detection antibody can be directly labeled and detected. Alternatively, the detection agent can be indirectly labeled and detected, such as through an enzyme linked secondary antibody that can react with the detection agent. A detection reagent or detection substrate can be added and the reaction detected, such as described in PCT Publication No. WO2009092386. In an illustrative example wherein the capture agent binds Rab-5b and the detection agent binds or detects CD63 or caveolin-1, the capture agent can be an anti-Rab 5b antibody and the detection agent can be an anti-CD63 or anti-caveolin-1 antibody. In some embodiments, the capture agent binds CD9, PSCA, TNFR, CD63, B7H3, MFG-E8, EpCam, Rab, CD81, STEAP, PCSA, PSMA, or 5T4. For example, the capture agent can be an antibody to CD9, PSCA, TNFR, CD63, B7H3, MFG-E8, EpCam, Rab, CD81, STEAP, PCSA, PSMA, or 5T4. The capture agent can also be an antibody to DR3, STEAP, epha2, TMEM211, unc93A, A33, CD24, NGAL,

EpCam, MUC17, TROP2, or TETS. The detection agent can be an agent that binds or detects CD63, CD9, CD81, B7H3, or EpCam, such as a detection antibody to CD63, CD9, CD81, B7H3, or EpCam. Various combinations of capture and/or detection agents can be used in concert. In an embodiment, the capture agents comprise PCSA, PSMA, B7H3 and optionally EpCam. The detection agents can be one or more tetraspanin such CD9, CD63 and CD81. Increasing numbers of such general vesicle markers can improve the detection signal in some cases. Proteins or other circulating biomarkers can also be detected using sandwich approaches. The captured vesicles can be collected and used to isolate microRNAs contained therein.

[00228] In some embodiments, the capture agent binds or targets EpCam, B7H3 or CD24, and the one or more biomarkers detected on the vesicle are CD9 and/or CD63. In one embodiment, the capture agent binds or targets EpCam, and the one or more biomarkers detected on the vesicle are CD9, EpCam and/or CD81. The single capture agent can be selected from CD9, PSCA, TNFR, CD63, B7H3, MFG-E8, EpCam, Rab, CD81, STEAP, PCSA, PSMA, or 5T4. The single capture agent can also be an antibody to DR3, STEAP, epha2, TMEM211, unc93A, A33, CD24, NGAL, EpCam, MUC17, TROP2, or TETS. In some embodiments, the single capture agent is selected from PCSA, PSMA, B7H3, CD81, CD9 and CD63.

[00229] In other embodiments, the capture agent targets PCSA, and the one or more biomarkers detected on the captured vesicle are B7H3 and/or PSMA. In other embodiments, the capture agent targets PSMA, and the one or more biomarkers detected on the captured vesicle are B7H3 and/or PCSA. In other embodiments, the capture agent targets B7H3, and the one or more biomarkers detected on the captured vesicle are PSMA and/or PCSA. In yet other embodiments, the capture agent targets CD63 and the one or more biomarkers detected on the vesicle are CD81, CD83, CD9 and/or CD63. The different capture agent and biomarker combinations disclosed herein can be used to characterize a phenotype, such as detecting, diagnosing or prognosing a disease, e.g., a cancer. In some embodiments, vesicles are analyzed to characterize prostate cancer using a capture agent targeting EpCam and detection of CD9 and CD63; a capture agent targeting PCSA and detection of B7H3 and PSMA; or a capture agent of CD63 and detection of CD81. In other embodiments, vesicles are used to characterize colon cancer using capture agent targeting CD63 and detection of CD63, or a capture agent targeting CD9 coupled with detection of CD63. One of skill will appreciate that targets of capture agents and detection agents can be used interchangeably. In an illustrative example, consider a capture agent targeting PCSA and detection agents targeting B7H3 and PSMA. Because all of these markers are useful for detecting PCa derived vesicles, B7H3 or PSMA could be targeted by the capture agent and PCSA could be recognized by a detection agent. For example, in some embodiments, the detection agent targets PCSA, and one or more biomarkers used to capture the vesicle comprise B7H3 and/or PSMA. In other embodiments, the detection agent targets PSMA, and the one or more biomarkers used to capture the vesicle comprise B7H3 and/or PCSA. In other embodiments, the detection agent targets B7H3, and the one or more biomarkers used to capture the vesicle comprise PSMA and/or PCSA. In some embodiments, the invention provides a method of detecting prostate cancer cells in bodily fluid using capture agents and/or detection agents to PSMA, B7H3 and/or PCSA. The bodily fluid can comprise blood, including serum or plasma. The bodily fluid can comprise ejaculate or sperm. In further embodiments, the methods of detecting prostate cancer further use capture agents and/or detection agents to CD81, CD83, CD9 and/or CD63. The method further provides a method of characterizing a GI disorder, comprising capturing vesicles with one or more of DR3, STEAP, epha2, TMEM211, unc93A, A33, CD24, NGAL, EpCam, MUC17, TROP2, and TETS, and detecting the captured vesicles with one or more

general vesicle antigen, such as CD81, CD63 and/or CD9. Additional agents can improve the test performance, e.g., improving test accuracy or AUC, either by providing additional biological discriminatory power and/or by reducing experimental noise.

[00230] Techniques of detecting biomarkers for use with the invention include the use of a planar substrate such as an array (e.g., biochip or microarray), with molecules immobilized to the substrate as capture agents that facilitate the detection of a particular biosignature. The array can be provided as part of a kit for assaying one or more biomarkers or vesicles. A molecule that identifies the biomarkers of interest, such as the antigens in **Table 2**, can be included in an array for detection and diagnosis of diseases including presymptomatic diseases. In some embodiments, an array comprises a custom array comprising biomolecules selected to specifically identify biomarkers of interest. Customized arrays can be modified to detect biomarkers that increase statistical performance, e.g., additional biomolecules that identifies a biosignature which lead to improved cross-validated error rates in multivariate prediction models (e.g., logistic regression, discriminant analysis, or regression tree models). In some embodiments, customized array(s) are constructed to study the biology of a disease, condition or syndrome and profile biosignatures in defined physiological states. Markers for inclusion on the customized array be chosen based upon statistical criteria, e.g., having a desired level of statistical significance in differentiating between phenotypes or physiological states. In some embodiments, standard significance of p-value = 0.05 is chosen to exclude or include biomolecules on the microarray. The p-values can be corrected for multiple comparisons. As an illustrative example, nucleic acids extracted from samples from a subject with or without a disease can be hybridized to a high density microarray that binds to thousands of gene sequences. Nucleic acids whose levels are significantly different between the samples with or without the disease can be selected as biomarkers to distinguish samples as having the disease or not. A customized array can be constructed to detect the selected biomarkers. In some embodiments, customized arrays comprise low density microarrays, which refer to arrays with lower number of addressable binding agents, e.g., tens or hundreds instead of thousands. Low density arrays can be formed on a substrate. In some embodiments, customizable low density arrays use PCR amplification in plate wells, e.g., TaqMan® Gene Expression Assays (Applied Biosystems by Life Technologies Corporation, Carlsbad, CA).

[00231] A planar array generally contains addressable locations (e.g., pads, addresses, or micro-locations) of biomolecules in an array format. The size of the array will depend on the composition and end use of the array. Arrays can be made containing from 2 different molecules to many thousands. Generally, the array comprises from two to as many as 100,000 or more molecules, depending on the end use of the array and the method of manufacture. A microarray for use with the invention comprises at least one biomolecule that identifies or captures a biomarker present in a biosignature of interest, e.g., a microRNA or other biomolecule or vesicle that makes up the biosignature. In some arrays, multiple substrates are used, either of different or identical compositions. Accordingly, planar arrays may comprise a plurality of smaller substrates.

[00232] The present invention can make use of many types of arrays for detecting a biomarker, e.g., a biomarker associated with a biosignature of interest. Useful arrays or microarrays include without limitation DNA microarrays, such as cDNA microarrays, oligonucleotide microarrays and SNP microarrays, microRNA arrays, protein microarrays, antibody microarrays, tissue microarrays, cellular microarrays (also called transfection microarrays), chemical compound microarrays, and carbohydrate arrays (glycoarrays). These arrays are described in more detail above. In some embodiments, microarrays comprise biochips that provide

high-density immobilized arrays of recognition molecules (e.g., antibodies), where biomarker binding is monitored indirectly (e.g., via fluorescence). **FIG. 2A** shows an illustrative configuration in which capture antibodies against a vesicle antigen of interest are tethered to a surface. The captured vesicles are then detected using detector antibodies against the same or different vesicle antigens of interest. The capture antibodies can be substituted with tethered aptamers as available and desirable. Fluorescent detectors are shown. Other detectors can be used similarly, e.g., enzymatic reaction, detectable nanoparticles, radiolabels, and the like. In other embodiments, an array comprises a format that involves the capture of proteins by biochemical or intermolecular interaction, coupled with detection by mass spectrometry (MS). The vesicles can be eluted from the surface and the payload therein, e.g., microRNA, can be analyzed.

[00233] An array or microarray that can be used to detect one or more biomarkers of a biosignature can be made according to the methods described in U.S. Pat. Nos. 6,329,209; 6,365,418; 6,406,921; 6,475,808; and 6,475,809, and U.S. Patent Application Ser. No. 10/884,269, each of which is herein incorporated by reference in its entirety. Custom arrays to detect specific selections of sets of biomarkers described herein can be made using the methods described in these patents. Commercially available microarrays can also be used to carry out the methods of the invention, including without limitation those from Affymetrix (Santa Clara, CA), Illumina (San Diego, CA), Agilent (Santa Clara, CA), Exiqon (Denmark), or Invitrogen (Carlsbad, CA). Custom and/or commercial arrays include arrays for detection proteins, nucleic acids, and other biological molecules and entities (e.g., cells, vesicles, virii) as described herein.

[00234] In some embodiments, molecules to be immobilized on an array comprise proteins or peptides. One or more types of proteins may be immobilized on a surface. In certain embodiments, the proteins are immobilized using methods and materials that minimize the denaturing of the proteins, that minimize alterations in the activity of the proteins, or that minimize interactions between the protein and the surface on which they are immobilized.

[00235] Array surfaces useful may be of any desired shape, form, or size. Non-limiting examples of surfaces include chips, continuous surfaces, curved surfaces, flexible surfaces, films, plates, sheets, or tubes. Surfaces can have areas ranging from approximately a square micron to approximately 500 cm^2. The area, length, and width of surfaces may be varied according to the requirements of the assay to be performed. Considerations may include, for example, ease of handling, limitations of the material(s) of which the surface is formed, requirements of detection systems, requirements of deposition systems (e.g., arrayers), or the like.

[00236] In certain embodiments, it is desirable to employ a physical means for separating groups or arrays of binding islands or immobilized biomolecules: such physical separation facilitates exposure of different groups or arrays to different solutions of interest. Therefore, in certain embodiments, arrays are situated within microwell plates having any number of wells. In such embodiments, the bottoms of the wells may serve as surfaces for the formation of arrays, or arrays may be formed on other surfaces and then placed into wells. In certain embodiments, such as where a surface without wells is used, binding islands may be formed or molecules may be immobilized on a surface and a gasket having holes spatially arranged so that they correspond to the islands or biomolecules may be placed on the surface. Such a gasket is preferably liquid tight. A gasket may be placed on a surface at any time during the process of making the array and may be removed if separation of groups or arrays is no longer necessary.

[00237] In some embodiments, the immobilized molecules can bind to one or more biomarkers or vesicles present in a biological sample contacting the immobilized molecules. In some embodiments, the immobilized molecules modify or are modified by molecules present in the one or more vesicles contacting the immobilized molecules. Contacting the sample typically comprises overlaying the sample upon the array.

[00238] Modifications or binding of molecules in solution or immobilized on an array can be detected using detection techniques known in the art. Examples of such techniques include immunological techniques such as competitive binding assays and sandwich assays; fluorescence detection using instruments such as confocal scanners, confocal microscopes, or CCD-based systems and techniques such as fluorescence, fluorescence polarization (FP), fluorescence resonant energy transfer (FRET), total internal reflection fluorescence (TIRF), fluorescence correlation spectroscopy (FCS); colorimetric/spectrometric techniques; surface plasmon resonance, by which changes in mass of materials adsorbed at surfaces are measured; techniques using radioisotopes, including conventional radioisotope binding and scintillation proximity assays (SPA); mass spectroscopy, such as matrix-assisted laser desorption/ionization mass spectroscopy (MALDI) and MALDI-time of flight (TOF) mass spectroscopy; ellipsometry, which is an optical method of measuring thickness of protein films; quartz crystal microbalance (QCM), a very sensitive method for measuring mass of materials adsorbing to surfaces; scanning probe microscopies, such as atomic force microscopy (AFM), scanning force microscopy (SFM) or scanning electron microscopy (SEM); and techniques such as electrochemical, impedance, acoustic, microwave, and IR/Raman detection. See, e.g., *Mere L, et al.,"Miniaturized FRET assays and microfluidics: key components for ultra-high-throughput screening," Drug Discovery Today 4(8):363-369 (1999),* and references cited therein*; Lakowicz J R, Principles of Fluorescence Spectroscopy, 2nd Edition, Plenum Press (1999), or Jain KK: Integrative Omics, Pharmacoproteomics, and Human Body Fluids. In: Thongboonkerd V, ed., ed. Proteomics of Human Body Fluids: Principles, Methods and Applications. Volume 1: Totowa, N.J.: Humana Press, 2007*, each of which is herein incorporated by reference in its entirety.

[00239] Microarray technology can be combined with mass spectroscopy (MS) analysis and other tools. Electrospray interface to a mass spectrometer can be integrated with a capillary in a microfluidics device. For example, one commercially available system contains eTag reporters that are fluorescent labels with unique and well-defined electrophoretic mobilities; each label is coupled to biological or chemical probes via cleavable linkages. The distinct mobility address of each eTag reporter allows mixtures of these tags to be rapidly deconvoluted and quantitated by capillary electrophoresis. This system allows concurrent gene expression, protein expression, and protein function analyses from the same sample *Jain KK: Integrative Omics, Pharmacoproteomics, and Human Body Fluids. In: Thongboonkerd V, ed., ed. Proteomics of Human Body Fluids: Principles, Methods and Applications. Volume 1: Totowa, N.J.: Humana Press, 2007, which* is herein incorporated by reference in its entirety.

[00240] A biochip can include components for a microfluidic or nanofluidic assay. A microfluidic device can be used for isolating or analyzing biomarkers, such as determining a biosignature. Microfluidic systems allow for the miniaturization and compartmentalization of one or more processes for isolating, capturing or detecting a vesicle, detecting a microRNA, detecting a circulating biomarker, detecting a biosignature, and other processes. The microfluidic devices can use one or more detection reagents in at least one aspect of the system, and such a detection reagent can be used to detect one or more biomarkers. In one embodiment, the device detects a biomarker on an isolated or bound vesicle. Various probes, antibodies, proteins, or other binding agents can be

used to detect a biomarker within the microfluidic system. The detection agents may be immobilized in different compartments of the microfluidic device or be entered into a hybridization or detection reaction through various channels of the device.

[00241] A vesicle in a microfluidic device can be lysed and its contents detected within the microfluidic device, such as proteins or nucleic acids, e.g., DNA or RNA such as miRNA or mRNA. The nucleic acid may be amplified prior to detection, or directly detected, within the microfluidic device. Thus microfluidic system can also be used for multiplexing detection of various biomarkers. In an embodiment, vesicles are captured within the microfluidic device, the captured vesicles are lysed, and a biosignature of microRNA from the vesicle payload is determined. The biosignature can further comprise the capture agent used to capture the vesicle.

[00242] Novel nanofabrication techniques are opening up the possibilities for biosensing applications that rely on fabrication of high-density, precision arrays, e.g., nucleotide-based chips and protein arrays otherwise know as heterogeneous nanoarrays. Nanofluidics allows a further reduction in the quantity of fluid analyte in a microchip to nanoliter levels, and the chips used here are referred to as nanochips. (See, e.g., *Unger M et al., Biotechniques 1999; 27(5):1008-14, Kartalov EP et al., Biotechniques 2006; 40(1):85-90*, each of which are herein incorporated by reference in their entireties.) Commercially available nanochips currently provide simple one step assays such as total cholesterol, total protein or glucose assays that can be run by combining sample and reagents, mixing and monitoring of the reaction. Gel-free analytical approaches based on liquid chromatography (LC) and nanoLC separations (*Cutillas et al. Proteomics, 2005;5:101-112 and Cutillas et al., Mol Cell Proteomics 2005;4:1038-1051*, each of which is herein incorporated by reference in its entirety) can be used in combination with the nanochips.

[00243] Further provided herein is a rapid detection device that facilitates the detection of a particular biosignature in a biological sample. The device can integrate biological sample preparation with polymerase chain reaction (PCR) on a chip. The device can facilitate the detection of a particular biosignature of a vesicle in a biological sample, and an example is provided as described in *Pipper et al., Angewandte Chemie, 47(21), p. 3900-3904 (2008)*, which is herein incorporated by reference in its entirety. A biosignature can be incorporated using micro-/nano-electrochemical system (MEMS/NEMS) sensors and oral fluid for diagnostic applications as described in *Li et al., Adv Dent Res 18(1): 3-5 (2005),* which is herein incorporated by reference in its entirety.

[00244] As an alternative to planar arrays, assays using particles, such as bead based assays as described herein, can be used in combination with flow cytometry. Multiparametric assays or other high throughput detection assays using bead coatings with cognate ligands and reporter molecules with specific activities consistent with high sensitivity automation can be used. In a bead based assay system, a binding agent for a biomarker or vesicle, such as a capture agent (e.g. capture antibody), can be immobilized on an addressable microsphere. Each binding agent for each individual binding assay can be coupled to a distinct type of microsphere (i.e., microbead) and the assay reaction takes place on the surface of the microsphere, such as depicted in **FIG. 2B**. A binding agent for a vesicle can be a capture antibody coupled to a bead. Dyed microspheres with discrete fluorescence intensities are loaded separately with their appropriate binding agent or capture probes. The different bead sets carrying different binding agents can be pooled as necessary to generate custom bead arrays. Bead arrays are then incubated with the sample in a single reaction vessel to perform the assay. Examples of microfluidic devices that may be used, or adapted for use with the invention, include but are not limited to those described herein.

[00245] Product formation of the biomarker with an immobilized capture molecule or binding agent can be detected with a fluorescence based reporter system (see for example, **FIG. 2A-B**). The biomarker can either be labeled directly by a fluorophore or detected by a second fluorescently labeled capture biomolecule. The signal intensities derived from captured biomarkers can be measured in a flow cytometer. The flow cytometer can first identify each microsphere by its individual color code. For example, distinct beads can be dyed with discrete fluorescence intensities such that each bead with a different intensity has a different binding agent. The beads can be labeled or dyed with at least 2 different labels or dyes. In some embodiments, the beads are labeled with at least 3, 4, 5, 6, 7, 8, 9, or 10 different labels. The beads with more than one label or dye can also have various ratios and combinations of the labels or dyes. The beads can be labeled or dyed externally or may have intrinsic fluorescence or signaling labels.

[00246] The amount of captured biomarkers on each individual bead can be measured by the second color fluorescence specific for the bound target. This allows multiplexed quantitation of multiple targets from a single sample within the same experiment. Sensitivity, reliability and accuracy are compared or can be improved to standard microtiter ELISA procedures. An advantage of a bead-based system is the individual coupling of the capture biomolecule or binding agent for a vesicle to distinct microspheres provides multiplexing capabilities. For example, as depicted in **FIG. 2C**, a combination of 5 different biomarkers to be detected (detected by antibodies to antigens such as CD63, CD9, CD81, B7H3, and EpCam) and 20 biomarkers for which to capture a vesicle, (using capture antibodies, such as antibodies to CD9, PSCA, TNFR, CD63, B7H3, MFG-E8, EpCam, Rab, CD81, STEAP, PCSA, PSMA, 5T4, and CD24) can result in approximately 100 combinations to be detected. As shown in **FIG. 2C** as "EpCam 2x," "CD63 2X," multiple antibodies to a single target can be used to probe detection against various epitopes. In another example, multiplex analysis comprises capturing a vesicle using a binding agent to CD24 and detecting the captured vesicle using a binding agent for CD9, CD63, and/or CD81. The captured vesicles can be detected using a detection agent such as an antibody. The detection agents can be labeled directly or indirectly, as described herein.

[00247] Multiplexing of at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 50, 75 or 100 different biomarkers may be performed. For example, an assay of a heterogeneous population of vesicles can be performed with a plurality of particles that are differentially labeled. There can be at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 50, 75 or 100 differentially labeled particles. The particles may be externally labeled, such as with a tag, or they may be intrinsically labeled. Each differentially labeled particle can be coupled to a capture agent, such as a binding agent, for a vesicle, resulting in capture of a vesicle. The multiple capture agents can be selected to characterize a phenotype of interest, including capture agents against general vesicle biomarkers, cell-of-origin specific biomarkers, and disease biomarkers. One or more biomarkers of the captured vesicle can then be detected by a plurality of binding agents. The binding agent can be directly labeled to facilitate detection. Alternatively, the binding agent is labeled by a secondary agent. For example, the binding agent may be an antibody for a biomarker on the vesicle. The binding agent is linked to biotin. A secondary agent comprises streptavidin linked to a reporter and can be added to detect the biomarker. In some embodiments, the captured vesicle is assayed for at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 50, 75 or 100 different biomarkers. For example, multiple detectors, i.e., detection of multiple biomarkers of a captured vesicle or population of vesicles, can increase the signal obtained, permitted increased sensitivity, specificity, or both, and the use of smaller amounts of samples. For example, detection with more

than one general vesicle marker can improve the signal as compared to using a lesser number of detection markers, such as a single marker. To illustrate, detection of vesicles with labeled binding agents to two or three of CD9, CD63 and CD81 can improve the signal compared to detection with any one of the tetraspanins individually.

[00248] An immunoassay based method or sandwich assay can also be used to detect a biomarker of a vesicle. An example includes ELISA. A binding agent or capture agent can be bound to a well. For example an antibody to an antigen of a vesicle can be attached to a well. A biomarker on the captured vesicle can be detected based on the methods described herein. **FIG. 2A** shows an illustrative schematic for a sandwich-type of immunoassay. The capture antibody can be against a vesicle antigen of interest, e.g., a general vesicle biomarker, a cell-of-origin marker, or a disease marker. In the figure, the captured vesicles are detected using fluorescently labeled antibodies against vesicle antigens of interest. Multiple capture antibodies can be used, e.g., in distinguishable addresses on an array or different wells of an immunoassay plate. The detection antibodies can be against the same antigen as the capture antibody, or can be directed against other markers. The capture antibodies can be substituted with alternate binding agents, such as tethered aptamers or lectins, and/or the detector antibodies can be similarly substituted, e.g., with detectable (e.g., labeled) aptamers, lectins or other binding proteins or entities. In an embodiment, one or more capture agents to a general vesicle biomarker, a cell-of-origin marker, and/or a disease marker are used along with detection agents against general vesicle biomarker, such as tetraspanin molecules including without limitation one or more of CD9, CD63 and CD81.

[00249] **FIG. 2D** presents an illustrative schematic for analyzing vesicles according to the methods of the invention. Capture agents are used to capture vesicles, detectors are used to detect the captured vesicles, and the level or presence of the captured and detected antibodies is used to characterize a phenotype. Capture agents, detectors and characterizing phenotypes can be any of those described herein. For example, capture agents include antibodies or aptamers tethered to a substrate that recognize a vesicle antigen of interest, detectors include labeled antibodies or aptamers to a vesicle antigen of interest, and characterizing a phenotype includes a diagnosis, prognosis, or theranosis of a disease. In the scheme shown in **FIG. 2D i)**, a population of vesicles is captured with one or more capture agents against general vesicle biomarkers (**200**). The captured vesicles are then labeled with detectors against cell-of-origin biomarkers (**201**) and/or disease specific biomarkers (**202**). If only cell-of-origin detectors are used (**201**), the biosignature used to characterize the phenotype (**203**) can include the general vesicle markers (**200**) and the cell-of-origin biomarkers (**201**). If only disease detectors are used (**202**), the biosignature used to characterize the phenotype (**203**) can include the general vesicle markers (**200**) and the disease biomarkers (**202**). Alternately, detectors are used to detect both cell-of-origin biomarkers (**201**) and disease specific biomarkers (**202**). In this case, the biosignature used to characterize the phenotype (**203**) can include the general vesicle markers (**200**), the cell-of-origin biomarkers (**201**) and the disease biomarkers (**202**). The biomarkers combinations are selected to characterize the phenotype of interest and can be selected from the biomarkers and phenotypes described herein.

[00250] In the scheme shown in **FIG. 2D ii)**, a population of vesicles is captured with one or more capture agents against cell-of-origin biomarkers (**210**) and/or disease biomarkers (**211**). The captured vesicles are then detected using detectors against general vesicle biomarkers (**212**). If only cell-of-origin capture agents are used (**210**), the biosignature used to characterize the phenotype (**213**) can include the cell-of-origin biomarkers (**210**)

and the general vesicle markers (**212**). If only disease biomarker capture agents are used (**211**), the biosignature used to characterize the phenotype (**213**) can include the disease biomarkers (**211**) and the general vesicle biomarkers (**212**). Alternately, capture agents to one or more cell-of-origin biomarkers (**210**) and one or more disease specific biomarkers (**211**) are used to capture vesicles. In this case, the biosignature used to characterize the phenotype (**213**) can include the cell-of-origin biomarkers (**210**), the disease biomarkers (**211**), and the general vesicle markers (**213**). The biomarkers combinations are selected to characterize the phenotype of interest and can be selected from the biomarkers and phenotypes described herein.

[00251] Biomarkers comprising vesicle payload can be analyzed to characterize a phenotype. Payload comprises the biological entities contained within a vesicle membrane. These entities include without limitation nucleic acids, e.g., mRNA, microRNA, or DNA fragments; protein, e.g., soluble and membrane associated proteins; carbohydrates; lipids; metabolites; and various small molecules, e.g., hormones. The payload can be part of the cellular milieu that is encapsulated as a vesicle is formed in the cellular environment. In some embodiments of the invention, the payload is analyzed in addition to detecting vesicle surface antigens. Specific populations of vesicles can be captured as described above then the payload in the captured vesicles can be used to characterize a phenotype. For example, vesicles captured on a substrate can be further isolated to assess the payload therein. Alternately, the vesicles in a sample are detected and sorted without capture. The vesicles so detected can be further isolated to assess the payload therein. In an embodiment, vesicle populations are sorted by flow cytometry and the payload in the sorted vesicles is analyzed. In the scheme shown in **FIG. 2E iii)**, a population of vesicles is captured and/or detected (**230**) using one or more of cell-of-origin biomarkers (**220**), disease biomarkers (**221**), and general vesicle markers (**222**). The payload of the isolated vesicles is assessed (**223**). A biosignature detected within the payload can be used to characterize a phenotype (**224**). In a non-limiting example, a vesicle population can be analyzed in a plasma sample from a patient using antibodies against one or more vesicle antigens of interest. The antibodies can be capture antibodies which are tethered to a substrate to isolate a desired vesicle population. Alternately, the antibodies can be directly labeled and the labeled vesicles isolated by sorting with flow cytometry. The presence or level of microRNA or mRNA extracted from the isolated vesicle population can be used to detect a biosignature. The biosignature is then used to diagnose, prognose or theranose the patient.

[00252] In other embodiments, vesicle payload is analyzed in a vesicle population without first capturing or detected subpopulations of vesicles. For example, vesicles can be generally isolated from a sample using centrifugation, filtration, chromatography, or other techniques as described herein. The payload of the isolated vesicles can be analyzed thereafter to detect a biosignature and characterize a phenotype. In the scheme shown in **FIG. 2E iv)**, a population of vesicles is isolated (**230**) and the payload of the isolated vesicles is assessed (**231**). A biosignature detected within the payload can be used to characterize a phenotype (**232**). In a non-limiting example, a vesicle population is isolated from a plasma sample from a patient using size exclusion and membrane filtration. The presence or level of microRNA or mRNA extracted from the vesicle population is used to detect a biosignature. The biosignature is then used to diagnose, prognose or theranose the patient.

[00253] A peptide or protein biomarker can be analyzed by mass spectrometry or flow cytometry. Proteomic analysis of a vesicle may be carried out by immunocytochemical staining, Western blotting, electrophoresis, SDS-PAGE, chromatography, x-ray crystallography or other protein analysis techniques in accordance with procedures well known in the art. In other embodiments, the protein biosignature of a vesicle may be analyzed

using 2 D differential gel electrophoresis as described in, *Chromy et al. J Proteome Res, 2004;3:1120-1127,* which is herein incorporated by reference in its entirety, or with liquid chromatography mass spectrometry as described in *Zhang et al. Mol Cell Proteomics, 2005;4:144-155*, which is herein incorporated by reference in its entirety. A vesicle may be subjected to activity-based protein profiling described for example, in *Berger et al., Am J Pharmacogenomics, 2004;4:371-381*, which is in incorporated by reference in its entirety. In other embodiments, a vesicle may be profiled using nanospray liquid chromatography-tandem mass spectrometry as described in *Pisitkun et al., Proc Natl Acad Sci U S A, 2004; 101:13368-13373,* which is herein incorporated by reference in its entirety. In another embodiment, the vesicle may be profiled using tandem mass spectrometry (MS) such as liquid chromatography/MS/MS (LC-MS/MS) using for example a LTQ and LTQ-FT ion trap mass spectrometer. Protein identification can be determined and relative quantitation can be assessed by comparing spectral counts as described in *Smalley et al., J Proteome Res, 2008;7:2088-2096*, which is herein incorporated by reference in its entirety.

[00254] The expression of circulating protein biomarkers or protein payload within a vesicle can also be identified. The latter analysis can optionally follow the isolation of specific vesicles using capture agents to capture populations of interest. In an embodiment, immunocytochemical staining is used to analyze protein expression. The sample can be resuspended in buffer, centrifuged at 100 x g for example, for 3 minutes using a cytocentrifuge on adhesive slides in preparation for immunocytochemical staining. The cytospins can be air-dried overnight and stored at -80°C until staining. Slides can then be fixed and blocked with serum-free blocking reagent. The slides can then be incubated with a specific antibody to detect the expression of a protein of interest. In some embodiments, the vesicles are not purified, isolated or concentrated prior to protein expression analysis.

[00255] Biosignatures comprising vesicle payload can be characterized by analysis of a metabolite marker or metabolite within the vesicle. Various metabolite-oriented approaches have been described such as metabolite target analyses, metabolite profiling, or metabolic fingerprinting, see for example, *Denkert et al., Molecular Cancer 2008; 7: 4598-4617, Ellis et al., Analyst 2006; 8: 875-885, Kuhn et al., Clinical Cancer Research 2007; 24: 7401-7406, Fiehn O., Comp Funct Genomics 2001;2:155-168, Fancy et al., Rapid Commun Mass Spectrom 20(15): 2271-80 (2006), Lindon et al., Pharm Res, 23(6): 1075-88 (2006), Holmes et al., Anal Chem. 2007 Apr 1;79(7):2629-40. Epub 2007 Feb 27. Erratum in: Anal Chem. 2008 Aug 1;80(15):6142-3, Stanley et al., Anal Biochem. 2005 Aug 15;343(2):195-202., Lehtimäki et al., J Biol Chem. 2003 Nov 14;278(46):45915-23*, each of which is herein incorporated by reference in its entirety.

[00256] Peptides can be analyzed by systems described in *Jain KK: Integrative Omics, Pharmacoproteomics, and Human Body Fluids. In: Thongboonkerd V, ed., ed. Proteomics of Human Body Fluids: Principles, Methods and Applications. Volume 1: Totowa, N.J.: Humana Press, 2007,* which is herein incorporated by reference in its entirety. This system can generate sensitive molecular fingerprints of proteins present in a body fluid as well as in vesicles. Commercial applications which include the use of chromatography/mass spectroscopy and reference libraries of all stable metabolites in the human body, for example Paradigm Genetic's Human Metabolome Project, may be used to determine a metabolite biosignature. Other methods for analyzing a metabolic profile can include methods and devices described in U.S. Patent No. 6,683,455 (Metabometrix), U.S. Patent Application Publication Nos. 20070003965 and 20070004044 (Biocrates Life Science), each of which is herein incorporated by reference in its entirety. Other proteomic profiling techniques are described in *Kennedy,*

Toxicol Lett 120:379-384 (2001), Berven et al., Curr Pharm Biotechnol 7(3): 147-58 (2006), Conrads et al., Expert Rev Proteomics 2(5): 693-703, Decramer et al., World J Urol 25(5): 457-65 (2007), Decramer et al., Mol Cell Proteomics 7(10): 1850-62 (2008), Decramer et al., Contrib Nephrol, 160: 127-41 (2008), Diamandis, J Proteome Res 5(9): 2079-82 (2006), Immler et al., Proteomics 6(10): 2947-58 (2006), Khan et al., J Proteome Res 5(10): 2824-38 (2006), Kumar et al., Biomarkers 11(5): 385-405 (2006), Noble et al., Breast Cancer Res Treat 104(2): 191-6 (2007), Omenn, Dis Markers 20(3): 131-4 (2004), Powell et al., Expert Rev Proteomics 3(1): 63-74 (2006), Rai et al., Arch Pathol Lab Med, 126(12): 1518-26 (2002), Ramstrom et al., Proteomics, 3(2): 184-90 (2003), Tammen et al., Breast Cancer Res Treat, 79(1): 83-93 (2003), Theodorescu et al., Lancet Oncol, 7(3): 230-40 (2006), or Zurbig et al., Electrophoresis, 27(11): 2111-25 (2006).

[00257] For analysis of mRNAs, miRNAs or other small RNAs, the total RNA can be isolated using any known methods for isolating nucleic acids such as methods described in U.S. Patent Application Publication No. 2008132694, which is herein incorporated by reference in its entirety. These include, but are not limited to, kits for performing membrane based RNA purification, which are commercially available. Generally, kits are available for the small-scale (30 mg or less) preparation of RNA from cells and tissues, for the medium scale (250 mg tissue) preparation of RNA from cells and tissues, and for the large scale (1 g maximum) preparation of RNA from cells and tissues . Other commercially available kits for effective isolation of small RNA-containing total RNA are available. Such methods can be used to isolate nucleic acids from vesicles.

[00258] Alternatively, RNA can be isolated using the method described in U.S. Patent No. 7,267,950, which is herein incorporated by reference in its entirety. U.S. Patent No. 7,267,950 describes a method of extracting RNA from biological systems (cells, cell fragments, organelles, tissues, organs, or organisms) in which a solution containing RNA is contacted with a substrate to which RNA can bind and RNA is withdrawn from the substrate by applying negative pressure. Alternatively, RNA may be isolated using the method described in U.S. Patent Application No. 20050059024, which is herein incorporated by reference in its entirety, which describes the isolation of small RNA molecules. Other methods are described in U.S. Patent Application No. 20050208510, 20050277121, 20070238118, each of which is incorporated by reference in its entirety.

[00259] In one embodiment, mRNA expression analysis can be carried out on mRNAs from a vesicle isolated from a sample. In some embodiments, the vesicle is a cell-of-origin specific vesicle. An expression pattern generated from a vesicle can be indicative of a given disease state, disease stage, therapy related signature, or physiological condition.

[00260] In one embodiment, once the total RNA has been isolated, cDNA can be synthesized and either qRT-PCR assays (e.g. Applied Biosystem's Taqman® assays) for specific mRNA targets can be performed according to manufacturer's protocol, or an expression microarray can be performed to look at highly multiplexed sets of expression markers in one experiment. Methods for establishing gene expression profiles include determining the amount of RNA that is produced by a gene that can code for a protein or peptide. This can be accomplished by quantitative reverse transcriptase PCR (qRT-PCR), competitive RT-PCR, real time RT-PCR, differential display RT-PCR, Northern Blot analysis or other related tests. While it is possible to conduct these techniques using individual PCR reactions, it is also possible to amplify complementary DNA (cDNA) or complementary RNA (cRNA) produced from mRNA and analyze it via microarray.

[00261] The level of a miRNA product in a sample can be measured using any appropriate technique that is suitable for detecting mRNA expression levels in a biological sample, including but not limited to Northern blot

analysis, RT-PCR, qRT-PCR, in situ hybridization or microarray analysis. For example, using gene specific primers and target cDNA, qRT-PCR enables sensitive and quantitative miRNA measurements of either a small number of target miRNAs (via singleplex and multiplex analysis) or the platform can be adopted to conduct high throughput measurements using 96-well or 384-well plate formats. See for example, *Ross JS et al, Oncologist. 2008 May;13(5):477-93,* which is herein incorporated by reference in its entirety. A number of different array configurations and methods for microarray production are known to those of skill in the art and are described in U.S. patents such as: U.S. Pat. Nos. 5,445,934; 5,532,128; 5,556,752; 5,242,974; 5,384,261; 5,405,783; 5,412,087; 5,424,186; 5,429,807; 5,436,327; 5,472,672; 5,527,681; 5,529,756; 5,545,531; 5,554,501; 5,561,071; 5,571,639; 5,593,839; 5,599,695; 5,624,711; 5,658,734; or 5,700,637; each of which is herein incorporated by reference in its entirety. Other methods of profiling miRNAs are described in *Taylor et al., Gynecol Oncol. 2008 Jul;110(1):13-21, Gilad et al, PLoS ONE. 2008 Sep 5;3(9):e3148, Lee et al., Annu Rev Pathol. 2008 Sep 25 and Mitchell et al, Proc Natl Acad Sci U S A. 2008 Jul 29;105(30):10513-8, Shen R et al, BMC Genomics. 2004 Dec 14;5(1):94, Mina L et al, Breast Cancer Res Treat. 2007 Jun;103(2):197-208, Zhang L et al, Proc Natl Acad Sci U S A. 2008 May 13;105(19):7004-9, Ross JS et al, Oncologist. 2008 May;13(5):477-93, Schetter AJ et al, JAMA. 2008 Jan 30;299(4):425-36, Staudt LM, N Engl J Med 2003;348:1777-85, Mulligan G et al, Blood. 2007 Apr 15;109(8):3177-88. Epub 2006 Dec 21, McLendon R et al, Nature. 2008 Oct 23;455(7216):1061-8,* and U.S. Patent Nos. 5,538,848, 5,723,591, 5,876,930, 6,030,787, 6,258,569, and 5,804,375, each of which is herein incorporated by reference. In some embodiments, arrays of microRNA panels are use to simultaneously query the expression of multiple miRs. The Exiqon mIRCURY LNA microRNA PCR system panel (Exiqon, Inc., Woburn, MA) or the TaqMan® MicroRNA Assays and Arrays systems from Applied Biosystems (Foster City, CA) can be used for such purposes.

[00262] Microarray technology allows for the measurement of the steady-state mRNA or miRNA levels of thousands of transcripts or miRNAs simultaneously thereby presenting a powerful tool for identifying effects such as the onset, arrest, or modulation of uncontrolled cell proliferation. Two microarray technologies, such as cDNA arrays and oligonucleotide arrays can be used. The product of these analyses are typically measurements of the intensity of the signal received from a labeled probe used to detect a cDNA sequence from the sample that hybridizes to a nucleic acid sequence at a known location on the microarray. Typically, the intensity of the signal is proportional to the quantity of cDNA, and thus mRNA or miRNA, expressed in the sample cells. A large number of such techniques are available and useful. Methods for determining gene expression can be found in U.S. Pat. No. 6,271,002 to Linsley, et al.; U.S. Pat. No. 6,218,122 to Friend, et al.; U.S. Pat. No. 6,218,114 to Peck et al.; or U.S. Pat. No. 6,004,755 to Wang, et al., each of which is herein incorporated by reference in its entirety.

[00263] Analysis of an expression level can be conducted by comparing such intensities. This can be performed by generating a ratio matrix of the expression intensities of genes in a test sample versus those in a control sample. The control sample may be used as a reference, and different references to account for age, ethnicity and sex may be used. Different references can be used for different conditions or diseases, as well as different stages of diseases or conditions, as well as for determining therapeutic efficacy.

[00264] For instance, the gene expression intensities of mRNA or miRNAs derived from a diseased tissue, including those isolated from vesicles, can be compared with the expression intensities of the same entities in normal tissue of the same type (e.g., diseased breast tissue sample versus normal breast tissue sample). A ratio

of these expression intensities indicates the fold-change in gene expression between the test and control samples. Alternatively, if vesicles are not normally present in from normal tissues (e.g. breast) then absolute quantitation methods, as is known in the art, can be used to define the number of miRNA molecules present without the requirement of miRNA or mRNA isolated from vesicles derived from normal tissue.

[00265] Gene expression profiles can also be displayed in a number of ways. A common method is to arrange raw fluorescence intensities or ratio matrix into a graphical dendogram where columns indicate test samples and rows indicate genes. The data is arranged so genes that have similar expression profiles are proximal to each other. The expression ratio for each gene is visualized as a color. For example, a ratio less than one (indicating down-regulation) may appear in the blue portion of the spectrum while a ratio greater than one (indicating up-regulation) may appear as a color in the red portion of the spectrum. Commercially available computer software programs are available to display such data.

[00266] mRNAs or miRNAs that are considered differentially expressed can be either over expressed or under expressed in patients with a disease relative to disease free individuals. Over and under expression are relative terms meaning that a detectable difference (beyond the contribution of noise in the system used to measure it) is found in the amount of expression of the mRNAs or miRNAs relative to some baseline. In this case, the baseline is the measured mRNA/miRNA expression of a non-diseased individual. The mRNA/miRNA of interest in the diseased cells can then be either over or under expressed relative to the baseline level using the same measurement method. Diseased, in this context, refers to an alteration of the state of a body that interrupts or disturbs, or has the potential to disturb, proper performance of bodily functions as occurs with the uncontrolled proliferation of cells. Someone is diagnosed with a disease when some aspect of that person's genotype or phenotype is consistent with the presence of the disease. However, the act of conducting a diagnosis or prognosis includes the determination of disease/status issues such as determining the likelihood of relapse or metastasis and therapy monitoring. In therapy monitoring, clinical judgments are made regarding the effect of a given course of therapy by comparing the expression of genes over time to determine whether the mRNA/miRNA expression profiles have changed or are changing to patterns more consistent with normal tissue.

[00267] Levels of over and under expression are distinguished based on fold changes of the intensity measurements of hybridized microarray probes. A 2X difference is preferred for making such distinctions or a p-value less than 0.05. That is, before an mRNA/miRNA is the to be differentially expressed in diseased/relapsing versus normal/non-relapsing cells, the diseased cell is found to yield at least 2 times more, or 2 times less intensity than the normal cells. The greater the fold difference, the more preferred is use of the gene as a diagnostic or prognostic tool. mRNA/miRNAs selected for the expression profiles of the instant invention have expression levels that result in the generation of a signal that is distinguishable from those of the normal or non-modulated genes by an amount that exceeds background using clinical laboratory instrumentation.

[00268] Statistical values can be used to confidently distinguish modulated from non-modulated mRNA/miRNA and noise. Statistical tests find the mRNA/miRNA most significantly different between diverse groups of samples. The Student's t-test is an example of a robust statistical test that can be used to find significant differences between two groups. The lower the p-value, the more compelling the evidence that the gene shows a difference between the different groups. Nevertheless, since microarrays measure more than one mRNA/miRNA at a time, tens of thousands of statistical tests may be performed at one time. Because of this,

one is unlikely to see small p-values just by chance and adjustments for this using a Sidak correction as well as a randomization/permutation experiment can be made. A p-value less than 0.05 by the t-test is evidence that the gene is significantly different. More compelling evidence is a p-value less then 0.05 after the Sidak correction is factored in. For a large number of samples in each group, a p-value less than 0.05 after the randomization/permutation test is the most compelling evidence of a significant difference.

[00269] In one embodiment, a method of generating a posterior probability score to enable diagnostic, prognostic, therapy-related, or physiological state specific biosignature scores can be arrived at by obtaining mRNA or miRNA (biomarker) expression data from a statistically significant number of patients; applying linear discrimination analysis to the data to obtain selected biomarkers; and applying weighted expression levels to the selected biomarkers with discriminate function factor to obtain a prediction model that can be applied as a posterior probability score. Other analytical tools can also be used to answer the same question such as, logistic regression and neural network approaches.

[00270] For instance, the following can be used for linear discriminant analysis:

where,

$I(p_s i_d)$ = The log base 2 intensity of the probe set enclosed in parenthesis. d(cp) = The discriminant function for the disease positive class $d(C_N)$ = The discriminant function for the disease negative class

$P_{(CP)}$ = The posterior p-value for the disease positive class

$P_{(CN)}$ = The posterior p-value for the disease negative class

[00271] Numerous other well-known methods of pattern recognition are available. The following references provide some examples: *Weighted Voting: Golub et al. (1999); Support Vector Machines: Su et al. (2001); and Ramaswamy et al. (2001); K-nearest Neighbors: Ramaswamy (2001); and Correlation Coefficients: van 't Veer et al. (2002),* all of which are herein incorporated by reference in their entireties.

[00272] A biosignature portfolio, further described below, can be established such that the combination of biomarkers in the portfolio exhibit improved sensitivity and specificity relative to individual biomarkers or randomly selected combinations of biomarkers. In one embodiment, the sensitivity of the biosignature portfolio can be reflected in the fold differences, for example, exhibited by a transcript's expression in the diseased state relative to the normal state. Specificity can be reflected in statistical measurements of the correlation of the signaling of transcript expression with the condition of interest. For example, standard deviation can be a used as such a measurement. In considering a group of biomarkers for inclusion in a biosignature portfolio, a small standard deviation in expression measurements correlates with greater specificity. Other measurements of variation such as correlation coefficients can also be used in this capacity.

[00273] Another parameter that can be used to select mRNA/miRNA that generate a signal that is greater than that of the non-modulated mRNA/miRNA or noise is the use of a measurement of absolute signal difference. The signal generated by the modulated mRNA/miRNA expression is at least 20% different than those of the normal or non-modulated gene (on an absolute basis). It is even more preferred that such mRNA/miRNA produce expression patterns that are at least 30% different than those of normal or non-modulated mRNA/miRNA.

[00274] MiRNA can also be detected and measured by amplification from a biological sample and measured using methods described in U.S. Patent No. 7,250,496, U.S. Application Publication Nos. 20070292878,

20070042380 or 20050222399 and references cited therein, each of which is herein incorporated by reference in its entirety. The microRNA can be assessed as in U.S. Patent Application Serial No. 12/609,847, entitled "METHODS FOR ASSESSING RNA PATTERNS" and filed October 30, 2009, which application is incorporated by reference herein in its entirety.

[00275] Peptide nucleic acids (PNAs) which are a new class of synthetic nucleic acid analogs in which the phosphate–sugar polynucleotide backbone is replaced by a flexible pseudo-peptide polymer may be utilized in analysis of a biosignature. PNAs are capable of hybridizing with high affinity and specificity to complementary RNA and DNA sequences and are highly resistant to degradation by nucleases and proteinases. Peptide nucleic acids (PNAs) are an attractive new class of probes with applications in cytogenetics for the rapid in situ identification of human chromosomes and the detection of copy number variation (CNV). Multicolor peptide nucleic acid-fluorescence in situ hybridization (PNA-FISH) protocols have been described for the identification of several human CNV-related disorders and infectious diseases. PNAs can also be utilized as molecular diagnostic tools to non-invasively measure oncogene mRNAs with tumor targeted radionuclide-PNA-peptide chimeras. Methods of using PNAs are described further in *Pellestor F et al, Curr Pharm Des. 2008;14(24):2439-44, Tian X et al, Ann N Y Acad Sci. 2005 Nov;1059:106-44, Paulasova P and Pellestor F, Annales de Génétique, 47 (2004) 349–358, Stender H. Expert Rev Mol Diagn. 2003 Sep;3(5):649-55. Review, Vigneault et al., Nature Methods, 5(9), 777 – 779 (2008),* each reference is herein incorporated by reference in its entirety. These methods can be used to screen the genetic materials isolated from a vesicle. When applying these techniques to a cell-of-origin specific vesicle, they can be used to identify a given molecular signal that directly pertains to the cell of origin.

[00276] Mutational analysis may be carried out for mRNAs and DNA, including those that are identified from a vesicle. For mutational analysis of a target or biomarker that is of RNA origin, the RNA (mRNA, miRNA or other) can be reverse transcribed into cDNA and subsequently sequenced or assayed, such as for known SNPs (by Taqman SNP assays, for example) or single nucleotide mutations, as well as using sequencing to look for insertions or deletions to determine mutations present in the cell-of-origin. Multiplexed ligation dependent probe amplification (MLPA) could alternatively be used for the purpose of identifying CNV in small and specific areas of interest. For example, once the total RNA has been obtained from isolated colon cancer-specific vesicles, cDNA can be synthesized and primers specific for exons 2 and 3 of the KRAS gene can be used to amplify these two exons containing codons 12, 13 and 61 of the KRAS gene. The same primers used for PCR amplification can be used for Big Dye Terminator sequence analysis on the ABI 3730 to identify mutations in exons 2 and 3 of KRAS. Mutations in these codons are known to confer resistance to drugs such as Cetuximab and Panitumimab. Methods of conducting mutational analysis are described in *Maheswaran S et al, July 2, 2008 (10.1056/NEJMoa0800668) and Orita, M et al, PNAS 1989, (86): 2766-70,* each of which is herein incorporated by reference in its entirety.

[00277] Other methods of conducting mutational analysis include miRNA sequencing. Applications for identifying and profiling miRNAs can be done by cloning techniques and the use of capillary DNA sequencing or "next-generation" sequencing technologies. The new sequencing technologies currently available allow the identification of low-abundance miRNAs or those exhibiting modest expression differences between samples, which may not be detected by hybridization-based methods. Such new sequencing technologies include the massively parallel signature sequencing (MPSS) methodology described in *Nakano et al. 2006, Nucleic Acids*

Res. 2006;34:D731–D735. doi: 10.1093/nar/gkj077, the Roche/454 platform described in *Margulies et al. 2005, Nature. 2005;437:376–380* or the Illumina sequencing platform described in *Berezikov et al. Nat. Genet. 2006b;38:1375–1377*, each of which is incorporated by reference in its entirety.

[00278] Additional methods to determine a biosignature includes assaying a biomarker by allele-specific PCR, which includes specific primers to amplify and discriminate between two alleles of a gene simultaneously, single-strand conformation polymorphism (SSCP), which involves the electrophoretic separation of single-stranded nucleic acids based on subtle differences in sequence, and DNA and RNA aptamers. DNA and RNA aptamers are short oligonucleotide sequences that can be selected from random pools based on their ability to bind a particular molecule with high affinity. Methods of using aptamers are described in *Ulrich H et al, Comb Chem High Throughput Screen. 2006 Sep;9(8):619-32, Ferreira CS et al, Anal Bioanal Chem. 2008 Feb;390(4):1039-50, Ferreira CS et al, Tumour Biol. 2006;27(6):289-301*, each of which is herein incorporated by reference in its entirety.

[00279] Biomarkers can also be detected using fluorescence in situ hybridization (FISH). Methods of using FISH to detect and localize specific DNA sequences, localize specific mRNAs within tissue samples or identify chromosomal abnormalities are described in *Shaffer DR et al, Clin Cancer Res. 2007 Apr 1;13(7):2023-9, Cappuzo F et al, Journal of Thoracic Oncology, Volume 2, Number 5, May 2007, Moroni M et al, Lancet Oncol. 2005 May;6(5):279-86*, each of which is herein incorporated by reference in its entirety.

[00280] An illustrative schematic for analyzing a population of vesicles for their payload is presented in **FIG. 2E**. In an embodiment, the methods of the invention include characterizing a phenotype by capturing vesicles (**230**) and determining a level of microRNA species contained therein (**231**), thereby characterizing the phenotype (**232**).

[00281] A biosignature comprising a circulating biomarker or vesicle can comprise a binding agent thereto. The binding agent can be a DNA, RNA, aptamer, monoclonal antibody, polyclonal antibody, Fabs, Fab', single chain antibody, synthetic antibody, aptamer (DNA/RNA), peptoid, zDNA, peptide nucleic acid (PNA), locked nucleic acid (LNA), lectin, synthetic or naturally occurring chemical compounds (including but not limited to drugs and labeling reagents).

[00282] A binding agent can used to isolate or detect a vesicle by binding to a component of the vesicle, as described above. The binding agent can be used to detect a vesicle, such as for detecting a cell-of-origin specific vesicle. A binding agent or multiple binding agents can themselves form a binding agent profile that provides a biosignature for a vesicle. One or more binding agents can be selected from **Table 3**. For example, if a vesicle population is detected or isolated using two, three or four binding agents in a differential detection or isolation of a vesicle from a heterogeneous population of vesicles, the particular binding agent profile for the vesicle population provides a biosignature for the particular vesicle population.

[00283] As an illustrative example, a vesicle for characterizing a cancer can be detected with one or more binding agents including, but not limited to, PSA, PSMA, PCSA, PSCA, B7H3, EpCam, TMPRSS2, mAB 5D4, XPSM-A9, XPSM-A10, Galectin-3, E-selectin, Galectin-1, or E4 (IgG2a kappa), or any combination thereof.

[00284] The binding agent can also be for a general vesicle biomarker, such as a "housekeeping protein" or antigen. The biomarker can be CD9, CD63, or CD81. For example, the binding agent can be an antibody for CD9, CD63, or CD81. The binding agent can also be for other proteins, such as for tissue specific or cancer specific vesicles. The binding agent can be for PCSA, PSMA, EpCam, B7H3, or STEAP. The binding agent

can be for DR3, STEAP, epha2, TMEM211, unc93A, A33, CD24, NGAL, EpCam, MUC17, TROP2, or TETS. For example, the binding agent can be an antibody or aptamer for PCSA, PSMA, EpCam, B7H3, DR3, STEAP, epha2, TMEM211, unc93A, A33, CD24, NGAL, EpCam, MUC17, TROP2, or TETS.

[00285] Various proteins are not typically distributed evenly or uniformly on a vesicle shell. See, e.g., **FIG. 3**, which illustrates a schematic of protein expression patterns. Vesicle-specific proteins are typically more common, while cancer-specific proteins are less common. In some embodiments, capture of a vesicle is accomplished using a more common, less cancer-specific protein, such as one or more housekeeping proteins or antigen or general vesicle antigen (e.g., a tetraspanin), and one or more cancer-specific biomarkers and/or one or more cell-of-origin specific biomarkers is used in the detection phase. In another embodiment, one or more cancer-specific biomarkers and/or one or more cell-of-origin specific biomarkers are used for capture, and one or more housekeeping proteins or antigen or general vesicle antigen (e.g., a tetraspanin) is used for detection. In embodiments, the same biomarker is used for both capture and detection. Different binding agents for the same biomarker can be used, such as antibodies or aptamers that bind different epitopes of an antigen.

[00286] Additional cellular binding partners or binding agents may be identified by any conventional methods known in the art, or as described herein, and may additionally be used as a diagnostic, prognostic or therapy-related marker.

Phenotypes

[00287] Analysis of microRNA, vesicles, and/or other biomarkers in a sample from a subject can be used to characterize a phenotype. A phenotype can be any observable characteristic or trait of a subject, such as a disease or condition, a disease stage or condition stage, susceptibility to a disease or condition, prognosis of a disease stage or condition, a physiological state, or response to therapeutics. A phenotype can result from a subject's gene expression as well as the influence of environmental factors and the interactions between the two, as well as from epigenetic modifications to nucleic acid sequences.

[00288] A phenotype in a subject can be characterized by obtaining a biological sample from said subject and analyzing one or more microRNA, vesicles, and/or other biomarkers in the sample. For example, characterizing a phenotype for a subject or individual may include detecting a disease or condition (including pre-symptomatic early stage detecting), determining the prognosis, diagnosis, or theranosis of a disease or condition, or determining the stage or progression of a disease or condition. Characterizing a phenotype can also include identifying appropriate treatments or treatment efficacy for specific diseases, conditions, disease stages and condition stages, predictions and likelihood analysis of disease progression, particularly disease recurrence, metastatic spread or disease relapse. A phenotype can also be a clinically distinct type or subtype of a condition or disease, such as a cancer or tumor. Phenotype determination can also be a determination of a physiological condition, or an assessment of organ distress or organ rejection, such as post-transplantation.

[00289] For example, the phenotype can comprise a tumor, neoplasm, or cancer. A cancer detected or assessed by products or processes described herein includes, but is not limited to, breast cancer, ovarian cancer, lung cancer, colon cancer, hyperplastic polyp, adenoma, colorectal cancer, high grade dysplasia, low grade dysplasia, prostatic hyperplasia, prostate cancer, melanoma, pancreatic cancer, brain cancer (such as a glioblastoma), hematological malignancy, hepatocellular carcinoma, cervical cancer, endometrial cancer, head and neck cancer, esophageal cancer, gastrointestinal stromal tumor (GIST), renal cell carcinoma (RCC) or gastric cancer. The colorectal cancer can be CRC Dukes B or Dukes C-D. The hematological malignancy can be B-Cell

Chronic Lymphocytic Leukemia, B-Cell Lymphoma-DLBCL, B-Cell Lymphoma-DLBCL-germinal center-like, B-Cell Lymphoma-DLBCL-activated B-cell-like, and Burkitt's lymphoma. The phenotype may also be a premalignant condition, such as Barrett's Esophagus or colorectal polyps.

[00290] The cancer characterized by the methods of the invention can comprise, without limitation, a carcinoma, a sarcoma, a lymphoma or leukemia, a germ cell tumor, a blastoma, or other cancers. Carcinomas include without limitation epithelial neoplasms, squamous cell neoplasms squamous cell carcinoma, basal cell neoplasms basal cell carcinoma, transitional cell papillomas and carcinomas, adenomas and adenocarcinomas (glands), adenoma, adenocarcinoma, linitis plastica insulinoma, glucagonoma, gastrinoma, vipoma, cholangiocarcinoma, hepatocellular carcinoma, adenoid cystic carcinoma, carcinoid tumor of appendix, prolactinoma, oncocytoma, hurthle cell adenoma, renal cell carcinoma, grawitz tumor, multiple endocrine adenomas, endometrioid adenoma, adnexal and skin appendage neoplasms, mucoepidermoid neoplasms, cystic, mucinous and serous neoplasms, cystadenoma, pseudomyxoma peritonei, ductal, lobular and medullary neoplasms, acinar cell neoplasms, complex epithelial neoplasms, warthin's tumor, thymoma, specialized gonadal neoplasms, sex cord stromal tumor, thecoma, granulosa cell tumor, arrhenoblastoma, sertoli leydig cell tumor, glomus tumors, paraganglioma, pheochromocytoma, glomus tumor, nevi and melanomas, melanocytic nevus, malignant melanoma, melanoma, nodular melanoma, dysplastic nevus, lentigo maligna melanoma, superficial spreading melanoma, and malignant acral lentiginous melanoma. Sarcoma includes without limitation Askin's tumor, botryodies, chondrosarcoma, Ewing's sarcoma, malignant hemangio endothelioma, malignant schwannoma, osteosarcoma, soft tissue sarcomas including: alveolar soft part sarcoma, angiosarcoma, cystosarcoma phyllodes, dermatofibrosarcoma, desmoid tumor, desmoplastic small round cell tumor, epithelioid sarcoma, extraskeletal chondrosarcoma, extraskeletal osteosarcoma, fibrosarcoma, hemangiopericytoma, hemangiosarcoma, kaposi's sarcoma, leiomyosarcoma, liposarcoma, lymphangiosarcoma, lymphosarcoma, malignant fibrous histiocytoma, neurofibrosarcoma, rhabdomyosarcoma, and synovialsarcoma. Lymphoma and leukemia include without limitation chronic lymphocytic leukemia/small lymphocytic lymphoma, B-cell prolymphocytic leukemia, lymphoplasmacytic lymphoma (such as waldenström macroglobulinemia), splenic marginal zone lymphoma, plasma cell myeloma, plasmacytoma, monoclonal immunoglobulin deposition diseases, heavy chain diseases, extranodal marginal zone B cell lymphoma, also called malt lymphoma, nodal marginal zone B cell lymphoma (nmzl), follicular lymphoma, mantle cell lymphoma, diffuse large B cell lymphoma, mediastinal (thymic) large B cell lymphoma, intravascular large B cell lymphoma, primary effusion lymphoma, burkitt lymphoma/leukemia, T cell prolymphocytic leukemia, T cell large granular lymphocytic leukemia, aggressive NK cell leukemia, adult T cell leukemia/lymphoma, extranodal NK/T cell lymphoma, nasal type, enteropathy-type T cell lymphoma, hepatosplenic T cell lymphoma, blastic NK cell lymphoma, mycosis fungoides / sezary syndrome, primary cutaneous CD30-positive T cell lymphoproliferative disorders, primary cutaneous anaplastic large cell lymphoma, lymphomatoid papulosis, angioimmunoblastic T cell lymphoma, peripheral T cell lymphoma, unspecified, anaplastic large cell lymphoma, classical hodgkin lymphomas (nodular sclerosis, mixed cellularity, lymphocyte-rich, lymphocyte depleted or not depleted), and nodular lymphocyte-predominant hodgkin lymphoma. Germ cell tumors include without limitation germinoma, dysgerminoma, seminoma, nongerminomatous germ cell tumor, embryonal carcinoma, endodermal sinus turmor, choriocarcinoma, teratoma, polyembryoma, and gonadoblastoma. Blastoma includes without limitation nephroblastoma, medulloblastoma, and retinoblastoma. Other cancers include without limitation labial

carcinoma, larynx carcinoma, hypopharynx carcinoma, tongue carcinoma, salivary gland carcinoma, gastric carcinoma, adenocarcinoma, thyroid cancer (medullary and papillary thyroid carcinoma), renal carcinoma, kidney parenchyma carcinoma, cervix carcinoma, uterine corpus carcinoma, endometrium carcinoma, chorion carcinoma, testis carcinoma, urinary carcinoma, melanoma, brain tumors such as glioblastoma, astrocytoma, meningioma, medulloblastoma and peripheral neuroectodermal tumors, gall bladder carcinoma, bronchial carcinoma, multiple myeloma, basalioma, teratoma, retinoblastoma, choroidea melanoma, seminoma, rhabdomyosarcoma, craniopharyngeoma, osteosarcoma, chondrosarcoma, myosarcoma, liposarcoma, fibrosarcoma, Ewing sarcoma, and plasmocytoma.

[00291] In a further embodiment, the cancer under analysis may be a lung cancer including non-small cell lung cancer and small cell lung cancer (including small cell carcinoma (oat cell cancer), mixed small cell/large cell carcinoma, and combined small cell carcinoma), colon cancer, breast cancer, prostate cancer, liver cancer, pancreas cancer, brain cancer, kidney cancer, ovarian cancer, stomach cancer, skin cancer, bone cancer, gastric cancer, breast cancer, pancreatic cancer, glioma, glioblastoma, hepatocellular carcinoma, papillary renal carcinoma, head and neck squamous cell carcinoma, leukemia, lymphoma, myeloma, or a solid tumor.

[00292] In embodiments, the cancer comprises an acute lymphoblastic leukemia; acute myeloid leukemia; adrenocortical carcinoma; AIDS-related cancers; AIDS-related lymphoma; anal cancer; appendix cancer; astrocytomas; atypical teratoid/rhabdoid tumor; basal cell carcinoma; bladder cancer; brain stem glioma; brain tumor (including brain stem glioma, central nervous system atypical teratoid/rhabdoid tumor, central nervous system embryonal tumors, astrocytomas, craniopharyngioma, ependymoblastoma, ependymoma, medulloblastoma, medulloepithelioma, pineal parenchymal tumors of intermediate differentiation, supratentorial primitive neuroectodermal tumors and pineoblastoma); breast cancer; bronchial tumors; Burkitt lymphoma; cancer of unknown primary site; carcinoid tumor; carcinoma of unknown primary site; central nervous system atypical teratoid/rhabdoid tumor; central nervous system embryonal tumors; cervical cancer; childhood cancers; chordoma; chronic lymphocytic leukemia; chronic myelogenous leukemia; chronic myeloproliferative disorders; colon cancer; colorectal cancer; craniopharyngioma; cutaneous T-cell lymphoma; endocrine pancreas islet cell tumors; endometrial cancer; ependymoblastoma; ependymoma; esophageal cancer; esthesioneuroblastoma; Ewing sarcoma; extracranial germ cell tumor; extragonadal germ cell tumor; extrahepatic bile duct cancer; gallbladder cancer; gastric (stomach) cancer; gastrointestinal carcinoid tumor; gastrointestinal stromal cell tumor; gastrointestinal stromal tumor (GIST); gestational trophoblastic tumor; glioma; hairy cell leukemia; head and neck cancer; heart cancer; Hodgkin lymphoma; hypopharyngeal cancer; intraocular melanoma; islet cell tumors; Kaposi sarcoma; kidney cancer; Langerhans cell histiocytosis; laryngeal cancer; lip cancer; liver cancer; malignant fibrous histiocytoma bone cancer; medulloblastoma; medulloepithelioma; melanoma; Merkel cell carcinoma; Merkel cell skin carcinoma; mesothelioma; metastatic squamous neck cancer with occult primary; mouth cancer; multiple endocrine neoplasia syndromes; multiple myeloma; multiple myeloma/plasma cell neoplasm; mycosis fungoides; myelodysplastic syndromes; myeloproliferative neoplasms; nasal cavity cancer; nasopharyngeal cancer; neuroblastoma; Non-Hodgkin lymphoma; nonmelanoma skin cancer; non-small cell lung cancer; oral cancer; oral cavity cancer; oropharyngeal cancer; osteosarcoma; other brain and spinal cord tumors; ovarian cancer; ovarian epithelial cancer; ovarian germ cell tumor; ovarian low malignant potential tumor; pancreatic cancer; papillomatosis; paranasal sinus cancer; parathyroid cancer; pelvic cancer; penile cancer; pharyngeal cancer; pineal parenchymal tumors of intermediate differentiation; pineoblastoma; pituitary

tumor; plasma cell neoplasm/multiple myeloma; pleuropulmonary blastoma; primary central nervous system (CNS) lymphoma; primary hepatocellular liver cancer; prostate cancer; rectal cancer; renal cancer; renal cell (kidney) cancer; renal cell cancer; respiratory tract cancer; retinoblastoma; rhabdomyosarcoma; salivary gland cancer; Sézary syndrome; small cell lung cancer; small intestine cancer; soft tissue sarcoma; squamous cell carcinoma; squamous neck cancer; stomach (gastric) cancer; supratentorial primitive neuroectodermal tumors; T-cell lymphoma; testicular cancer; throat cancer; thymic carcinoma; thymoma; thyroid cancer; transitional cell cancer; transitional cell cancer of the renal pelvis and ureter; trophoblastic tumor; ureter cancer; urethral cancer; uterine cancer; uterine sarcoma; vaginal cancer; vulvar cancer; Waldenström macroglobulinemia; or Wilm's tumor. The methods of the invention can be used to characterize these and other cancers. Thus, characterizing a phenotype can be providing a diagnosis, prognosis or theranosis of one of the cancers disclosed herein.

[00293] The phenotype can also be an inflammatory disease, immune disease, or autoimmune disease. For example, the disease may be inflammatory bowel disease (IBD), Crohn's disease (CD), ulcerative colitis (UC), pelvic inflammation, vasculitis, psoriasis, diabetes, autoimmune hepatitis, Multiple Sclerosis, Myasthenia Gravis, Type I diabetes, Rheumatoid Arthritis, Psoriasis, Systemic Lupus Erythematosis (SLE), Hashimoto's Thyroiditis, Grave's disease, Ankylosing Spondylitis Sjogrens Disease, CREST syndrome, Scleroderma, Rheumatic Disease, organ rejection, Primary Sclerosing Cholangitis, or sepsis.

Cancer of the GI Tract

[00294] In some embodiments, the methods of the invention are used to characterize a phenotype, wherein the phenotype comprises a disease or disorder of the gastrointestinal (GI) tract. The gastrointestinal tract includes without limitation the oral cavity, gums, pharynx, tongue, salivary glands, esophagus, pancreas, liver, gallbladder, small intestine (duodenum, jejunum, ileum), bile duct, stomach, large intestine (cecum, colon, rectum), appendix and anus. A biosignature can be used to detect or characterize cancers of such components, e.g., colorectal cancer (CRC), stomach cancer, intestinal cancer, liver cancer or esophageal cancer.

[00295] A colorectal cancer (CRC) biosignature can comprise any one or more antigens for colon cancer as listed in **Table 2**, any one or more binding agents associated with isolating or detecting a vesicle for characterizing colon cancer (for example, as shown in **Table 3**). The biosignature can comprise one or more miRNA selected from the group consisting of miR-24-1, miR-29b-2, miR-20a, miR-10a, miR-32, miR-203, miR-106a, miR-17-5p, miR-30c, miR-223, miR-126, miR-128b, miR-21, miR-24-2, miR-99b, miR-155, miR-213, miR-150, miR-107, miR-191, miR-221, miR-20a, miR-510, miR-92, miR-513, miR-19a, miR-21, miR-20, miR-183, miR-96, miR-135b, miR-31, miR-21, miR-92, miR-222, miR-181b, miR-210, miR-20a, miR-106a, miR-93, miR-335, miR-338, miR-133b, miR-346, miR-106b, miR-153a, miR-219, miR-34a, miR-99b, miR-185, miR-223, miR-211, miR-135a, miR-127, miR-203, miR-212, miR-95, or miR-17-5p, or any combination thereof. The biosignature can also comprise one or more underexpressed miRs such as miR-143, miR-145, miR-143, miR-126, miR-34b, miR-34c, let-7, miR-9-3, miR-34a, miR-145, miR-455, miR-484, miR-101, miR-145, miR-133b, miR-129, miR-124a, miR-30-3p, miR-328, miR-106a, miR-17-5p, miR-342, miR-192, miR-1, miR-34b, miR-215, miR-192, miR-301, miR-324-5p, miR-30a-3p, miR-34c, miR-331, miR-148b, miR-548c-5p, miR-362-3p and miR422a.

[00296] The biosignature can comprise assessing one or more genes, such as EFNB1, ERCC1, HER2, VEGF, and EGFR. A biomarker mutation for colorectal cancer that can be assessed in a vesicle can also include one or more mutations of EGFR, KRAS, VEGFA, B-Raf, APC, or p53. The biosignature can also comprise one or

more proteins, ligands, or peptides that can be assessed of a vesicle, such as AFRs, Rabs, ADAM10, CD44, NG2, ephrin-B1, MIF, b-catenin, Junction, plakoglobin, galectin-4, RACK1, tetraspanin-8, FasL, TRAIL, A33, CEA, EGFR, dipeptidase 1, hsc-70, tetraspanins, ESCRT, TS, PTEN, or TOPO1.

[00297] Gastrointestinal disorders can be detected using various biomarkers indicative of the origin of interest. The biomarkers can be detected in various forms, including without limitation as circulating biomarkers, embedded in a cellular or vesicle membrane, associated with a membrane fragment, or as vesicle payload. Biomarkers that can be included in a biosignature for colorectal cancer comprise without limitation one or more of CEA, MUC2, GPA33, CEACAM5, ENFB1, CCSA-3, CCSA-4, ADAM10, CD44, NG2, ephrin B1, plakoglobin, galectin 4, RACK1, tetraspanin-8, FASL, A33, CEA, EGFR, dipeptidase 1, PTEN, Na(+)-dependent glucose transporter, UDP-glucuronosyltransferase 1A, fragments thereof, or other antigens useful for identifying a colon cancer cell. Biomarkers that can be included in a biosignature for liver cancer comprise without limitation one or more of HBxAg, HBsAg, NLT, fragments thereof, or other antigens useful for identifying a hepatocellular carcinoma cell. Biomarkers that can be included in a biosignature for irritable bowel disease (IBD) or syndrome (IBS) comprise without limitation one or more of IL-16, IL-1beta, IL-12, TNF-alpha, interferon-gamma, IL-6, Rantes, II-12, MCP-1, 5HT, fragments thereof, or other antigens useful for identifying IBD or IBS. Biomarkers that can be included in a biosignature for esophageal cancer comprise without limitation CaSR, fragments thereof, or other antigens useful for identifying an esophageal cell. Biomarkers that can be included in a biosignature for Barrett's Esophagus comprise without limitation one or more of p53, MUC1, MUC6, fragments thereof, or other antigens useful for identifying an esophageal cell. Biomarkers that can be included in a biosignature for gastrointestinal stromal tumor (GIST) comprise without limitation one or more of c-kit, PDGFRA, NHE-3, fragments thereof, or other antigens useful for identifying a GIST cell. Biomarkers that can be included in a biosignature for cirrhosis comprise without limitation one or more of NLT, HBsAg, fragments thereof, or other antigens useful for identifying cirrhosis.

[00298] A microRNA, biomarker and/or vesicle can be isolated and assayed to provide a diagnostic, prognostic or theranostic profile, such as the stage of the cancer, the efficacy of the cancer, or other characteristics of the cancer. The microRNA and/or biomarkers can be circulating in a bodily fluid, can be membrane associated, or can be associated with a vesicle, such as vesicle payload or surface markers. The vesicles can be isolated as described herein prior to analysis. Vesicles can also be directly assayed from a sample, such that the vesicles are not purified or concentrated prior to assaying for a biosignature associated with colorectal cancer.

[00299] The one or more miRNAs used to characterize a phenotype may be selected from those disclosed in PCT Publication No. WO2009/036236. For example, one or more miRNAs listed in Tables I-VI (Figures 6-11) therein can be used to characterize colon adenocarcinoma, colorectal cancer, and other diseases and disorders, as further described therein. In an embodiment, microRNAs useful for characterizing colorectal cancer include one or more of the following, which can be upregulated in CRC compared to normal: miR-19a, miR-21, miR-127, miR-31, miR-96, miR- 135b and miR-183. MicroRNAs useful for characterizing colorectal cancer also include one or more of the following, which can be downregulated in CRC compared to normal: miR-30c, miR-133a, mirl43, miR-133b and miR-145. In some embodiments, microRNAs useful for characterizing colon adenocarcinomas include one or more of the following, which can be upregulated: miR-20a, miR-21, miR-106a, miR-181b and miR-203. MicroRNAs that can be used to characterize a pancreatic cancer include one or more of the following upregulated miRs: miR-103, miR-107, miR-18a, miR-31, miR-93, miR- 221, miR-224 and

miR-155; and/or one or more of the following downregulated miRs: miR- 133a, miR-216 miR-217. The one or more biomarker can be used to characterize a colorectal cancer can be an upregulated or overexpressed miRNA selected from the group consisting of: miR-548c-5p, miR-362-3p, miR-422a, miR-597, miR-429, miR-200a, and miR-200b, or any combination thereof. The microRNAs can be detected directly in a bodily fluid or in vesicles isolated from such sample.

[00300] Other biomarkers that can be used in a biosignature for characterizing GI disorders comprise without limitation overexpressed miRs, underexpressed miRs, mRNAs, genetic mutations, proteins, ligands, peptides, and snoRNA. The markers can be detected in bodily fluids and/or in association with a vesicle. Biomarkers that can be used in a biosignature to differentiate adenoma and hyperplastic polyps include without limitation ABCA8, KIAA1199, GCG, MAMDC2, C2orf32, 229670_at, IGF1, PCDH7, PRDX6, PCNA, COX2, MUC6, hTERT, or a combination thereof. Mutation of KRAS and/or B-Raf can also be used for distinguishing between adenoma and hyperplastic polyp.

[00301] Biomarkers that can be used in a biosignature to characterize IBD/IBS include without limitation REG1A, MMP3, or a combination thereof.

[00302] Biomarkers that can be used in a biosignature to differentiate adenoma and colorectal cancer include without limitation GREM1, DDR2, GUCY1A3, TNS1, ADAMTS1, FBLN1, FLJ38028, RDX, FAM129A, ASPN, FRMD6, MCC, RBMS1, SNAI2, MEIS1, DOCK10, PLEKHC1, FAM126A, TBC1D9, VWF, DCN, ROBO1, MSRB3, LATS2, MEF2C, IGFBP3, GNB4, RCN3, AKAP12, RFTN1, 226834_at, COL5A1, GNG2, NR3C1*, SPARCL1, MAB21L2, AXIN2, 236894_at, AEBP1, AP1S2, C10orf56, LPHN2, AKT3, FRMD6, COL15A1, CRYAB, COL14A1, LOC286167, QKI, WWTR1, GNG11, PAPPA, or ELDT1, or a combination thereof.

[00303] Biomarkers that can be used in a biosignature to differentiate IBD and colorectal cancer include without limitation 227458_at, INDO, CXCL9, CCR2, CD38, RARRES3, CXCL10, FAM26F, TNIP3, NOS2A, CCRL1, TLR8, IL18BP, FCRL5, SAMD9L, ECGF1, TNFSF13B, GBP5, or GBP1, or a combination thereof. Biomarkers that can be used in a biosignature to differentiate Dukes B and Dukes C-D colorectal cancer include without limitation TMEM37*, IL33, CA4, CCDC58, CLIC6, VERSUSNL1, ESPN, APCDD1, C13orf18, CYP4X1, ATP2A3, LOC646627, MUPCDH, ANPEP, C1orf115, HSD3B2, GBA3, GABRB2, GYLTL1B, LYZ, SPC25, CDKN2B, FAM89A, MOGAT2, SEMA6D, 229376_at, TSPAN5, IL6R, or SLC26A2, or a combination thereof.

[00304] Biomarkers that can be used in a biosignature to differentiate adenoma with low grade dysplasia versus adenoma with high grade dysplasia include without limitation SI, DMBT1, CFI*, AQP1, APOD, TNFRSF17, CXCL10, CTSE, IGHA1, SLC9A3, SLC7A1, BATF2, SOCS1, DOCK2, NOS2A, HK2, CXCL2, IL15RA, POU2AF1, CLEC3B, ANI3BP, MGC13057, LCK*, C4BPA, HOXC6, GOLT1A, C2orf32, IL10RA,, 240856_at, SOCS3,, MEIS3P1, HIPK1, GLS, CPLX1, 236045_x_at, GALC, AMN, CCDC69, CCL28, CPA3, TRIB2, HMGA2, PLCL2, NR3C1, EIF5A, LARP4, RP5-1022P6.2, PHLDB2, FKBP1B, INDO, CLDN8, CNTN3, PBEF1, SLC16A9, CDC25B, TPSB2, PBEF1, ID4, GJB5, CHN2, LIMCH1, or CXCL9, or a combination thereof.

[00305] Biomarkers that can be used in a biosignature to differentiate ulcerative colitis (UC) and Crohn's disease (CD) include without limitation IFITM1, IFITM3, STAT1, STAT3, TAP1, PSME2, PSMB8, HNF4G, KLF5, AQP8, APT2B1, SLC16A, MFAP4, CCNG2, SLC44A4, DDAH1, TOB1, 231152_at, MKNK1,

CEACAM7*, 1562836_at, CDC42SE2, PSD3, 231169_at, IGL@*, GSN, GPM6B, CDV3*, PDPK1, ANP32E, ADAM9, CDH1, NLRP2, 215777_at, OSBPL1, VNN1, RABGAP1L, PHACTR2, ASH1L, 213710_s_at, CDH1, NLRP2, 215777_at, OSBPL1, VNN1, RABGAP1L, PHACTR2, ASH1, 213710_s_at, ZNF3, FUT2, IGHA1, EDEM1, GPR171, 229713_at, LOC643187, FLVCR1, SNAP23*, ETNK1, LOC728411, POSTN, MUC12, HOXA5, SIGLEC1, LARP5, PIGR, SPTBN1, UFM1, C6orf62, WDR90, ALDH1A3, F2RL1, IGHV1-69, DUOX2, RAB5A, CP, (P)ASCA or a combination thereof. Mutation of CARD15 can be used for distinguishing UC versus CD.

[00306] Biomarkers that can be used in a biosignature to characterize hyperplastic polyp or distinguish the polyps from normal include without limitation SLC6A14, ARHGEF10, ALS2, IL1RN, SPRY4, PTGER3, TRIM29, SERPINB5, 1560327_at, ZAK, BAG4, TRIB3, TTL, FOXQ1, or a combination thereof.

[00307] Biomarkers that can be used in a biosignature to differentiate adenoma with low grade dysplasia versus normal include without limitation UGT2A3, KLK11, KIAA1199, FOXQ1, CLDN8, ABCA8, PYY, or a combination thereof. snoRNA that can be used as a biomarker for adenoma low grade dysplasia versus normal include, but is not limited to, GAS5.

[00308] Biomarkers that can be used in a biosignature to differentiate adenoma versus normal include without limitation KIAA1199, FOXQ1, CA7, Clusterin, or a combination thereof.

[00309] Biomarkers that can be used in a biosignature to differentiate colorectal cancer versus normal include without limitation VWF, IL8, CHI3L1, S100A8, GREM1, ODC, or a combination thereof. A mutation of KRAS, BRAF, APC, MSH2, or MLH1 can further be used to characterize colorectal cancer. Other markers for colorectal cancer include the peptides cytokeratin 13, calcineurin, CHK1, clathrin light chain, phospho-ERK, phospho-PTK2, or MDM2.

[00310] Biomarkers that can be used in a biosignature to characterize pancreatic cancer include without limitation miR-221, miR-181a, miR-155, miR-210, miR-213, miR-181b, miR-222, miR-181b-2, miR-21, miR-181b-1, miR-220, miR-181d, miR-223, miR-100-1/2, miR-125a, miR-143, miR-10a, miR-146, miR-99, miR-100, miR-199a-1, miR-10b, miR-199a-2, miR-221, miR-181a, miR-155, miR-210, miR-213, miR-181b, miR-222, miR-181b-2, miR-21, miR-181b-1, miR-181c, miR-220, miR-181d, miR-223, miR-100-1/2, miR-125a, miR-143, miR-10a, miR-146, miR-99, miR-100, miR-199a-1, miR-10b, miR-199a-2, miR-107, miR-103, miR-103-2, miR-125b-1, miR-205, miR-23a, miR-221, miR-424, miR-301, miR-100, miR-376a, miR-125b-1, miR-21, miR-16-1, miR-181a, miR-181c, miR-92, miR-15, miR-155, let-7f-1, miR-212, miR-107, miR-024-1/2, miR-18a, miR-31, miR-93, miR-224, or let-7d, or a combination thereof. In embodiments, overexpression of these miRs as compared to a reference is indicative of pancreatic cancer. One or more of the following miRs can be used in a biosignature for pancreatic cancer: miR-103-2, miR-107, miR-103-1, miR-342, miR-100, miR-24-2, miR-23a, miR-125a, miR-26a-l, miR-24-1, miR-191, miR-15a, miR-368, miR-26b, miR-125b-2, miR-125b- 1, miR-26a-2, miR-335, miR-126. miR-1-2, miR-21, miR-25, miR-92-2, miR-130a, miR-93, miR-16-1, miR-145, miR-17, miR-99b, miR-181b-l, miR-146, miR-181b-2, miR- 16-2, miR-99a, miR- 197, miR- 10a, miR-224, miR-92-1, miR-27a, miR-221, miR- 320, miR-7-1 , miR-29b-2, miR-150, miR-30d, miR-29a, miR-23b, miR-135a-2, miR- 223, miR-3p21-v, miR-128b, miR-30b, miR-29b-l, miR-106b, miR-132, miR-214, miR-7-3, miR-29c, miR-367, miR-30c-2, miR-27b, miR-140, miR-10b, miR-20, miR- 129-1, miR-340, miR-30a, miR-30c-l, miR-106a, miR-32, miR-95, miR-222, miR-30e, miR- 129-2, miR-345, miR- 143, miR- 182, miR-1-1, miR-133a-l, miR-200c, miR- 194- 1, miR-210, miR-181c, miR-192, miR-220, miR-213, miR-323, and miR-

375, wherein high expression or overexpression of the one or more miRs can be used to characterize pancreatic cancer. Additional biomarkers that can be used in a biosignature to characterize pancreatic cancer include without limitation miR-148a, miR-148b, miR-375, miR-345, miR-142, miR-133a, miR-216, miR-217, miR-139, or a combination thereof. In embodiments, underexpression of these miRs as compared to a reference is indicative of pancreatic cancer. miR-375 expression can be used to characterize pancreatic insular or acinar tumors. Still other biomarkers that can be used in a biosignature to characterize pancreatic cancer include without limitation CA 19-9, prostate stem cell antigen (PSCA), Mesothelin, Osteopontin, or a combination thereof. Mutation of KRAS, CTNNLB1, AKT, NCOA3, or B-RAF can further be used to characterize pancreatic cancer. Other biomarkers for pancreatic cancer include BRCA2, PALB2, or p16.

[00311] Other biomarkers that can be used in a biosignature to characterize hepatocellular carcinoma include without limitation miR-221. In embodiments, overexpression of this miR as compared to a reference is indicative of hepatocellular carcinoma. Other biomarkers that can be used in a biosignature to characterize hepatocellular carcinoma include without limitation let-7a-1, let-7a-2, let-7a-3, let-7b, let-7c, let-7d, let-7e, let-7f-2, let-fg, miR-122a, miR-124a-2, miR-130a, miR-132, miR-136, miR-141, miR-142, miR-143, miR-145, miR-146, miR-150, miR-155(BIC), miR-181a-1, miR-181a-2, miR-181c, miR-195, miR-199a-1-5p, miR-199a-2-5p, miR-199b, miR-200b, miR-214, miR-223, or pre-miR-594, or a combination thereof. In embodiments, underexpression of these miRs as compared to a reference is indicative of hepatocellular carcinoma. Other biomarkers for hepatocellular carcinoma include FAT10. Biomarkers that can be used in a biosignature to characterize hepatitis C virus-associated hepatocellular carcinoma include without limitation miR-122, miR-100, miR-10a, or a combination thereof. In embodiments, overexpression of these miRs as compared to a reference is indicative of hepatocellular carcinoma. Other biomarkers that can be used in a biosignature to characterize hepatitis C virus-associated hepatocellular carcinoma include without limitation miR-198, miR-145, or a combination thereof. In embodiments, underexpression of these miRs as compared to a reference is indicative of hepatocellular carcinoma.

[00312] Biomarkers that can be used in a biosignature to characterize inflammatory bowel disease (IBD) include without limitation Trypsinogen IV, SERT, or a combination thereof. Mutation of CARD15 can further be used to characterize IBD. Other biomarkers for IBD include Il-16, Il-1beta, Il-12, TNF-alpha, interferon gamma, Il-6, Rantes, MCP-1, Resistin, or 5-HT.

[00313] Biomarkers that can be used in a biosignature to characterize Barrett's Esophagus include without limitation miR-21, miR-143, miR-145, miR-194, miR-215, or a combination thereof. In embodiments, overexpression of these miRs as compared to a reference is indicative of Barrett's Esophagus. Additional biomarkers that can be used in a biosignature to characterize Barrett's Esophagus include without limitation S100A2, S100A4, or a combination thereof. Mutation of p53 can further be used to characterize Barrett's Esophagus. Other markers for Barrett's Esophagus include p53, MUC1, MUC2.

[00314] Biomarkers that can be used in a biosignature to characterize gastrointestinal stromal tumor (GIST) include without limitation DOG-1, PKC-theta, KIT, GPR20, PRKCQ, KCNK3, KCNH2, SCG2, TNFRSF6B, CD34, or a combination thereof. Mutation of PKC-theta can further be used to characterize GIST. Other markers for GIST include PDGFRA, c-kit, or a combination thereof.

[00315] Biomarkers that can be used in a biosignature to characterize cirrhosis include without limitation NLT, HBsAG, AST, YKL-40, Hyaluronic acid, TIMP-1, alpha 2 macroglobulin, a-1-antitrypsin PlZ allele, haptoglobin, acid phosphatase ACP AC, or a combination thereof.

[00316] Biomarkers that can be used in a biosignature to characterize esophageal cancer include without limitation miR-192, miR-194, miR-21, miR-200c, miR-93, miR-342, miR-152, miR-93, miR-25, miR-424, miR-151, or a combination thereof. In embodiments, overexpression of these miRs as compared to a reference is indicative of esophageal cancer. Additional biomarkers that can be used in a biosignature to characterize esophageal cancer include without limitation miR-27b, miR-205, miR-203, miR-342, let-7c, miR-125b, miR-100, miR-152, miR-192, miR-194, miR-27b, miR-205, miR-203, miR-200c, miR-99a, miR-29c, miR-140, miR-103, or miR-107, or a combination thereof. In embodiments, underexpression of these miRs as compared to a reference is indicative of esophageal cancer. Other markers useful for characterizing esophageal cancer include the MTHFR.

[00317] Biomarkers that can be used in a biosignature to characterize gastric cancer include without limitation miR-106a, miR-21, miR-191, miR-223, miR-24-1, miR-24-2, miR-107, miR-92-2, miR-214, miR-25, miR-221, or a combination thereof. In embodiments, overexpression of these miRs as compared to a reference is indicative of gastric cancer. Additional biomarkers that can be used in a biosignature to characterize gastric cancer include without limitation let-7a. In embodiments, underexpression of this miR as compared to a reference is indicative of gastric cancer. Mutation of gastric cancer can further be used to characterize APC. Other useful biomarkers for characterizing gastric cancer include CEA (Carcinoembryonic Antigen), CA 19-9, CA 72-4, CA 125, RRM2, EphA4, and/or survivin.

[00318] The microRNAs miR-194, miR-148 and miR-192 are highly enriched in liver. A biosignature comprising these miRs can be used to distinguish positive and negative lymph nodes from a subject with GI cancer. Additional microRNAs useful in the methods of the invention include those described in U.S. Pat. No. 7,592,441, which describes microRNAs related to liver cancer. U.S. Pat. No. 7,592,441 is herein incorporated by reference in its entirety.

[00319] Cell proliferation has been correlated with the expression of miR-31, miR-92, miR-99a, miR-100, miR-125a, miR-129, miR-130a, miR-150, miR-187, miR-190, miR-191, miR-193, miR 204, miR-210, miR-21 1, miR-212, miR-213, miR-215, miR-216, miR-217, miR 218, miR-224, miR-292, miR-294, miR-320, miR-324, miR-325, miR-326, miR-330, miR-331, miR-338, miR-341, miR-369, miR-370, et-7a, Let-7b, Let-7c, Let-7d, Let-7g, miR-7, miR-9, miR-10a, miR-10b, miR-15a, miR-18, miR-19a, miR-17-3p, miR-20, miR-23b, miR-25, miR-26a, miR-26a, miR-30e-5p, miR-31, miR-32, miR-92, miR-93, miR-100, miR-125a, miR-125b, miR-126, miR-127, miR-128, miR-129, miR-130a, miR-135, miR-138, miR-139, miR-140, miR-141, miR-143, miR-145, miR-146, miR-150, miR-154, miR-155, miR-181a, miR-182, miR-186, miR-187, miR-188, miR-190, miR-191, miR-193, miR-194, miR-196, miR-197, miR-198, miR-199, miR-201, miR-204, miR-216, miR-218, miR-223, miR-293, miR-291-3p, miR-294, miR-295, miR-322, miR-333, miR-335, miR-338, miR-341, miR-350, miR-369, miR-373, miR-410, and miR-412. Detection one or more of the above miRs can be used to characterize a GI cancer.

[00320] Other microRNAs that are expressed commonly in solid cancer, such as colon cancer, lung cancer, breast cancer, stomach (gastric) cancer, prostate cancer, and pancreatic cancer, can also be detected in a vesicle and used to characterize a cancer. For example, one or more of the following miRs: miR-21, miR-17-5p, miR-

191, miR-29b-2, miR-223, miR-128b, miR-199a-1, miR-24-1, miR-24-2, miR-146, miR-155, miR-181b-1, miR-20a, miR-107, miR-32, miR-92-2, miR-214, miR-30c, miR-25, miR-221, and miR-106a, can be detected in a vesicle and used to characterize a solid cancer.

[00321] Other examples of microRNAs that can be detected in solution or in a vesicle are disclosed in PCT Publication Nos. WO2006126040, WO2006033020, WO2005116250, and WO2005111211, US Publication Nos. US20070042982 and US20080318210; and EP Publication Nos. EP1784501A2 and EP1751311A2, each of which is incorporated by reference.

[00322] As depicted in **FIG. 4**, a GI cancer, such as colon cancer, a biosignature can comprise detection of EpCam, CD63, CD81, CD9, CD66, or any combination thereof, of a vesicle. Furthermore, a colon cancer-biosignature for various stages of cancer can comprise CD63, CD9, EpCam, or any combination thereof. See, e.g., **FIGs. 5** and **6**. For example, the biosignature can comprise CD9 and EpCam. In some embodiments, the GI cancer biosignature comprises one or more miRNA selected from the group consisting of miR-548c-5p, miR-362-3p, miR-422a, miR-597, miR-429, miR-200a, and miR-200b. These miRNAs can be overexpressed in GI cancers, as shown in **FIG. 7**. The miRNA signature can be combined with the biomarkers listed above. The biosignatures can provide a diagnostic, prognostic or theranostic profile, such as the stage of the cancer, the efficacy of the cancer, or other characteristics of the cancer.

[00323] A biosignature for gastrointestinal cancer can be used to assess the efficacy of a therapy. For example, biomarkers that are elevated in a cancer of the GI tract can be monitored before and after a treatment. A reduction in the level of the biomarker post-treatment can indicate that the treatment is efficacious. The same biosignature can be monitored overtime, e.g., to detect recurrence or relapse post-treatment. As a non-limiting example, mutation in K-ras can be used to guide treatment of colorectal cancer with Panitumumab or Cetuximab. In yet another embodiment, a biosignature comprising one or more biomarkers selected from the group consisting of: miR-548c-5p, miR-362-3p, miR-422a, miR-597, miR-429, miR-200a, and miR-200b, is used for the theranosis of colorectal cancer.

[00324] Assessing a plurality of biomarkers, such as those listed above, can provide increased sensitivity, specificity, or signal intensity, as compared to assessing less than a plurality of biomarkers. For example, assessing CD24 and EpCam can provide increased sensitivity in detection as compared to assessing CD24 or EpCam alone. Assessing CD9 and CD63 can provide increased sensitivity in detection as compared to assessing CD9 or CD63 alone. In one embodiment, one or more the following biomarkers are detected to characterize a GI disorder: EpCam, CD9, CD63, CD81, B7H3, ICAM, STEAP, and EGFR. In another embodiment, one or more of the following biomarkers are detected to characterize a GI disorder: DR3, STEAP, epha2, TMEM211, unc93A, A33, CD24, NGAL, EpCam, MUC17, TROP2, TETS. The GI disorder can be colorectal cancer.

[00325] In some embodiments, characterization of a GI disorder comprises the analysis of one or more vesicles. The markers used to capture and/or detect the vesicles include one or more general vesicle markers, one or more markers indicative of the cell-of-origin of the GI disorder, and one or more disease specific marker. In other embodiments, the markers used to capture and/or detect characterization of a GI disorder the vesicles include one or more general vesicle markers, and one or more disease specific marker. In still other embodiments, the markers used to capture and/or detect the vesicles include one or more general vesicle markers, and one or more markers indicative of the cell-of-origin of the GI disorder. In an embodiment, the markers used to capture

and/or detect the vesicles include one or more markers indicative of the cell-of-origin of the GI disorder, and one or more disease specific marker. The general vesicle markers can include a tetraspanin, e.g., one or more of CD9, CD63 and CD81. The cell-of-origin or disease markers can be any of those described herein, such as DR3, STEAP, epha2, TMEM211, unc93A, A33, CD24, NGAL, EpCam, MUC17, TROP2, or TETS. The characterization of the GI disorder can further comprise analysis of the payload in the captured vesicles. In an embodiment, the payload comprises one or more of miR-548c-5p, miR-362-3p, miR-422a, miR-597, miR-429, miR-200a, and miR-200b. The level of vesicles and/or payload therein can be compared to a reference, thereby characterizing the GI disorder. The analysis can be performed in vitro on a sample from a subject. The sample can be a bodily fluid, such as blood, plasma or serum. The sample can also comprise stool, bile, pancreatic fluid, or other fluids of the GI tract. The invention further provides use of an agent in preparing a composition for use in carrying out the subject invention.

EXAMPLES

Example 1: Purification of Vesicles From Cell Lines

[00326] Prostate cancer cell lines are cultured for 3-4 days in culture media containing 20% FBS (fetal bovine serum) and 1% P/S/G. The cells are then pre-spun for 10 minutes at 400x g at 4°C. The supernatant is kept and centrifuged for 20 minutes at 2000 x g at 4. The supernatant containing vesicles can be concentrated using a Millipore Centricon Plus-70 (Cat # UFC710008 Fisher).

[00327] The Centricon is pre washed with 30mls of PBS at 1000 x g for 3 minutes at room temperature. Next, 15- 70 mls of the pre-spun cell culture supernatant is poured into the Concentrate Cup and is centrifuged in a Swing Bucket Adapter (Fisher Cat # 75-008-144) for 30 minutes at 1000 x g at room temperature.

[00328] The flow through in the Collection Cup is poured off. The volume in the Concentrate Cup is brought back up to 60mls with any additional supernatant. The Concentrate Cup is centrifuged for 30 minutes at 1000 x g at room temperature to concentrate the cell supernatant.

[00329] The Concentrate Cup is washed by adding 70mls of PBS and centrifuged for 30-60 minutes at 1000 x g until approximately 2 mls remains. The vesicles are removed from the filter by inverting the concentrate into the small sample cup and centrifuge for 1 minute at 4°C. The volume is brought up to 25 mls with PBS. The vesicles are now concentrated and are added to a 30% Sucrose Cushion.

[00330] To make a cushion, 4 mls of Tris/30%Sucrose/D2O solution (30g protease-free sucrose, 2.4g Tris base, 50ml D2O, adjust pH to 7.4 with 10N NCL drops, adjust volume to 100mls with D2O, sterilize by passing thru a 0.22-um filter) is loaded to the bottom of a 30ml V bottom thin walled Ultracentrifuge tube. The diluted 25 mls of concentrated vesicles is gently added above the sucrose cushion without disturbing the interface and is centrifuged for 75 minutes at 100,000 x g at 4°C. The ~25mls above the sucrose cushion is carefully removed with a 10ml pipet and the ~3.5mls of vesicles is collected with a fine tip transfer pipet (SAMCO 233) and transferred to a fresh ultracentrifuge tube, where 30 mls PBS is added. The tube is centrifuged for 70 minutes at 100,000 x g at 4°C. The supernatant is poured off carefully. The pellet is resuspended in 200ul PBS and can be stored at 4°C or used for assays. A BCA assay (1:2) can be used to determine protein content and Western blotting or electron micrography can be used to determine vesicle purification.

Example 2: Purification of Vesicles from VCaP and 22Rv1

[00331] Vesicles from Vertebral-Cancer of the Prostate (VCaP) and 22Rv1, a human prostate carcinoma cell line, derived from a human prostatic carcinoma xenograft (CWR22R) were collected by ultracentrifugation by first diluting plasma with an equal volume of PBS (1 ml). The diluted fluid was transferred to a 15 ml falcon tube and centrifuged 30 minutes at 2000 x g 4°C. The supernatant (~2 mls) was transferred to an ultracentrifuge tube 5.0 ml PA thinwall tube (Sorvall # 03127) and centrifuged at 12,000 x g, 4°C for 45 minutes.

[00332] The supernatant (~2 mls) was transferred to a new 5.0 ml ultracentrifuge tubes and filled to maximum volume with addition of 2.5 mls PBS and centrifuged for 90 minutes at 110,000 x g, 4°C. The supernatant was poured off without disturbing the pellet and the pellet resuspended with 1 ml PBS. The tube was filled to maximum volume with addition of 4.5 ml of PBS and centrifuged at 110,000 x g, 4°C for 70 minutes.

[00333] The supernatant was poured off without disturbing the pellet and an additional 1 ml of PBS was added to wash the pellet. The volume was increased to maximum volume with the addition of 4.5 mls of PBS and centrifuged at 110,000 x g for 70 minutes at 4°C. The supernatant was removed with P-1000 pipette until ~ 100 µl of PBS was in the bottom of the tube. The ~ 90 µl remaining was removed with P-200 pipette and the pellet collected with the ~10 µl of PBS remaining by gently pipetting using a P-20 pipette into the microcentrifuge tube. The residual pellet was washed from the bottom of a dry tube with an additional 5 µl of fresh PBS and collected into microcentrifuge tube and suspended in phosphate buffered saline (PBS) to a concentration of 500 µg/ml.

Example 3: Plasma Collection and Vesicle Purification

[00334] Blood is collected via standard veinpuncture in a 7ml K2-EDTA tube. The sample is spun at 400g for 10 minutes in a 4°C centrifuge to separate plasma from blood cells (SORVALL Legend RT+ centrifuge). The supernatant (plasma) is transferred by careful pipetting to 15ml Falcon centrifuge tubes. The plasma is spun at 2,000g for 20 minutes and the supernatant is collected.

[00335] For storage, approximately 1ml of the plasma (supernatant) is aliquoted to a cryovials, placed in dry ice to freeze them and stored in -80°C. Before vesicle purification, if samples were stored at -80°C, samples are thawed in a cold water bath for 5 minutes. The samples are mixed end over end by hand to dissipate insoluble material.

[00336] In a first prespin, the plasma is diluted with an equal volume of PBS (example, approximately 2 ml of plasma is diluted with 2 ml of PBS). The diluted fluid is transferred to a 15 ml Falcon tube and centrifuged for 30 minutes at 2000 x g at 4°C.

[00337] For a second prespin, the supernatant (approximately 4 mls) is carefully transferred to a 50 ml Falcon tube and centrifuged at 12,000 x g at 4°C for 45 minutes in a Sorval.

[00338] In the isolation step, the supernatant (approximately 2 mls) is carefully transferred to a 5.0 ml ultracentrifuge PA thinwall tube (Sorvall # 03127) using a P1000 pipette and filled to maximum volume with an additional 0.5 mls of PBS. The tube is centrifuged for 90 minutes at 110,000 x g at 4°C.

[00339] In the first wash, the supernatant is poured off without disturbing the pellet. The pellet is resuspended or washed with 1 ml PBS and the tube is filled to maximum volume with an additional 4.5 ml of PBS. The tube is centrifuged at 110,000 x g at 4°C for 70 minutes. A second wash is performed by repeating the same steps.

[00340] The vesicles are collected by removing the supernatant with P-1000 pipette until approximately 100 µl of PBS is in the bottom of the tube. Approximately 90 µl l of the PBS is removed and discarded with P-200

pipette. The pellet and remaining PBS is collected by gentle pipetting using a P-20 pipette. The residual pellet is washed from the bottom of the dry tube with an additional 5 µl of fresh PBS and collected into a microcentrifuge tube.

Example 4: Analysis of Vesicles Using Antibody-Coupled Microspheres and Directly Conjugated Antibodies

[00341] This example demonstrates the use of particles coupled to an antibody, where the antibody captures the vesicles (see for example, **FIG. 2B**). An antibody, the detector antibody, is directly coupled to a label, and is used to detect a biomarker on the captured vesicle.

[00342] First, an antibody-coupled microsphere set is selected (Luminex, Austin, TX). The microsphere set can comprise various antibodies, and thus allows multiplexing. The microspheres are resuspended by vortex and sonication for approximately 20 seconds. A Working Microsphere Mixture is prepared by diluting the coupled microsphere stocks to a final concentration of 100 microspheres of each set/µL in Startblock (Pierce (37538)). (Note: 50 µL of Working Microsphere Mixture is required for each well.) Either PBS-1% BSA or PBS-BN (PBS, 1% BSA, 0.05% Azide, pH 7.4) may be used as Assay Buffer.

[00343] A 1.2 µm Millipore filter plate is pre-wet with 100 µl/well of PBS-1% BSA (Sigma (P3688-10PAK + 0.05% NaAzide (S8032))) and aspirated by vacuum manifold. An aliquot of 50 µl of the Working Microsphere Mixture is dispensed into the appropriate wells of the filter plate (Millipore Multiscreen HTS (MSBVN1250)). A 50 µl aliquot of standard or sample is dispensed into to the appropriate wells. The filter plate is covered and incubated for 60 minutes at room temperature on a plate shaker. The plate is covered with a sealer, placed on the orbital shaker and set to 900 for 15-30 seconds to re-suspend the beads. Following that the speed is set to 550 for the duration of the incubation.

[00344] The supernatant is aspirated by vacuum manifold (less than 5 inches Hg in all aspiration steps). Each well is washed twice with 100 µl of PBS-1% BSA (Sigma (P3688-10PAK + 0.05% NaAzide (S8032))) and is aspirated by vacuum manifold. The microspheres are resuspended in 50 µL of PBS-1% BSA (Sigma (P3688-10PAK + 0.05% NaAzide (S8032))). The PE conjugated detection antibody is diluted to 4 µg/mL (or appropriate concentration) in PBS-1% BSA (Sigma (P3688-10PAK + 0.05% NaAzide (S8032))). (Note: 50 µL of diluted detection antibody is required for each reaction.) A 50 µl aliquot of the diluted detection antibody is added to each well. The filter plate is covered and incubated for 60 minutes at room temperature on a plate shaker. The filter plate is covered with a sealer, placed on the orbital shaker and set to 900 for 15-30 seconds to re-suspend the beads. Following that the speed is set to 550 for the duration of the incubation. The supernatant is aspirated by vacuum manifold. The wells are washed twice with 100 µl of PBS-1% BSA (Sigma (P3688-10PAK + 0.05% NaAzide (S8032))) and aspirated by vacuum manifold. The microspheres are resuspended in 100 µl of PBS-1% BSA (Sigma (P3688-10PAK + 0.05% NaAzide (S8032))). The microspheres are analyzed on a Luminex analyzer according to the system manual.

Example 5: Analysis of Vesicles Using Antibody-Coupled Microspheres and Biotinylated Antibody

[00345] This example demonstrates the use of particles coupled to an antibody, where the antibody captures the vesicles. An antibody, the detector antibody, is biotinylated. A label coupled to streptavidin is used to detect the biomarker.

[00346] First, the appropriate antibody-coupled microsphere set is selected (Luminex, Austin, TX). The microspheres are resuspended by vortex and sonication for approximately 20 seconds. A Working Microsphere

Mixture is prepared by diluting the coupled microsphere stocks to a final concentration of 50 microspheres of each set/μL in Startblock (Pierce (37538)). (Note: 50 μl of Working Microsphere Mixture is required for each well.) Beads in Start Block should be blocked for 30 minutes and no more than 1 hour.

[00347] A 1.2 μm Millipore filter plate is pre-wet with 100 μl /well of PBS-1% BSA + Azide (PBS-BN)((Sigma (P3688-10PAK + 0.05% NaAzide (S8032))) and is aspirated by vacuum manifold. A 50 μl aliquot of the Working Microsphere Mixture is dispensed into the appropriate wells of the filter plate (Millipore Multiscreen HTS (MSBVN1250)). A 50 μl aliquot of standard or sample is dispensed to the appropriate wells. The filter plate is covered with a seal and is incubated for 60 minutes at room temperature on a plate shaker. The covered filter plate is placed on the orbital shaker and set to 900 for 15-30 seconds to re-suspend the beads. Following that, the speed is set to 550 for the duration of the incubation.

[00348] The supernatant is aspirated by a vacuum manifold (less than 5 inches Hg in all aspiration steps). Aspiration can be done with the Pall vacuum manifold. The valve is place in the full off position when the plate is placed on the manifold. To aspirate slowly, the valve is opened to draw the fluid from the wells, which takes approximately 3 seconds for the 100 μl of sample and beads to be fully aspirated from the well. Once the sample drains, the purge button on the manifold is pressed to release residual vacuum pressure from the plate.

[00349] Each well is washed twice with 100 μl of PBS-1% BSA + Azide (PBS-BN) (Sigma (P3688-10PAK + 0.05% NaAzide (S8032))) and is aspirated by vacuum manifold. The microspheres are resuspended in 50 μl of PBS-1% BSA+ Azide (PBS-BN) ((Sigma (P3688-10PAK + 0.05% NaAzide (S8032))).

[00350] The biotinylated detection antibody is diluted to 4 μg/mL in PBS-1% BSA + Azide (PBS-BN) (Sigma (P3688-10PAK + 0.05% NaAzide (S8032))). (Note: 50 μl of diluted detection antibody is required for each reaction.) A 50 μl aliquot of the diluted detection antibody is added to each well.

[00351] The filter plate is covered and incubated with shaking as described above. The supernatant is aspirated by vacuum manifold as described above. The wells are washed and resuspended with PBS-BN as described above.

[00352] The streptavidin-R-phycoerythrin reporter (Molecular Probes 1 mg/ml) is diluted to 4 μg/mL in PBS-1% BSA+ Azide (PBS-BN). 50 μl of diluted streptavidin-R-phycoerythrin is required for each reaction. A 50 μl aliquot of the diluted streptavidin-R-phycoerythrin is added to each well.

[00353] The filter plate is covered and incubated with shaking as described above. The supernatant is aspirated by vacuum manifold as described above.

[00354] Each well is washed twice with 100 μl of PBS-1% BSA + Azide (PBS-BN) ((Sigma (P3688-10PAK + 0.05% NaAzide (S8032))) and is aspirated by vacuum manifold. The microspheres are resuspended in 100 μl of PBS-1% BSA+ Azide (PBS-BN) and analyzed on the Luminex analyzer according to the system manual.

Example 6: Reference Values for Cancer Detection

[00355] Fourteen stage 3 prostate cancer subjects, eleven benign prostate hyperplasia (BPH) samples, and 15 normal samples were tested. Vesicle samples were obtained using methods as described in Example 3 and used in multiplexing assays, such as described in Examples 4 and 5. The samples were analyzed to determine four criteria 1) if the sample has overexpressed vesicles, 2) if the sample has overexpressed prostate vesicles, 3) if the sample has overexpressed cancer vesicles, and 4) if the sample is reliable. If the sample met all four criteria, the categorization of the sample as positive for prostate cancer had varying sensitivities and specificities, depending on the different biosignatures present for a sample as shown in **Table 4**.

[00356] The four criteria were as follows:

[00357] *Vesicle Overexpression*

[00358] The mean fluorescence intensities (MFIs) for a sample in three assays were averaged to determine a value for the sample. Each assay used a different capture antibody. The first used a CD9 capture antibody, the second a CD81 capture antibody, and the third a CD63 antibody. The same combination of detection antibodies was used for each assay, antibodies for CD9, CD81, and CD63. If the average value obtained for the three assays was greater than 3000, the sample was categorized as having overexpressed vesicles (**Table 4**, Vesicle).

[00359] *Prostate Vesicle Overexpression*

[00360] The MFIs for a sample in two assays were averaged to determine a value for the sample. Each assay used a different capture antibody. The first used a PCSA capture antibody and the second used a PSMA capture antibody. The same combination of detection antibodies was used for each assay, antibodies for CD9, CD81, and CD63. If the average value obtained for the two assays was greater than 100, the sample was categorized as having prostate vesicles overexpressed (**Table 4**, Prostate).

[00361] *Cancer Vesicle Overexpression*

[00362] Three different cancer biosignatures were used to determine if cancer vesicles were overexpressed in a sample. The first, Cancer-1, used an EpCam capture antibody and detection antibodies for CD81, CD9, and CD63. The second, Cancer-2, used a CD9 capture antibody with detection antibodies for EpCam and B7H3. If the MFI value of a sample for any two of the three cancer biosignatures was above a reference value, the sample was categorized as having overexpressed cancer (see **Table 4**, Ca-1, Ca-2, Ca-3).

[00363] *Reliability of Sample*

[00364] Two quality control measures, QC-1 and QC-2, were determined for each sample. If the sample met one of them, the sample was categorized as reliable.

[00365] For QC-1, the sum of all the MFIs of 7 assays was determined. Each of the 7 assays used detection antibodies for CD59 and PSMA. The capture antibody used for each assay was CD63, CD81, PCSA, PSMA, STEAP, B7H3, and EpCam. If the sum was greater than 4000, the sample was not reliable and not included.

[00366] For QC-2, the sum of all the MFIs of 5 assays was determined. Each of the 5 assays used detection antibodies for CD9, CD81 and CD63. The capture antibody used for each assay was PCSA, PSMA, STEAP, B7H3, and EpCam. If the sum was greater than 8000, the sample was not reliable and not included.

[00367] **Table 4** shows the sensitivity and specificity for different prostate signatures. "Vesicle" lists the threshold value or reference value of vesicle levels, "Prostate" lists the threshold value or reference value used for prostate vesicles, "Ca-1," "Ca-2," and "Ca-3" lists the threshold values or reference values for the three different biosignatures for prostate cancer, the "QC-1" and "QC-2" columns list the threshold values or reference values for quality control, or reliability, and the last four columns list the specificities ("Spec") and sensitivities ("Sens") for benign prostate hyperplasia (BPH).

Table 4: Sensitivity and Specificity for Cancer Signatures

Vesicle	Prostate	Ca-1	Ca-2	Ca-3	QC-1	QC-2	Sens	Spec	Sens	Spec
							With BPH	With BPH	Without BPH	Without BPH
3000	100	na	200	na	4000	na	85.70%	58.00%	85.70%	71.40%

3000	100	350	100	na	4000	na	85.70%	74.10%	85.70%	85.70%
3000	100	125	125	50	4000	na	71.40%	83.00%	71.40%	90.40%
3000	100	100	100	50	4000	8000	71.40%	87.00%	71.40%	90.40%
3000	100	100	150	50	4000	na	64.30%	90.30%	64.20%	90.40%
3000	100	100	150	150	4000	na	35.70%	93.40%	35.70%	95.20%

Example 7: Determining Biosignatures for Colorectal Cancer Using Multiplexing

[00368] The samples obtained using methods as described in Example 1-3 are used in multiplexing assays following methodology as described in Examples 4 and 5. The detection antibodies used are CD63, CD9, CD81, B7H3 and EpCam. The capture antibodies used are CD9, PSCA, TNFR, CD63 (2 antibodies), B7H3, MFG-E8, EpCam (2 antibodies), CD63, Rab, CD81, STEAP, PCSA, PSMA, 5T4, CD24, TMEM211, Rab IgG (control) and IgG (control), resulting in 100 combinations to be screened (**FIG. 2C**).

[00369] The results are depicted in **FIGs. 4**, **5,** and **6**. The sensitivity of the different combinations is depicted in **FIG. 8**.

Example 8: Capture of Vesicles Using Magnetic Beads

[00370] Vesicles isolated as described in Example 2 are used. Approximately 40 ul of the vesicles are incubated with approximately 5 μg (~50 μl) of EpCam antibody coated Dynal beads (Invitrogen, Carlsbad, CA) and 50 μl of Starting Block. The vesicles and beads are incubated with shaking for 2 hours at 45°C in a shaking incubator. The tube containing the Dynal beads is placed on the magnetic separator for 1 minute and the supernatant removed. The beads are washed twice and the supernatant removed each time. Wash beads twice, discarding the supernatant each time.

Example 9: Vesicle Cancer Assay/Test

[00371] In this example, the vesicle (e.g., exosome, microvesicle, etc) PCa test is a microsphere based immunoassay for the detection of a set of protein biomarkers present on the vesicles from plasma of patients with prostate cancer. The test employs specific antibodies to the following protein biomarkers: CD9, CD59, CD63, CD81, PSMA, PCSA, B7H3 and EpCAM (**FIG. 9A**). After capture of the vesicles by antibody coated microspheres, phycoerythrin-labeled antibodies are used for the detection of vesicle specific biomarkers. Depending on the level of binding of these antibodies to the vesicles from a patient's plasma a determination of the presence or absence of prostate cancer is made.

[00372] Vesicles are isolated as described in Example 1.

[00373] *Microspheres*

[00374] Specific antibodies are conjugated to microspheres (Luminex) after which the microspheres are combined to make a Microsphere Master Mix consisting of L100-C105-01; L100-C115-01; L100-C119-01; L100-C120-01; L100-C122-01; L100-C124-01; L100-C135-01; and L100-C175-01. xMAP® Classification Calibration Microspheres L100-CAL1 (Luminex) are used as instrument calibration reagents for the Luminex LX200 instrument. xMAP® Reporter Calibration Microspheres L100-CAL2 (Luminex) are used as instrument reporter calibration reagents for the Luminex LX200 instrument. xMAP® Classification Control Microspheres L100-CON1 (Luminex) are used as instrument control reagents for the Luminex LX200 instrument. xMAP

Reporter Control Microspheres L100-CON2 (Luminex) and are used as reporter control reagents for the Luminex LX200 instrument.

[00375] *Capture Antibodies*

[00376] The following antibodies are used to coat Luminex microspheres for use in capturing certain populations of vesicles by binding to their respective protein targets on the vesicles in this Example: a. Mouse anti-human CD9 monoclonal antibody is an IgG2b used to coat microsphere L100-C105 to make *EPCLMACD9-C105; b. Mouse anti-human PSMA monoclonal antibody is an IgG1 used to coat microsphere L100-C115 to make EPCLMAPSMA-C115; c. Mouse anti-human PCSA monoclonal antibody is an IgG1 used to coat microsphere L100-C119 to make EPCLMAPCSA-C119; d. Mouse anti-human CD63 monoclonal antibody is an IgG1 used to coat microsphere L100-C120 to make EPCLMACD63-C120; e. Mouse anti-human CD81 monoclonal antibody is an IgG1 used to coat microsphere L100-C124 to make EPCLMACD81-C124; f. Goat anti-human B7-H3 polyclonal antibody is an IgG purified antibody used to coat microsphere L100-C125 to make EPCLGAB7-H3-C125; and g. Mouse anti-human EpCAM monoclonal antibody is an IgG2b purified antibody used to coat microsphere L100-C175 to make EPCLMAEpCAM-C175.

[00377] *Detection Antibodies*

[00378] The following phycoerythrin (PE) labeled antibodies are used as detection probes in this assay: a. EPCLMACD81PE: Mouse anti-human CD81 PE labeled antibody is an IgG1 antibody used to detect CD81 on captured vesicles; b. EPCLMACD9PE: Mouse anti-human CD9 PE labeled antibody is an IgG1 antibody used to detect CD9 on captured vesicles; c. EPCLMACD63PE: Mouse anti-human CD63 PE labeled antibody is an IgG1 antibody used to detect CD63 on captured vesicles; d. EPCLMAEpCAMPE: Mouse anti-human EpCAM PE labeled antibody is an IgG1 antibody used to detect EpCAM on captured vesicles; e. EPCLMAPSMAPE: Mouse anti-human PSMA PE labeled antibody is an IgG1 antibody used to detect PSMA on captured vesicles; f. EPCLMACD59PE: Mouse anti-human CD59 PE labeled antibody is an IgG1 antibody used to detect CD59 on captured vesicles; and g. EPCLMAB7-H3PE: Mouse anti-human B7-H3 PE labeled antibody is an IgG1 antibody used to detect B7-H3 on captured vesicles.

[00379] *Reagent Preparation*

[00380] Antibody Purification: The following antibodies in **Table 5** are received from vendors and purified and adjusted to the desired working concentrations according to the following protocol.

Table 5: Antibodies for PCa Assay

Antibody	Use
EPCLMACD9	Coating of microspheres for vesicle capture
EPCLMACD63	Coating of microspheres for vesicle capture
EPCLMACD81	Coating of microspheres for vesicle capture
EPCLMAPSMA	Coating of microspheres for vesicle capture
EPCLGAB7-H3	Coating of microspheres for vesicle capture
EPCLMAEpCAM	Coating of microspheres for vesicle capture
EPCLMAPCSA	Coating of microspheres for vesicle capture
EPCLMACD81PE	PE coated antibody for vesicle biomarker detection
EPCLMACD9PE	PE coated antibody for vesicle biomarker detection
EPCLMACD63PE	PE coated antibody for vesicle biomarker detection

EPCLMAEpCAMPE	PE coated antibody for vesicle biomarker detection
EPCLMAPSMAPE	PE coated antibody for vesicle biomarker detection
EPCLMACD59PE	PE coated antibody for vesicle biomarker detection
EPCLMAB7-H3PE	PE coated antibody for vesicle biomarker detection

[00381] Antibody Purification Protocol: Antibodies are purified using Protein G resin from Pierce (Protein G spin kit, prod # 89979). Micro-chromatography columns made from filtered P-200 tips are used for purification.

[00382] One hundred µl of Protein G resin is loaded with 100µl buffer from the Pierce kit to each micro column. After waiting a few minutes to allow the resin to settle down, air pressure is applied with a P-200 Pipettman to drain buffer when needed, ensuring the column is not let to dry. The column is equilibrated with 0.6ml of Binding Buffer (pH 7.4, 100mM Phosphate Buffer, 150mM NaCl; (Pierce, Prod # 89979). An antibody is applied to the column (<1mg of antibody is loaded on the column). The column is washed with 1.5ml of Binding Buffer. Five tubes (1.5 ml micro centrifuge tubes) are prepared and 10 µl of neutralization solution (Pierce, Prod # 89979) is applied to each tube. The antibody is eluted with the elution buffer from the kit to each of the five tubes, 100ul for each tube (for a total of 500 µl). The relative absorbance of each fraction is measured at 280nm using Nanodrop (Thermo scientific, Nanodrop 1000 spectrophotometer). The fractions with highest OD reading are selected for downstream usage. The samples are dialyzed against 0.25 liters PBS buffer using Pierce Slide-A-Lyzer Dialysis Cassette (Pierce, prod 66333, 3KDa cut off). The buffer is exchanged every 2 hours for minimum three exchanges at 4°C with continuous stirring. The dialyzed samples are then transferred to 1.5ml microcentifuge tubes, and can be labeled and stored at 4°C (short term) or -20°C (long term).

[00383] Microsphere Working Mix Assembly: A microsphere working mix MWM101 includes the first four rows of antibody, microsphere and coated microsphere of **Table 6**.

Table 6: Antibody-Microsphere Combinations

Antibody	**Microsphere**	**Coated Microsphere**
EPCLMACD9	L100-C105	EPCLMACD9-C105
EPCLMACD63	L100-C120	EPCLMACD63-C120
EPCLMACD81	L100-C124	EPCLMACD81-C124
EPCLMAPSMA	L100-C115	EPCLMAPSMA-C115
EPCLGAB7-H3	L100-C125	EPCLGAB7-H3-C125
bEPCLMAEpCAM	L100-C175	EPCLMAEpCAM-C175
EPCLMAPCSA	L100-C119	EPCLMAPCSA-C119

[00384] Microspheres are coated with their respective antibodies as listed above according to the following protocol.

[00385] Protocol for Two-Step Carbodiimide Coupling of Protein to Carboxylated Microspheres: The microspheres should be protected from prolonged exposure to light throughout this procedure. The stock uncoupled microspheres are resuspended according to the instructions described in the Product Information Sheet provided with the microspheres (xMAP technologies, MicroPlex TM Microspheres). Five x 106 of the stock microspheres are transferred to a USA Scientific 1.5ml microcentrifuge tube. The stock microspheres are pelleted by microcentrifugation at ≥ 8000 x g for 1-2 minutes at room temperature. The supernatant is removed and the pelleted microspheres are resuspended in 100 µl of dH2O by vortex and sonication for approximately 20

seconds. The microspheres are pelleted by microcentrifugation at ≥ 8000 x g for 1-2 minutes at room temperature. The supernatant is removed and the washed microspheres are resuspended in 80 µl of 100 mM Monobasic Sodium Phosphate, pH 6.2 by vortex and sonication (Branson 1510, Branson ULTrasonics Corp.) for approximately 20 seconds. Ten µl of 50 mg/ml Sulfo-NHS (Thermo Scientific, Cat#24500) (diluted in dH20) is added to the microspheres and is mixed gently by vortex. Ten µl of 50 mg/ml EDC (Thermo Scientific, Cat# 25952-53-8) (diluted in dH20) is added to the microspheres and gently mixed by vortexing. The microspheres are incubated for 20 minutes at room temperature with gentle mixing by vortex at 10 minute intervals. The activated microspheres are pelleted by microcentrifugation at ≥ 8000 x g for 1-2 minutes at room temperature. The supernatant is removed and the microspheres are resuspended in 250 µl of 50 mM MES, pH 5.0 (MES, Sigma, Cat# M2933) by vortex and sonication for approximately 20 seconds. (Only PBS-1% BSA+ Azide (PBS-BN)((Sigma (P3688-10PAK + 0.05% NaAzide (S8032))) should be used as assay buffer as well as wash buffer.). The microspheres are then pelleted by microcentrifugation at ≥ 8000 x g for 1-2 minutes at room temperature.

[00386] The supernatant is removed and the microspheres are resuspended in 250 µl of 50 mM MES, pH 5.0 (MES, Sigma, Cat# M2933) by vortex and sonication for approximately 20 seconds. (Only PBS-1% BSA+ Azide (PBS-BN) ((Sigma (P3688-10PAK + 0.05% NaAzide (S8032))) should be used as assay buffer as well as wash buffer.). The microspheres are then pelleted by microcentrifugation at ≥ 8000 x g for 1-2 minutes at room temperature, thus completing two washes with 50 mM MES, pH 5.0.

[00387] The supernatant is removed and the activated and washed microspheres are resuspended in 100 µl of 50 mM MES, pH 5.0 by vortex and sonication for approximately 20 seconds. Protein in the amount of 125, 25, 5 or 1 µg is added to the resuspended microspheres. (Note: Titration in the 1 to 125 µg range can be performed to determine the optimal amount of protein per specific coupling reaction.). The total volume is brought up to 500 µl with 50 mM MES, pH 5.0. The coupling reaction is mixed by vortex and is incubated for 2 hours with mixing (by rotating on Labquake rotator, Barnstead) at room temperature. The coupled microspheres are pelleted by microcentrifugation at ≥ 8000 x g for 1-2 minutes at room temperature. The supernatant is removed and the pelleted microspheres are resuspended in 500 µL of PBS-TBN by vortex and sonication for approximately 20 seconds. (Concentrations can be optimized for specific reagents, assay conditions, level of multiplexing, etc. in use.).

[00388] The microspheres are incubated for 30 minutes with mixing (by rotating on Labquake rotator, Barnstead) at room temperature. The coupled microspheres are pelleted by microcentrifugation at ≥ 8000 x g for 1-2 minutes at room temperature. The supernatant is removed and the microspheres are resuspended in 1 ml of PBS-TBN by vortex and sonication for approximately 20 seconds. (Each time there is the addition of samples, detector antibody or SA-PE the plate is covered with a sealer and light blocker (such as aluminum foil), placed on the orbital shaker and set to 900 for 15-30 seconds to re-suspend the beads. Following that the speed should be set to 550 for the duration of the incubation.).

[00389] The microspheres are pelleted by microcentrifugation at ≥ 8000 x g for 1-2 minutes. The supernatant is removed and the microspheres are resuspended in 1 ml of PBS-TBN by vortex and sonication for approximately 20 seconds. The microspheres are pelleted by microcentrifugation at ≥ 8000 x g for 1-2 minutes (resulting in a total of two washes with 1 ml PBS-TBN).

[00390] *Protocol for microsphere assay:* For multiple phycoerythrin (PE) detector antibody, the preparation is as described in Example 4. One hundred µl is analyzed on the Luminex analyzer (Luminex 200, xMAP technologies) according to the system manual. (High PMT setting).

[00391] *Decision Tree:* A decision tree (**FIG. 9B**) using the results from the Luminex assay to determine if a subject has cancer. Threshold limits on the MFI is established and samples classified according to the result of MFI scores for the antibodies, to classify if a sample is PCa positive. **FIG. 9C** shows a decision tree in which a sample is classified as indeterminate if the MFI is within the standard deviation of the predetermined threshold. For validation, the sum of the MFI signal from PCSA, EpCAM and B7-H3 must be greater than 200 or the test is a 'No Test' meaning no result can be obtained.

Example 10: Detection of Cancer using Vesicle Multiplex Analysis

[00392] High quality training set samples were obtained from commercial suppliers. The samples comprised plasma from 42 normal prostate, 42 prostate cancer (PCa) and 15 BPH patients. The PCa samples included 4 stage III and the remainder state II. The samples were blinded until all laboratory work was completed.

[00393] The vesicles from the samples were obtained by filtration to eliminate particles greater than 1.5 microns, followed by column concentration and purification using hollow fiber membrane tubes. The samples were analyzed using a multiplexed bead-based assay system.

[00394] Antibodies to the following proteins were analyzed:

a. General Vesicle (MV) markers: CD9, CD81, and CD63
b. Prostate MV markers: PCSA
c. Cancer-Associated MV markers: EpCam and B7H3

[00395] Samples were required to pass a quality test as follows: if multiplexed fluorescence intensity (MFI) PSCA + MFI B7H3 + MFI EpCam < 200 then sample fails due to lack of signal above background. In the training set, six samples (three normals and three prostate cancers) did not achieve an adequate quality score and were excluded. An upper limit on the MFI was also established as follows: if MFI of EpCam is > 6300 then test is over the upper limit score and samples are deemed not cancer (i.e., "negative" for purposes of the test).

[00396] The samples were classified according to the result of MFI scores for the six antibodies to the training set proteins, wherein the following conditions must be met to classified as PCa positive:

a. Average MFI of General MV markers > 1500
b. PCSA MFI > 300
c. B7H3 MFI > 550
d. EpCam MFI 550 – 6300

[00397] Using the 84 normal and PCa training data samples, the test was found to be 98% sensitive and 95% specific for PCa vs normal samples. See **FIG. 10A**. The increased MFI of the PCa samples compared to normals is shown in **FIG. 10B**.

Example 11: Indentifying microRNA Payload in Vesicle Subpopulations

[00398] This Example identifies various vesicle subpopulations by their particular surface protein topography. Plasma-derived vesicle RNA content of each subpopulation was characterized for association with a cancer phenotype. The protein topography and RNA content of vesicles found in plasma from patients with cancer, benign prostatic hyperplasia (BPH), and unaffected individuals were profiled to characterize and identify the vesicle subpopulations that are indicative of a given disease state. The biosignatures of these vesicle

subpopulations can be used to develop a diagnostic platform to aid in the screening and diagnosis of various cancers. Plasma-derived vesicles are separated using flow cytometry (FACS) and cell sorting techniques into protein-specific subpopulations by using membrane-specific protein biomarkers (e.g. EpCam). Vesicles from prostate cancer (PCa) patients had the highest percentage of vesicles labeled with EpCam, PSMA and CD9 compared to vesicles from normal, BPH and colorectal cancer (CRC) patients. Additionally, for each subpopulation of vesicles separated by FACS, quantitative expression profiling of miRs was used to identify expression signatures specific to cancer patients.

[00399] The RNA content of various subpopulations of vesicles, defined by their membrane protein biosignature, can be unique. In a vesicle subpopulation with proteins CD9 and CD8l on the surface, miR 141 is significantly overexpressed in vesicles from prostate cancer (PCa) patient plasma compared to vesicles derived from normal plasma. miR 9 was significantly overexpressed in vesicles from BPH plasma in EpCam and PSMA vesicles when compared to vesicles of the same subpopulation isolated from normal and PCa plasma, thereby providing a signature to separate BPH and PCa samples. miR 491 was overexpressed in EpCam expressing vesicles derived from colon cancer plasma compared to normal and PCa.

Example 12: Vesicle Biosignature for Colorectal Cancer

[00400] Although colonoscopy is the gold standard to screen and identify colorectal cancer (CRC), it is estimated half of patients who are recommended for colonoscopy are not compliant. Often the lack of compliance is because many perceive a colonoscopy as an uncomfortable and invasive procedure. An ideal first step toward increasing participation in preventive strategies would be the development of a less invasive diagnostic test to identify those patients that have a blood-based biosignature indicative of the need for detection and biopsy by colonoscopy. This strategy would result in cancers being identified earlier and prevent disease-free individuals from undergoing an unnecessary invasive procedure. Current blood-based tests rely on increased levels of either carcinoembryonic antigen (CEA) or carbohydrate antigenic determinant (CA 19-9). Unfortunately, CEA and CA 19-9 are neither organ-specific nor tumor-specific.

[00401] The present invention provides a vesicle-based platform to identify patients with CRC using a vesicle-based biosignature derived from plasma samples. The vesicles comprise exosomes, which are endosome-derived vesicles between 40-100 nm in diameter that are secreted by most cell types, including tumor cells. The present invention provides a vesicle-specific assay that can diagnose CRC from surface membrane protein biosignatures on vesicles derived from peripheral blood of patients with CRC.

[00402] Biosignatures were derived from vesicles isolated from plasma of patients with and without CRC. Vesicle surface proteins (CD9, CD8l, CD63, EpCam, EGFR, and STEAP) were used in a multiplexed microsphere assay to capture and detect vesicles as described herein. The quantity of vesicles with significant concentrations of these surface proteins lead to the development of a vesicle-specific biosignature that differentiated CRC samples from normal.

[00403] Vesicles present in blood plasma of CRC patients provide a signature by which CRC can be diagnosed as early as histological grade 1. The biosignature comprises different vesicle surface membrane protein markers, which include both general vesicle and cancer-specific proteins. Measurement of the vesicle biosignature in plasma differentiated patients with CRC (n=20) diagnosed by biopsy from individuals from the general population (n= 20) with a sensitivity of 85% and specificity of 85%. The CRC samples analyzed were comprised of AJCC/UICC stage I (n = 10), IIA (n = 6), and IIIB (n = 4).

[00404] Biosignatures identified in vesicles derived from the blood of patients with CRC provide a sensitive and specific test that can assist physicians screen, diagnose and treat patients with CRC.

Example 13: Vesicle FACS Analysis

[00405] Purified plasma exosomes are assayed using the MoFlo XDP (BC, Fort Collins, CO, USA) and the median fluorescent intensity analyzed using the Summit 4.3 Software (BC, Fort Collins, CO, USA). Cells are labeled directly with antibodies, or beads or microspheres (e.g., magnetic, polystyrene, including BD FACS 7-color setup, catalog no. 335775) can be incorporated. Microplex microspheres are obtained from Luminex (Austin, TX, USA) and conjugated to the following antibodies, CD9 (Mouse anti-human CD9 , MAB1880, R&D Systems, Minneapolis, MN, USA), PSM (Mouse anti-human PSM , sc-73651, Santa Cruz, Santa Cruz, CA, USA), PCSA (Mouse anti-human Prostate Cell Surface Antigen , MAB4089, Millipore, MA, USA), CD63 (Mouse anti-human CD63 , 556019, BD Biosciences, San Jose, CA, USA), CD81 (Mouse anti-human CD81 , 555675, BD Biosciences, San Jose, CA, USA) B7-H3 (Goat anti-human B7-H3 , AF1027, R&D Systems, Minneapolis, MN, USA), EpCAM (Mouse anti-human EpCAM , MAB9601, R&D Systems, Minneapolis, MN, USA) using Sulfo-NHS, and EDC obtained from Pierce Thermo (Cat. No. 24510 and 22981, respectively, Rockford, Ill, USA).

[00406] Purified membrane vesicles (10ug/ml) are incubated with 5,000 microspheres for one hour at room temperature with shaking. The samples ae washed in FACS buffer (0.5% FBS/PBS) for 10 minutes at 1700 rpms. The detection antibodies are incubated at the manufacturer's recommended concentrations for one hour at room temperature with shaking. Following another wash with FACS buffer for 10 minutes at 1700 rpms, the samples are resuspended in 100ul FACS buffer and run on the FACS machine. Microspheres are sorted according to their detection antibody content into four different tubes. The first contains the population of microspheres with no detectors, the second with PE detectors, the third with FITC detectors, and the fourth with both PE and FITC detectors.

Example 14: Obtaining Serum Samples from Subjects

[00407] Blood is collected from subjects (both healthy subjects and subjects with prostate cancer) in EDTA tubes, citrate tubes or in a 10 ml Vacutainer SST plus Blood Collection Tube (BD367985 or BD366643, BD Biosciences). Blood is processed for plasma isolation within 2 h of collection.

[00408] Samples are allowed to sit at room temperature for a minimum of 30 min and a max of 2 h. Separation of the clot is accomplished by centrifugation at 1,000–1,300 x *g* at 4°C for 15–20 min. The serum is removed and dispensed in aliquots of 500 µl into 500 to 750 µl cryotubes. Specimens are stored at -80°C.

[00409] At a given sitting, the amount of blood drawn can range from ~20 to ~90 ml. Blood from several EDTA tubes is pooled and transferred to RNase/DNase-free 50-ml conical tubes (Greiner), and centrifuged at 1,200 x *g* at room temperature in a Hettich Rotanta 460R benchtop centrifuge for 10 min. Plasma is transferred to a fresh tube, leaving behind a fixed height of 0.5 cm plasma supernatant above the pellet to avoid disturbing the pellet. Plasma is aliquoted, with inversion to mix between each aliquot, and stored at -80°C.

Example 15: RNA Isolation From Human Plasma and Serum Samples

[00410] Four hundred µl of human plasma or serum is thawed on ice and lysed with an equal volume of 2X Denaturing Solution (Ambion). RNA is isolated using the *mir*Vana PARIS kit following the manufacturer's protocol for liquid samples (Ambion), modified such that samples are extracted twice with an equal volume of

acid-phenol chloroform (as supplied by the Ambion kit). RNA is eluted with 105 µl of Ambion elution solution according to the manufacturer's protocol. The average volume of eluate recovered from each column is about 80 µl.

[00411] A scaled-up version of the *mir*Vana PARIS (Ambion) protocol is also used: 10 ml of plasma is thawed on ice, two 5-ml aliquots are transferred to 50-ml tubes, diluted with an equal volume of *mir*Vana PARIS 2X Denaturing Solution, mixed thoroughly by vortexing for 30 s and incubated on ice for 5 min. An equal volume (10 ml) of acid/phenol/chloroform (Ambion) is then added to each aliquot. The resulting solutions are vortexed for 1 min and spun for 5 min at 8,000 rpm, 20°C in a JA17 rotor. The acid/phenol/chloroform extraction is repeated three times. The resulting aqueous volume is mixed thoroughly with 1.25 volumes of 100% molecular-grade ethanol and passed through a *mir*Vana PARIS column in sequential 700-µl aliquots. The column is washed following the manufacturer's protocol, and RNA is eluted in 105 µl of elution buffer (95°C). A total of 1.5 µl of the eluate is quantified by Nanodrop.

Example 16: Measurement of miRNA Levels in RNA from Plasma and Serum using qRT-PCR

[00412] A fixed volume of 1.67 µl of RNA solution from about ~80 µl -eluate from RNA isolation of a given sample is used as input into the reverse transcription (RT) reaction. For samples in which RNA is isolated from a 400- µl plasma or serum sample, for example, 1.67 µl of RNA solution represents the RNA corresponding to (1.67/80) X 400 = 8.3 µl plasma or serum. For generation of standard curves of chemically synthesized RNA oligonucleotides corresponding to known miRNAs, varying dilutions of each oligonucleotide are made in water such that the final input into the RT reaction has a volume of 1.67 µl. Input RNA is reverse transcribed using the TaqMan miRNA Reverse Transcription Kit and miRNA-specific stem-loop primers (Applied BioSystems) in a small-scale RT reaction comprised of 1.387 µl of H2O, 0.5 µl of 10X Reverse-Transcription Buffer, 0.063 µl of RNase-Inhibitor (20 units / µl), 0.05 µl of 100 mM dNTPs with dTTP, 0.33 µl of Multiscribe Reverse-Transcriptase, and 1.67 µl of input RNA; components other than the input RNA can be prepared as a larger volume master mix, using a Tetrad2 Peltier Thermal Cycler (BioRad) at 16°C for 30 min, 42°C for 30 min and 85°C for 5 min. Real-time PCR is carried out on an Applied BioSystems 7900HT thermocycler at 95°C for 10 min, followed by 40 cycles of 95°C for 15 s and 60°C for 1 min. Data is analyzed with SDS Relative Quantification Software version 2.2.2 (Applied BioSystems.), with the automatic Ct setting for assigning baseline and threshold for Ct determination.

[00413] The protocol can also be modified to include a preamplification step, such as for detecting miRNA. A 1.25-µl aliquot of undiluted RT product is combined with 3.75 µl of Preamplification PCR reagents [comprised, per reaction, of 2.5 µl of TaqMan PreAmp Master Mix (2X) and 1.25 µl of 0.2X TaqMan miRNA Assay (diluted in TE)] to generate a 5.0-µl preamplification PCR, which is carried out on a Tetrad2 Peltier Thermal Cycler (BioRad) by heating to 95°C for 10 min, followed by 14 cycles of 95°C for 15 s and 60°C for 4 min. The preamplification PCR product is diluted (by adding 20 µl of H2O to the 5-µl preamplification reaction product), following which 2.25 µl of the diluted material is introduced into the real-time PCR and carried forward as described.

Example 17: Generation of Standard Curves for Absolute Quantification of miRNAs

[00414] Synthetic single-stranded RNA oligonucleotides corresponding to the mature miRNA sequence (miRBase Release v.10.1) are purchased from Sigma. Synthetic miRNAs are input into the RT reaction over an empirically-derived range of copies to generate standard curves for each of the miRNA TaqMan assays listed

above. In general, the lower limit of accurate quantification for each assay is designated based on the minimal number of copies input into an RT reaction that results in a Ct value within the linear range of the standard curve and that is also not equivalent to or higher than a Ct obtained from an RT input of lower copy number. A line is fit to data from each dilution series using Ct values within the linear range, from which y =mln(x) +b equations are derived for quantification of absolute miRNA copies (x) from each sample Ct (y). Absolute copies of miRNA input into the RT reaction are converted to copies of miRNA per microliter plasma (or serum) based on the knowledge that the material input into the RT reaction corresponds to RNA from 2.1% of the total starting volume of plasma [i.e., 1.67 µl of the total RNA eluate volume (80 µl on average) is input into the RT reaction]. An example of a synthetic miRNA sequence is for miR-141 which can be obtained commercially such as from Sigma (St. Louis, MO).

Example 18: Extracting microRNA from Vesicles

[00415] MicroRNA is extracted from vesicles isolated from patient samples. Methods for isolation and concentration of vesicles are presented in PCT Patent Application PCT/US09/06095, entitled "METHODS AND SYSTEMS OF USING EXOSOMES FOR DETERMINING PHENOTYPES" and filed November 12, 2009; and U.S. Patent Application Serial No. 12/609,847, entitled "METHODS FOR ASSESSING RNA PATTERNS" and filed October 30, 2009; both of which applications are incorporated by reference herein in their entirety. The methods can be used to isolate microRNA from patient samples without first isolating vesicles as well.

[00416] Protocol Using QIAzol & Midi Column

[00417] This protocol uses the QIAzol Lysis Reagent and RNeasy Midi Kit from Qiagen Inc., Valencia CA to extract microRNA from concentrated vesicles. The steps of the method comprise:

1. Add 2 µl of RNase A to 50 µl of vesicle concentrate, incubate at 37°C for 20 min.
2. Add 700 µl of QIAzol Lysis Reagent, vortex 1 minute. Spike samples with 25 fmol/µL of *C. elegans* microRNA (1 µL) after the addition of QIAzol, making a 75 fmol/µL spike in for each total sample (3 aliquots combined).
3. Incubate at 55°C for 5 min.
4. Add 140 µl chloroform and shake vigorously for 15 sec.
5. Cool on ice for 2-3 min.
6. Centrifuge @ 12,000 x g at 4°C for 15 min.
7. Transfer aqueous phase (300 µL) to a new tube and add 1.5 volumes of 100% EtOH (i.e., 450 µL).
8. Pipet up to 4 ml of sample into an RNeasy Midi spin column in a 15 ml collection tube (combining lysis from 3 50 µl of concentrate)
9. Spin at 2700 x g for 5 min at room temperature.
10. Discard flowthrough from the spin.
11. Add 1 ml of Buffer RWT to column and centrifuge at 2700 x g for 5 min at room temperature. Do not use Buffer RW1 supplied in the Midi kit. Buffer RW1 can wash away miRNA. Buffer RWT is supplied in the Mini kit from Qiagen Inc.
12. Discard flowthrough.
13. Add 1 ml of Buffer RPE onto the column and centrifuge at 2700 x g for 2 min at room temperature.
14. Repeat steps 12 and 13.

16. Place column into a new 15 ml collection tube and add 150 ul Elution Buffer. Incubate at room temperature for 3 min.

17. Centrifuge at 2700 x g for 3 min at room temperature.

18. Vortex the sample and transfer to 1.7 mL tube. Store the extracted sample at -80°C.

[00418] Protocol using MagMax

[00419] This protocol uses the MagMAX™ RNA Isolation Kit from Applied Biosystems/Ambion, Austin, TX to extract microRNA from concentrated vesicles. The steps of the method comprise:

1. Add 700 ml of QIAzol Lysis Reagent and vortex 1 minute.

2. Incubate on benchtop at room temperature for 5 min.

3. Add 140 µl chloroform and shake vigorously for 15 sec.

4. Incubate on benchtop for 2-3 min.

5. Centrifuge at 12,000 x g at 4°C for 15 min.

6. Transfer aqueous phase to a deep well plate and add 1.25 volumes of 100% Isopropanol.

7. Shake MagMAX™ binding beads well. Pipet 10 µl of RNA binding beads into each well.

8. Gather two elution plates and two additional deep well plates.

9. Label one elution plate "Elution" and the other "Tip Comb."

10. Label one deep well as "1st Wash 2" and the other as "2nd Wash 2."

11. Fill both Wash 2 deep well plates with 150 µl of Wash 2, being sure to add ethanol to wash beforehand. Fill in the same number of wells as there are samples.

12. Select the appropriate collection program on the MagMax Particle Processor.

13. Press start and load each appropriate plate.

14. Transfer samples to microcentrifuge tubes.

15. Vortex and store at -80°C. Residual beads will be seen in sample.

Example 19: MicroRNA Arrays

[00420] TaqMan Low Density Array

[00421] TaqMan Low Density Array (TLDA) miRNA cards are used to compare expression of miRNA in various sample groups as desired. The miRNA are collected and analyzed using the TaqMan® MicroRNA Assays and Arrays systems from Applied Biosystems, Foster City, CA. Applied Biosystems TaqMan® Human MicroRNA Arrays are used according to the Megaplex™ Pools Quick Reference Card protocol supplied by the manufacturer.

[00422] Exiqon mIRCURY LNA microRNA

[00423] The Exiqon miRCURY LNA™ Universal RT microRNA PCR Human Panels I and II (Exiqon, Inc, Woburn, MA) are used to compare expression of miRNA in various sample groups as desired. The Exiqon 384 well panels include 750 miRs. Samples are normalized to control primers towards synthetic RNA spike-in from Universal cDNA synthesis kit (UniSp6 CP). Results were normalized to inter-plate calibrator probes.

[00424] With either system, quality control standards are implemented. Normalized values for each probe across three data sets for each indication are averaged. Probes with an average CV% higher than 20% are not used for analysis. Results are subjected to a paired t-test to find differentially expressed miRs between two sample groups. P-values are corrected with a Benjamini and Hochberg false-discovery rate test. Results are analyzed using using GeneSpring GX 11.0 software (Agilent Technologies, Inc., Santa Clara, CA).

Example 20: MicroRNA Profiles in Vesicles

[00425] Vesicles were collected by ultracentrifugation from 22Rv1, LNCaP, Vcap and normal plasma (pooled from 16 donors) as described in Examples 1 and 2. RNA was extracted using the Exiqon miR isolation kit (Cat. No. 300110, 300111). Equals amounts of vesicles (30μg) were used as determined by BCA assay.

[00426] Equal volumes (5 μl) were put into a reverse-transcription reaction for microRNA. The reverse-transcriptase reactions were diluted in 81 μl of nuclease-free water and then 9 μl of this solution was added to each individual miR assay. MiR-629 was found to only be expressed in PCa (prostate cancer) vesicles and was virtually undetectable in normal plasma vesicles. MiR-9 was found to be highly overexpressed (~704 fold increase over normal as measured by copy number) in all PCa cell lines, and has very low expression in normal plasma vesicles. The top ten differentially expressed miRNAs are depicted in **FIG. 11**.

Example 21: MicroRNA Profiles of Magnetic EpCam-Captured Vesicles

[00427] The bead-bound vesicles of Example 8 were placed in QIAzol™ Lysis Reagent (Cat. #79306). An aliquot of 125 fmol of c. elegans miR-39 was added. The RNA was isolated using the Qiagen miRneasy™ kit, (Qiagen Inc., Valencia CA; Cat. # 217061), according to the manufacturer's instructions, and eluted in 30 ul RNAse free water.

[00428] 10 μl of the purified RNA was placed into a pre-amplifiation reaction for miR-9, miR-141 and miR-629 using a Veriti 96-well thermocycler. A 1:5 dilution of the pre-amplification solution was used to set up a qRT-PCR reaction for miR9 (ABI 4373285), miR-141 (ABI 4373137) and miR-629 (ABI 4380969) as well as c. elegans miR-39 (ABI 4373455). The results were normalized to the c. elegans results for each sample.

Example 22: MicroRNA Profiles of CD9-Captured Vesicles

[00429] CD9 coated Dynal beads (Invitrogen, Carlsbad, CA) are described above. Vesicles from prostate cancer patients, LNCaP, or normal purified vesicles were incubated with the CD9 coated beads and the RNA isolated as in Example 21. The expression of miR-21 and miR-141 was detected by qRT-PCR and the results depicted in **FIG. 12A** and **FIG. 12B**, respectively.

Example 23: MicroRNAs Overexpressed in Colorectal Cancer

[00430] TaqMan Low Density Array (TLDA) miRNA cards were used to compare expression of miRNA in CRC cell lines versus normal vesicles. The miRNA was collected and analyzed using the TaqMan® MicroRNA Assays and Arrays systems from Applied Biosystems, Foster City, CA. Applied Biosystems TaqMan® Human MicroRNA Arrays were used according to the Megaplex™ Pools Quick Reference Card protocol supplied by the manufacturer.

[00431] **FIG. 7** illustrates TLDA miRNA card comparison of colorectal cancer (CRC) cell lines versus normal vesicles. The cell lines include LOVO, HT29, SW260, COLO205, HCT116 and RKO. The Y-axis shows a fold-change in expression in the CRC cell lines compared to normal controls. The plot shows a 2-3 fold increase in expression in the CRC cell lines compared to normal controls. These miRNAs were not overexpressed in melanoma cells.

[00432] The sequences assayed in **FIG. 7** include miR-548c-5p, miR-362-3p, miR-422a, miR-597, miR-429, miR-200a, and miR-200b. Sequences of the miRNAs are shown in **Table 7**:

Table 7: microRNAs Differentially Regulated in CRC

Name	Sequence	miRNA Base Accession	SEQ ID NO.

hsa-miR-548c-5p	aaaaguaauugcgguuuuugcc	MIMAT0004806	1
hsa-miR-362-3p	aacacaccuauucaaggauuca	MI0000762	2
hsa-miR-422a	acuggacuuagggucagaaggc	MIMAT0001339	3
hsa-miR-597	ugugucacucgaugaccacugu	MIMAT0003265	4
hsa-miR-429	uaauacugucugguaaaaccgu	MIMAT0001536	5
hsa-miR-200a	uaacacugucugguaacgaugu	MI0000737	6
hsa-miR-200b	uaauacugccugguaaugauga	MI0000342	7

[00433] MicroRNA Threshold Levels

[00434] The microRNAs in **Table 7** are used to detect colorectal cancer in patient samples. Blood is withdrawn from a cohort of colorectal cancer patients and normal patients (i.e., not having CRC). MicroRNA is obtained from the samples as in Example 18. The level of the microRNAs in the CRC and normals is determined and optimal threshold levels are determined for each that differentiate CRC and normal using methodology as in Example 10. A naïve patient blood sample is obtained and the levels of the microRNAs therein are assessed. The levels of the microRNAs are compared to the thresholds to classify the sample as CRC or normal.

[00435] MicroRNA Classifier

[00436] The microRNAs in **Table 7** are used to detect colorectal cancer in patient samples. Blood is withdrawn from a cohort of colorectal cancer patients and normal patients (i.e., not having CRC). MicroRNA is obtained from the samples as in Example 18. The TLDA cards are run on the samples as above. A microRNA profile is obtained from the cards for the microRNAs in **Table 7** that differentiates the colorectal cancer and normal patients. A statistical classifier is developed using the microRNA profile using logistic regression. A naïve patient blood sample is obtained and the level of the microRNAs is assessed. The classifier is used to classify the microRNA profile in the naïve patient sample as CRC or normal.

Example 24: Vesicle Detection of Colorectal Cancer (CRC)

[00437] Concentrated vesicle plasma samples were run on a bead-based detection platform as described above. Antibodies to various vesicle surface antigens were attached to beads and used to capture vesicles. Several antibodies showed significant differences between samples derived from CRC and normal patients. The captured vesicles were labeled with PE-labeled antibodies to CD9, CD63, and CD81. The captured and labeled vesicles were detected using laser fluorescence.

[00438] CRC detection using antibody capture of vesicles was performed using 128 total samples consisting of vesicles isolated from plasma of 49 normals (i.e., healthy, non-CRC), 20 confounders (diseased, non-CRC), and 59 CRC. Confounder samples included those having rheumatoid arthritis, asthma, diabetes, bladder cell carcinoma, renal cell carcinoma, and chronic or acute diverticulitis. Of the CRC samples, 16 were Stage I, 19 were Stage II, and 24 were Stage III. **Table 8** below shows sensitivity and specificity obtained with various capture antibodies:

Table 8: CRC Detection using Antibody Capture of Vesicles

Specificity	86%	71%	83%	84%	84%	75%	77%	81%	62%	77%	80%	80%
Sensitivity	82%	100%	90%	100%	86%	100%	98%	94%	90%	86%	96%	86%
Marker	DR3	STEAP	Epha2	TMEM211	unc93A	A33	CD24	NGAL	EpCam	MUC17	TROP2	TETS

[00439] The transmembrane protein 211 (TMEM211) gene encodes a transmembrane protein. The mRNA has four splice variants, including the following gene and protein sequence:

PROTEIN (SEQ ID NO. 8)

```
  1 mllggwllla fnaifllswa vapkglcprr ssvpmpgvqa vaatamivgl lifpiglasp
 61 fikevceass myyggkcrlg wgymtailna vlasllpiis wphttkvqgr tiifssater
121 iifvpemnk
```

cDNA sequence (SEQ ID NO. 9)

```
  1 ctttgcctgg aaggtctcag ctgtgatgct cctcggaggc tggctcctgt tggccttcaa
 61 tgcaattttc ctcctgtctt gggctgtggc ccccaaaggg ctgtgcccaa ggagaagcag
121 tgttccaatg ccaggggtgc aggcagtggc agctactgcc atgattgtgg gtctgctgat
181 tttcccaatc ggccttgcct ccccattcat caaggaagtg tgcgaagcct cctccatgta
241 ttatggtggg aagtgccggc tgggttgggg ttacatgact gctatcctca atgcagtcct
301 ggccagcctc ctgcccatca tcagctggcc ccacacaacc aaggtccaag ggaggaccat
361 catcttctcc agtgccaccg agagaatcat ctttgtgcca gaaatgaaca aataaaaatc
421 tcctgggagt agcacaaagg gcacactcca gagttttatg aaatcatcat gtagccaact
481 tcaaatccca tctctgctcc ttcttgc
```

[00440] Using TMEM211 as the capture antibody and detection antibodies CD9, CD63, CD81, the following performance was obtained for detection of CRC:

Table 9: CRC Detection using TMEM211

True Positive	59
True Negative	58
False Positive	11
False Negative	0
Total	128
Sensitivity	**Specificity**
100.00%	84.06%

Example 25: MicroRNA Assessment in Colorectal Vesicles

[00441] Vesicles are obtained from the blood samples from colorectal patients using methodology as in Example 24. The vesicles are contacted with beads that are conjugated to antibodies to the vesicle surface antigens in **Table 8**. The vesicles captured by the beads are contacted with PE labeled antibodies to the tetraspanins CD9, CD63 and CD81. Flow cytometry is used to sort the beads having captured and labeled vesicles bound thereto as in Example 13. RNA is obtained from the sorted vesicles. The levels of the microRNA vesicle payload in **Table 7** are assessed for each sample. The MFI of the captured vesicles are used together with the concentration of the microRNA to construct a biosignature for colorectal cancer.

[00442] The same approach is used to construct a biosignature for a gastrointestinal cancer, gastric cancer, hepatocellular carcinoma, liver cancer, gastrointestinal stromal tumor (GIST), esophageal cancer, pancreatic

cancer, colorectal cancer, Barrett's Esophagus, hyperplastic polyp, adenoma, inflammatory bowel disease (IBD) or cirrhosis. In such cases, the cohort of samples used to derive the biosignature is selected based on the comparison of interest, e.g., disease 1 versus disease 2 or disease versus normal.

[00443] While preferred embodiments of the present invention have been shown and described herein, it will be obvious to those skilled in the art that such embodiments are provided by way of example only. Numerous variations, changes, and substitutions will now occur to those skilled in the art without departing from the invention. It should be understood that various alternatives to the embodiments of the invention described herein may be employed in practicing the invention. It is intended that the following claims define the scope of the invention and that methods and structures within the scope of these claims and their equivalents be covered thereby.

CLAIMS

WHAT IS CLAIMED IS:

1. A method for characterizing a phenotype in a sample, comprising:

   i) identifying a biosignature in the sample, wherein the biosignature comprises a level of one or more microRNA selected from the group consisting of miR-548c-5p, miR-362-3p, miR-422a, miR-597, miR-429, miR-200a, and miR-200b; and

   ii) comparing the biosignature to a reference, wherein a difference in the level of the one or more microRNA in the sample compared to the reference indicate the phenotype, thereby characterizing the phenotype.

2. The method of claim **1**, wherein the one or more microRNA consist of mature microRNA.

3. The method of claim **1**, wherein the sample comprises a biological sample from a cell line.

4. The method of claim **2**, wherein the phenotype comprises cancer.

5. The method of claim **1**, wherein the sample comprises a biological sample from a subject.

6. The method of claim **5**, wherein the phenotype comprises cancer.

7. The method of claim **3** or **5**, wherein the identifying is performed in vitro.

8. The method of claim **6**, wherein the characterizing comprises a diagnosis, prognosis, determination of drug efficacy, monitoring the status of the subject's response or resistance to a treatment or selection of a treatment for the cancer.

9. The method of claim **1**, wherein the characterizing comprises diagnosing and the phenotype comprises a cancer.

10. The method of claim **6**, wherein the subject is non-responsive to a treatment being administered to the subject.

11. The method of claim **10**, wherein the treatment comprises administering a cancer therapeutic.

12. The method of claim **5**, wherein the sample comprises a bodily fluid.

13. The method of claim **12**, wherein the bodily fluid is peripheral blood, sera, plasma, ascites, urine, cerebrospinal fluid (CSF), sputum, saliva, bone marrow, synovial fluid, aqueous humor, amniotic fluid, cerumen, breast milk, broncheoalveolar lavage fluid, semen, prostatic fluid, cowper's fluid or pre-ejaculatory fluid, female ejaculate, sweat, fecal matter, hair, tears, cyst fluid, pleural and peritoneal fluid, pericardial fluid, lymph, chyme, chyle, bile, interstitial fluid, menses, pus, sebum, vomit, vaginal secretions, mucosal secretion, stool, stool water, pancreatic juice, lavage fluids from sinus cavities, bronchopulmonary aspirates, blastocyl cavity fluid, or umbilical cord blood.

14. The method of claim **12**, wherein the bodily fluid comprises peripheral blood, sera, plasma, saliva or stool.

15. The method of claim **4** or **6**, wherein the cancer comprises an acute lymphoblastic leukemia; acute myeloid leukemia; adrenocortical carcinoma; AIDS-related cancers; AIDS-related lymphoma; anal cancer; appendix cancer; astrocytomas; atypical teratoid/rhabdoid tumor; basal cell carcinoma; bladder cancer; brain stem glioma; brain tumor (including brain stem glioma, central nervous system atypical teratoid/rhabdoid tumor, central nervous system embryonal tumors, astrocytomas, craniopharyngioma, ependymoblastoma, ependymoma, medulloblastoma, medulloepithelioma, pineal parenchymal tumors of intermediate differentiation, supratentorial primitive neuroectodermal tumors and pineoblastoma); breast cancer; bronchial tumors; Burkitt lymphoma; cancer of unknown primary site; carcinoid tumor; carcinoma of unknown primary site; central nervous system atypical teratoid/rhabdoid tumor; central nervous system embryonal tumors; cervical cancer; childhood cancers; chordoma; chronic lymphocytic leukemia; chronic myelogenous leukemia; chronic myeloproliferative disorders; colon cancer; colorectal cancer; craniopharyngioma; cutaneous T-cell lymphoma; endocrine pancreas islet cell tumors; endometrial cancer; ependymoblastoma; ependymoma; esophageal cancer; esthesioneuroblastoma; Ewing sarcoma; extracranial germ cell tumor; extragonadal germ cell tumor; extrahepatic bile duct cancer; gallbladder cancer; gastric (stomach) cancer; gastrointestinal carcinoid tumor; gastrointestinal stromal cell tumor; gastrointestinal stromal tumor (GIST); gestational trophoblastic tumor; glioma; hairy cell leukemia; head and neck cancer; heart cancer; Hodgkin lymphoma; hypopharyngeal cancer; intraocular melanoma; islet cell tumors; Kaposi sarcoma; kidney cancer; Langerhans cell histiocytosis; laryngeal cancer; lip cancer; liver cancer; malignant fibrous histiocytoma bone cancer; medulloblastoma; medulloepithelioma; melanoma; Merkel cell carcinoma; Merkel cell skin carcinoma; mesothelioma; metastatic squamous neck cancer with occult primary; mouth cancer; multiple endocrine neoplasia syndromes; multiple myeloma; multiple myeloma/plasma cell neoplasm; mycosis fungoides; myelodysplastic syndromes; myeloproliferative neoplasms; nasal cavity cancer; nasopharyngeal cancer; neuroblastoma; Non-Hodgkin lymphoma; nonmelanoma skin cancer; non-small cell lung cancer; oral cancer; oral cavity cancer; oropharyngeal cancer; osteosarcoma; other brain and spinal cord tumors; ovarian cancer; ovarian epithelial cancer; ovarian germ cell tumor; ovarian low malignant potential tumor; pancreatic cancer; papillomatosis; paranasal sinus cancer; parathyroid cancer; pelvic cancer; penile cancer; pharyngeal cancer; pineal parenchymal tumors of intermediate differentiation; pineoblastoma; pituitary tumor; plasma cell neoplasm/multiple myeloma; pleuropulmonary blastoma; primary central nervous system (CNS) lymphoma; primary hepatocellular liver cancer; prostate cancer; rectal cancer; renal cancer; renal cell (kidney) cancer; renal cell cancer; respiratory tract cancer; retinoblastoma; rhabdomyosarcoma; salivary gland cancer; Sézary syndrome; small cell lung cancer; small intestine cancer; soft tissue sarcoma; squamous cell carcinoma; squamous neck cancer; stomach (gastric) cancer; supratentorial primitive neuroectodermal tumors; T-cell lymphoma; testicular cancer; throat cancer; thymic carcinoma; thymoma; thyroid cancer; transitional cell cancer; transitional cell cancer of the renal pelvis and ureter; trophoblastic tumor; ureter cancer; urethral cancer; uterine cancer; uterine sarcoma; vaginal cancer; vulvar cancer; Waldenström macroglobulinemia; or Wilm's tumor.

16. The method of claim **4** or **6**, wherein the cancer comprises a gastrointestinal cancer, gastric cancer, hepatocellular carcinoma, liver cancer, gastrointestinal stromal tumor (GIST), esophageal cancer, pancreatic cancer or colorectal cancer.

17. The method of claim **4** or **6**, wherein the cancer comprises colorectal cancer.

18. The method of claim **17**, wherein the colorectal cancer is Dukes B, Dukes C or Dukes D.

19. The method of claim **5**, wherein the reference comprises a different individual or group of individuals as compared to the subject.

20. The method of claim **5**, wherein the reference comprises samples obtained from the subject over a time course.

21. The method of claim **1**, wherein the sample comprises a vesicle.

22. The method of claim **21**, wherein the vesicle has a diameter of about 30 nm to about 800 nm.

23. The method of claim **21**, wherein the vesicle has a diameter of about 30 nm to about 200 nm.

24. The method of claim **21**, wherein the vesicle is isolated prior to step (i) by size exclusion chromatography, density gradient centrifugation, differential centrifugation, flow cytometry, high pressure liquid chromatography, flow pressure liquid chromatography, membrane ultrafiltration, affinity capture, microfluidic device, or combinations thereof.

25. The method of claim **21**, wherein the vesicle is a cell-of-origin specific vesicle.

26. The method of claim **25**, wherein the cell-of-origin is a tumor or cancer cell.

27. The method of claim **25**, wherein the cell-of-origin is a lung, pancreas, stomach, intestine, bladder, kidney, ovary, testis, skin, colorectal, breast, prostate, brain, esophagus, liver, placenta, or fetal cell.

28. The method of claim **25**, wherein the cell-of-origin is a stomach, intestine, colorectal, esophagus, or liver cell.

29. The method of any of claims **21-27**, wherein the one or more microRNA is present in the vesicle.

30. The method of any of claims **21**, wherein identifying the biosignature comprises:

    i) isolating the vesicle by:

        a. filtering the biological sample with a filter that retains the vesicle;

        b. capturing the retained vesicle with a capture binding agent;

        c. contacting the captured vesicle with a labeled binding agent; and

        d. isolating the labeled vesicle;

    ii) extracting nucleic acid from the isolated vesicle; and

    iii) detecting the one or more microRNA in the extracted nucleic acid.

31. The method of claim **30**, wherein the capture binding agent and/or labeled binding agent comprises an antibody, antibody fragment or an aptamer.

32. The method of claim **30**, wherein the capture binding agent is tethered to a substrate.

33. The method of claim **32**, wherein the substrate comprises an array, well, or particle.

34. The method of claim **32**, wherein the substrate comprises a labeled particle.

35. The method of claim **34**, wherein the label is a fluorescent label.

36. The method of claim **30**, wherein the labeled binding agent is fluorescently labeled.

37. The method of claim **30**, wherein the labeled binding agent is to one or more of CD9, CD63 and CD81.

38. The method of claim **30**, wherein the labeled vesicle is isolated using flow cytometry.

39. The method of claim **1**, wherein the biosignature further comprises a level or presence of one or more general vesicle biomarker and a level or presence of one or more cell-of-origin biomarker.

40. The method of claim **1**, wherein the biosignature further comprises a level or presence of one or more general vesicle biomarkers, and a level or presence of one or more disease specific biomarkers.

41. The method of claim **1**, wherein the biosignature further comprises a level or presence of one or more general vesicle biomarkers, a level or presence of one or more cell-of-origin biomarkers, and a level or presence of one or more disease specific biomarkers.

42. The method of claim **1**, wherein the biosignature further comprises a level or presence of one or more cell-of-origin biomarkers, and a level or presence of one or more disease specific biomarkers.

43. The method of claim **39**, **40**, or **41**, wherein the one or more general vesicle biomarker comprises one or more of CD9, CD63, CD81, CD37, CD53, CD82, or Rab-5b.

44. The method of claim **1**, wherein the biosignature further comprises a level or presence of one or more additional biomarker selected from the group consisting of DR3, STEAP, epha2, TMEM211, unc93A, A33, CD24, NGAL, EpCam, MUC17, TROP2, and TETS.

45. The method of claim **1**, wherein the biosignature further comprises a level or presence of one or more additional biomarker, wherein the one or more additional biomarker comprises TMEM211.

46. The method of claim **1**, wherein the biosignature further comprises a level or presence of one or more additional biomarker, wherein the one or more additional biomarker comprises CD24.

47. The method of claim **44**, **45** or **46**, wherein the one or more additional biomarker comprises an mRNA, a circulating biomarker, or a protein.

48. The method of claim **44**, **45** or **46**, wherein the one or more additional biomarker is associated with a vesicle.

49. The method of claim **48**, wherein the one or more additional biomarker is a surface antigen of the vesicle.

50. The method of claim **48**, wherein the one or more additional biomarker is payload within the vesicle.

51. The method of claim **1**, wherein the biosignature further comprises an expression level, presence, absence, mutation, copy number variation, truncation, duplication, insertion, modification, sequence variation, or molecular association of one or more biomarker.

52. The method of claim **51**, wherein the one or more biomarker comprises a nucleic acid, peptide, protein, lipid, antigen, carbohydrate, proteoglycan, vesicle, or a combination thereof.

53. The method of claim **51**, wherein the one or more biomarker is detected using microarray analysis, PCR, hybridization with allele-specific probes, enzymatic mutation detection, ligation chain reaction (LCR), oligonucleotide ligation assay (OLA), flow-cytometric heteroduplex analysis, chemical cleavage of mismatches, mass spectrometry, nucleic acid sequencing, single strand conformation polymorphism (SSCP), denaturing gradient gel electrophoresis (DGGE), temperature gradient gel electrophoresis (TGGE), restriction fragment polymorphisms, serial analysis of gene expression (SAGE), image cytometry, qRT-PCR, real-time PCR, PCR, flow cytometry, mass spectrometry, or a combination thereof.

54. The method of claim **51**, wherein the one or more biomarker is detected using a binding agent.

55. The method of claim **54**, wherein the binding agent is selected from the group consisting of: DNA, RNA, monoclonal antibodies, polyclonal antibodies, Fabs, Fab', single chain antibodies, synthetic antibodies, DNA aptamers, RNA aptamers, lectin, peptoids, zDNA, peptide nucleic acids (PNAs), locked nucleic acids (LNAs), synthetic occurring chemical compounds, naturally occurring chemical compounds, dendrimers, and combinations thereof.

56. A use of an agent for determining the level of at least one microRNA in preparation of a composition for carrying out a method according to any of claims **1** to **55**.

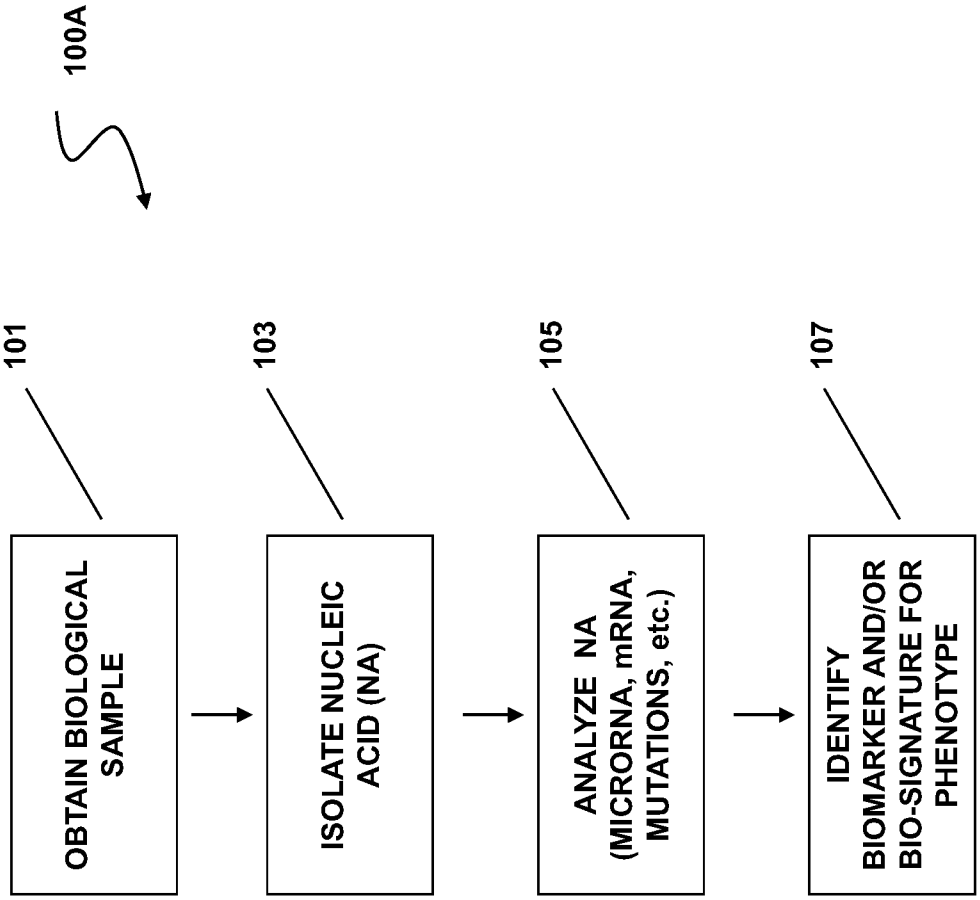

100A
101
OBTAIN BIOLOGICAL SAMPLE
103
ISOLATE NUCLEIC ACID (NA)
105
ANALYZE NA (MICRORNA, mRNA, MUTATIONS, etc.)
107
IDENTIFY BIOMARKER AND/OR BIO-SIGNATURE FOR PHENOTYPE

FIG. 1A

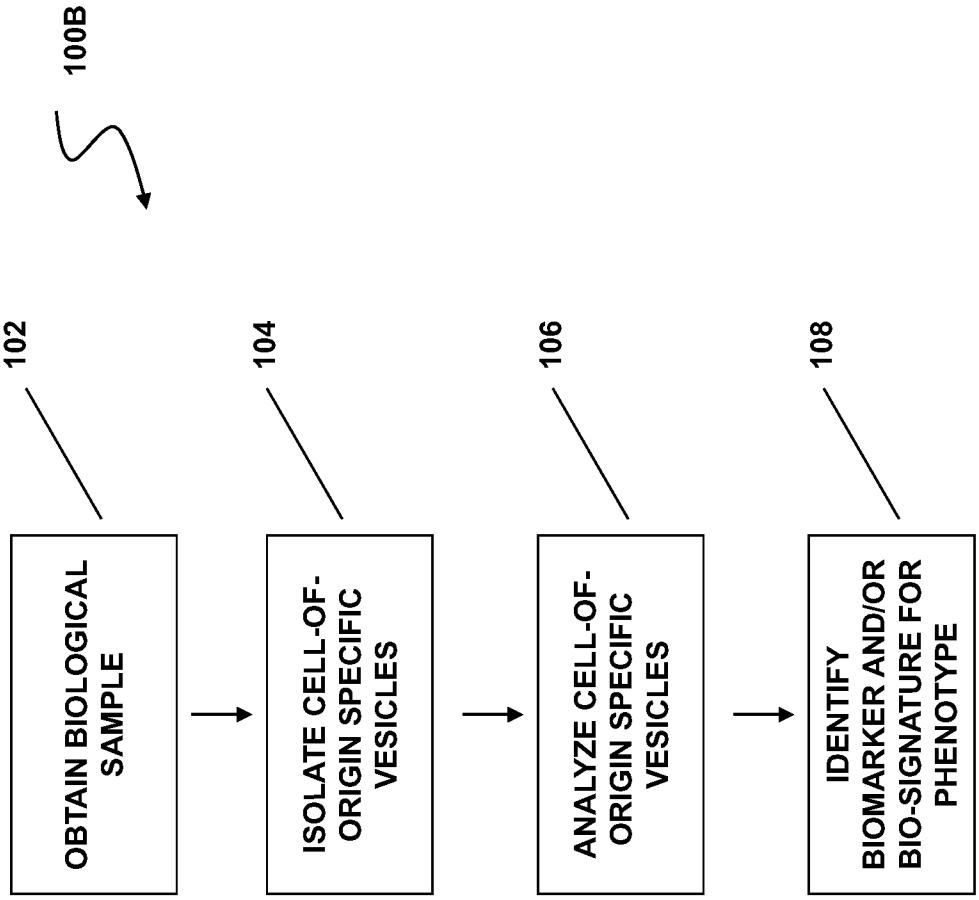

100B
102
OBTAIN BIOLOGICAL SAMPLE
104
ISOLATE CELL-OF-ORIGIN SPECIFIC VESICLES
106
ANALYZE CELL-OF-ORIGIN SPECIFIC VESICLES
108
IDENTIFY BIOMARKER AND/OR BIO-SIGNATURE FOR PHENOTYPE

FIG. 1B

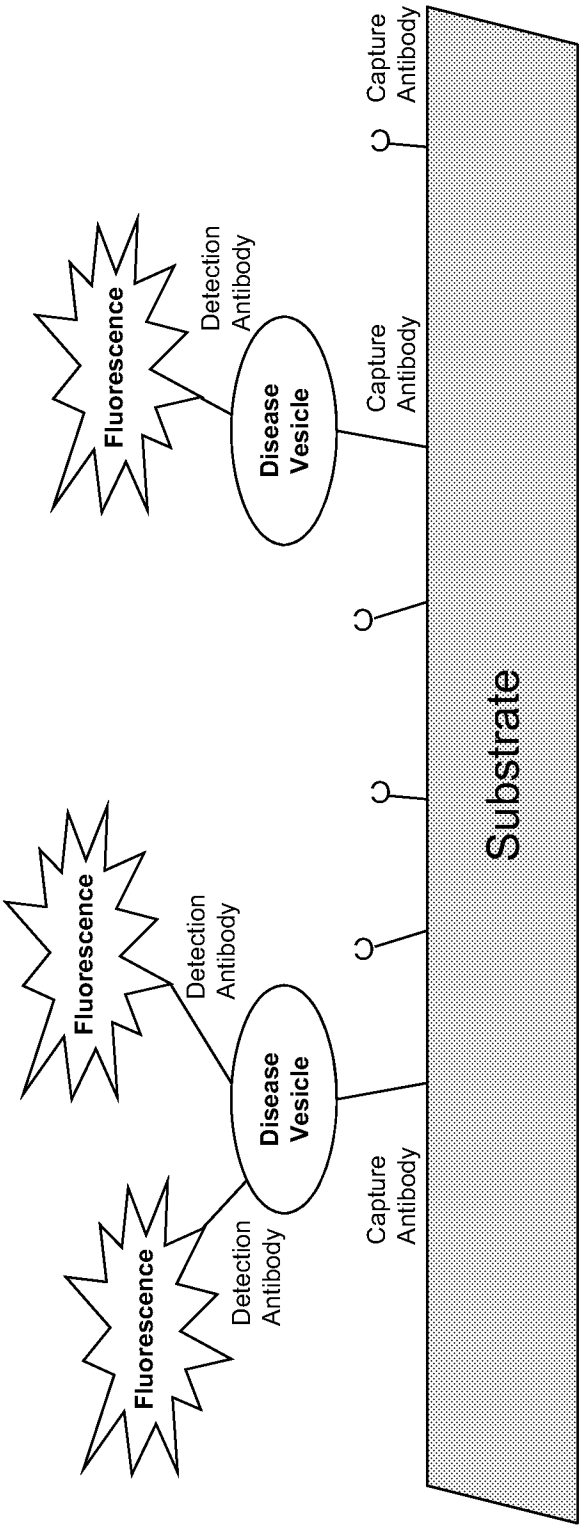

Substrate
Capture Antibody
Capture Antibody
Capture Antibody
Disease Vesicle
Disease Vesicle
Disease Vesicle
Detection Antibody
Detection Antibody
Detection Antibody
Fluorescence
Fluorescence
Fluorescence

FIG. 2A

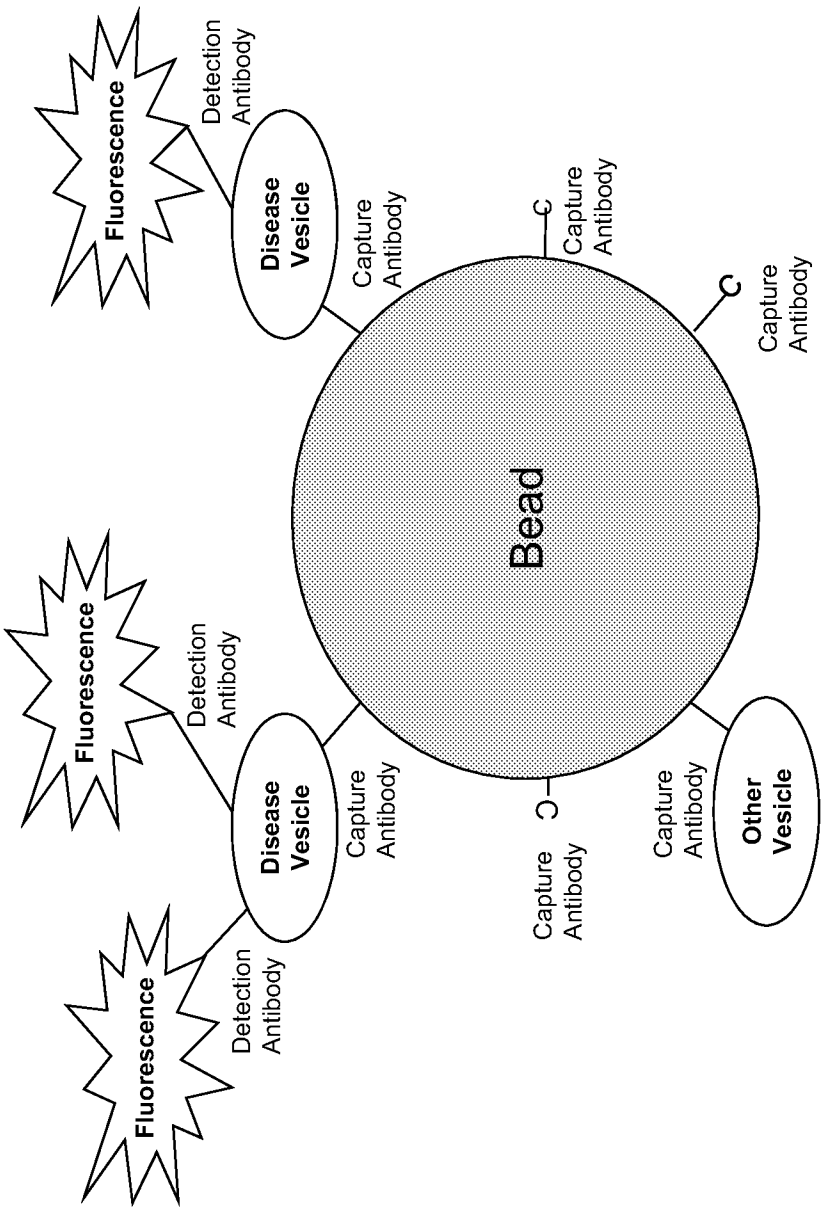

Fluorescence
Detection Antibody
Disease Vesicle
Capture Antibody
Capture Antibody
Capture Antibody
Bead
Capture Antibody
Capture Antibody
Other Vesicle
Fluorescence
Detection Antibody
Disease Vesicle
Capture Antibody
Fluorescence
Detection Antibody

FIG. 2B

Screening Scheme

5	X	20	=	100
Detection Antibodies		Capture Antibodies		Combinations Screened

CD63	CD9	Rab
CD9	PSCA	IgG
CD81	TNFR	CD81
B7H3	CD63 2X	STEAP
EpCam	B7H3	PCSA
	Rab IgG	PSMA
	MFG-E8	5T4
	EpCam 2X	CD24
	CD63	TMEM211

General vesicle biomarker antibodies: CD9, CD63, CD81
Cell of Origin biomarker antibodies: PSCA, MFG-E8, Rab, STEAP, PCSA, PSMA, 5T4, TMEM211
Cancer biomarker antibodies: EpCam, B7H3, CD24
Control antibodies: Rab IgG, IgG

FIG. 2C

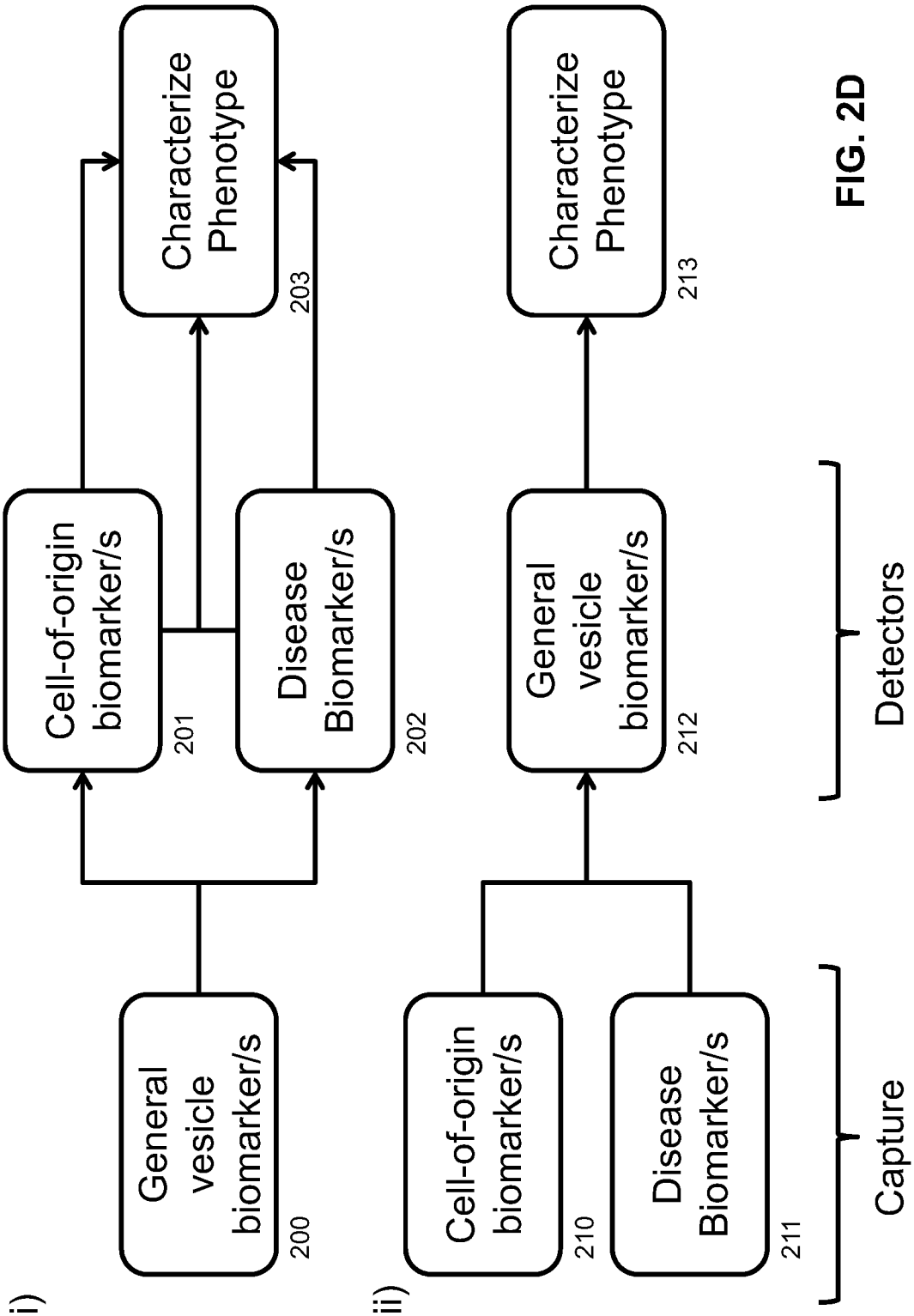

i)
General vesicle biomarker/s
200
Cell-of-origin biomarker/s
201
Disease Biomarker/s
202
Characterize Phenotype
203
ii)
Cell-of-origin biomarker/s
210
Disease Biomarker/s
211
General vesicle biomarker/s
212
Characterize Phenotype
213
Capture
Detectors

FIG. 2D

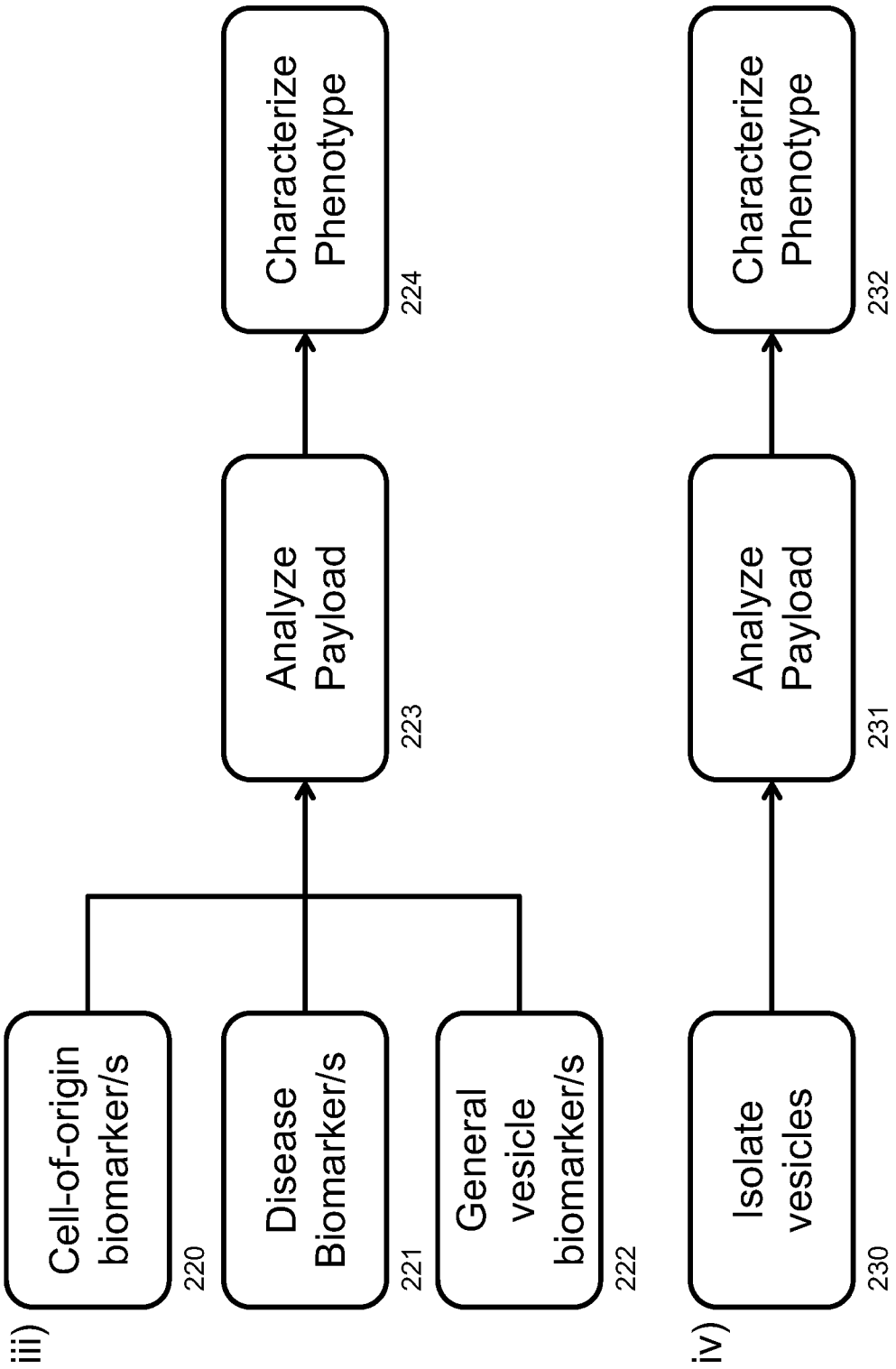

iii)
Cell-of-origin biomarker/s
220
Disease Biomarker/s
221
General vesicle biomarker/s
222
Analyze Payload
223
Characterize Phenotype
224
iv)
Isolate vesicles
230
Analyze Payload
231
Characterize Phenotype
232

FIG. 2E

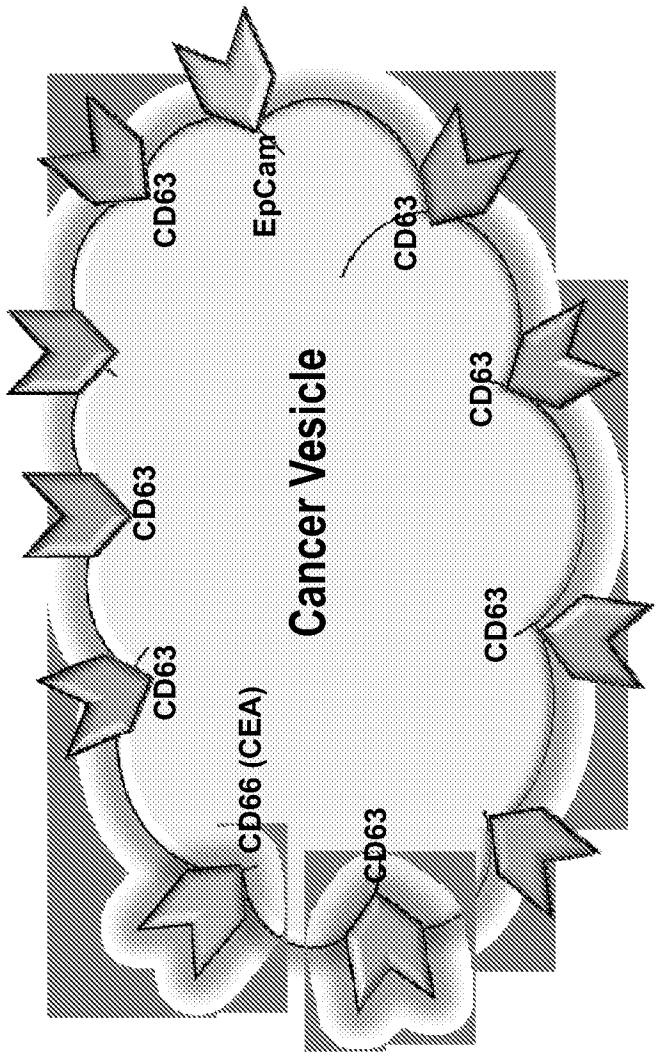

CD63
EpCam
CD63
CD63
CD63
Cancer Vesicle
CD63
CD63
CD66 (CEA)
CD63

FIG. 3

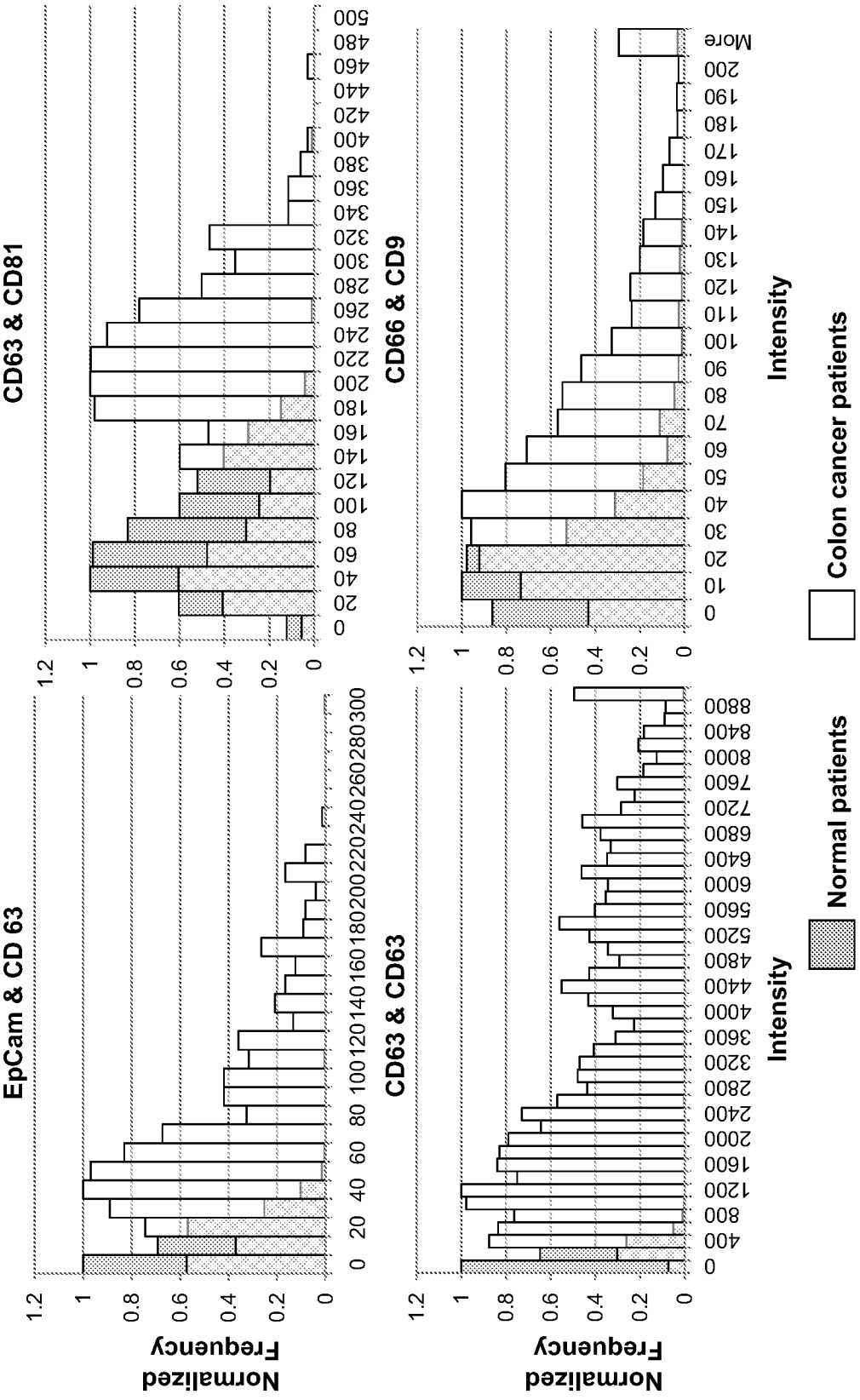

EpCam & CD 63
CD63 & CD81
CD63 & CD63
CD66 & CD9
Normalized Frequency
Intensity
Normal patients
Colon cancer patients

FIG. 4A

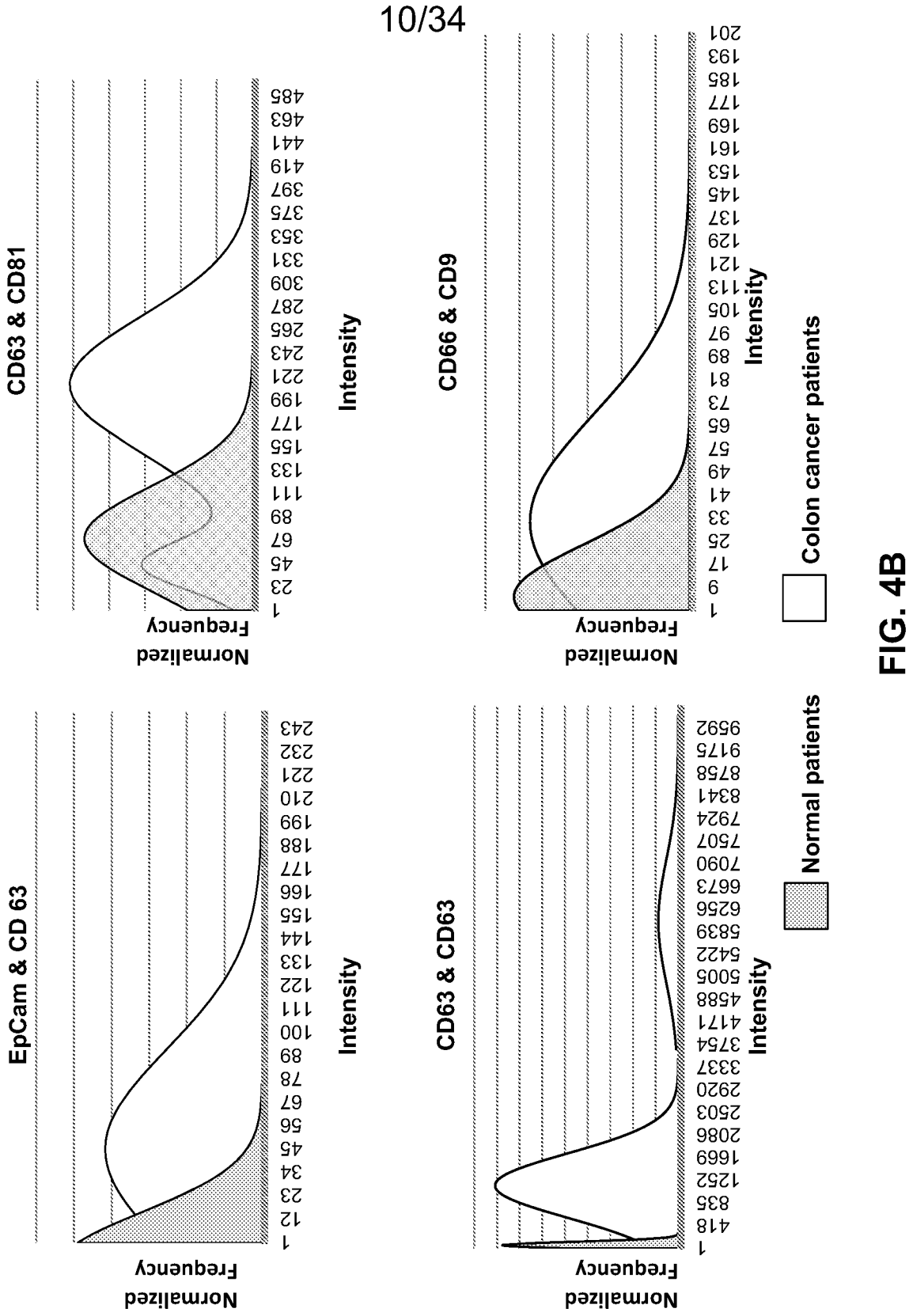

CD63 & CD81
CD66 & CD9
EpCam & CD 63
CD63 & CD63
Normalized Frequency
Intensity
Colon cancer patients
Normal patients

FIG. 4B

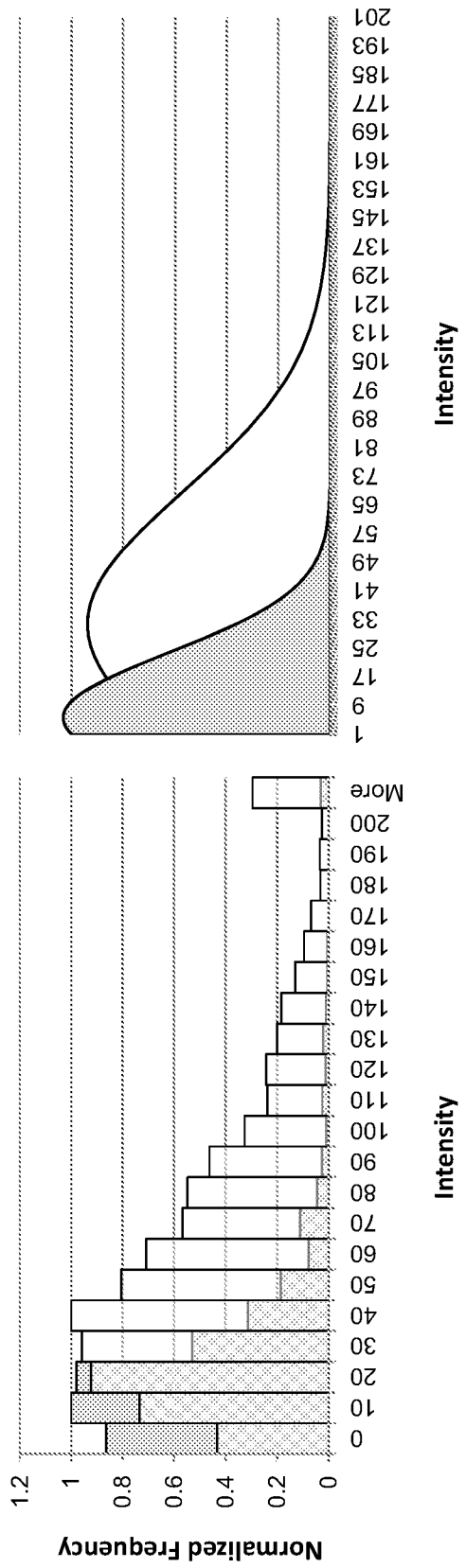

Normalized Frequency
1.2
1
0.8
0.6
0.4
0.2
0
Intensity
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
200
More
Intensity
1
9
17
25
33
41
49
57
65
73
81
89
97
105
113
121
129
137
145
153
161
169
177
185
193
201
Normal patients
Colon cancer patients

FIG. 4C

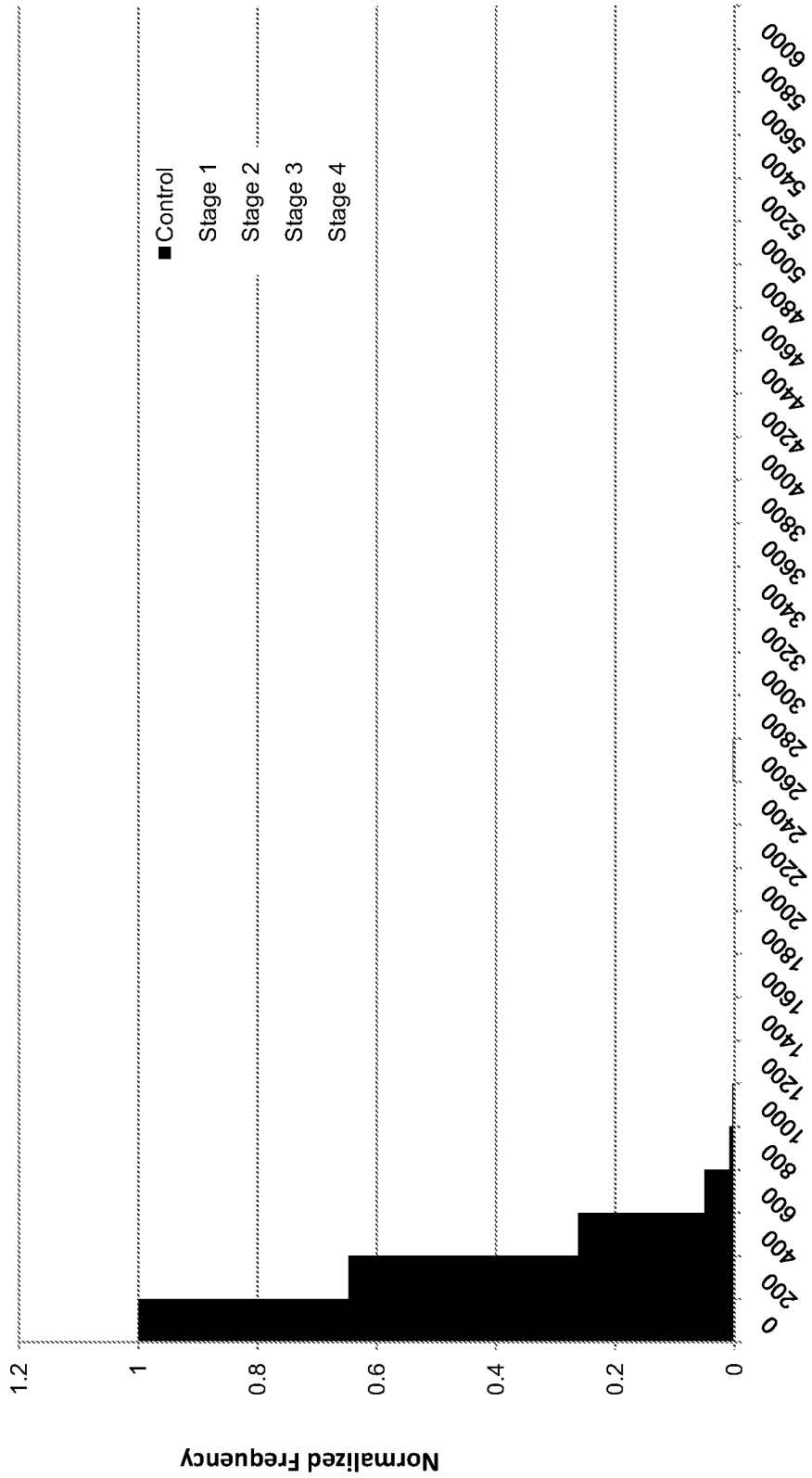

Normalized Frequency
1.2
1
0.8
0.6
0.4
0.2
0
Control
Stage 1
Stage 2
Stage 3
Stage 4
0
200
400
600
800
1000
1200
1400
1600
1800
2000
2200
2400
2600
2800
3000
3200
3400
3600
3800
4000
4200
4400
4600
4800
5000
5200
5400
5600
5800
6000
Intensity

FIG. 5A

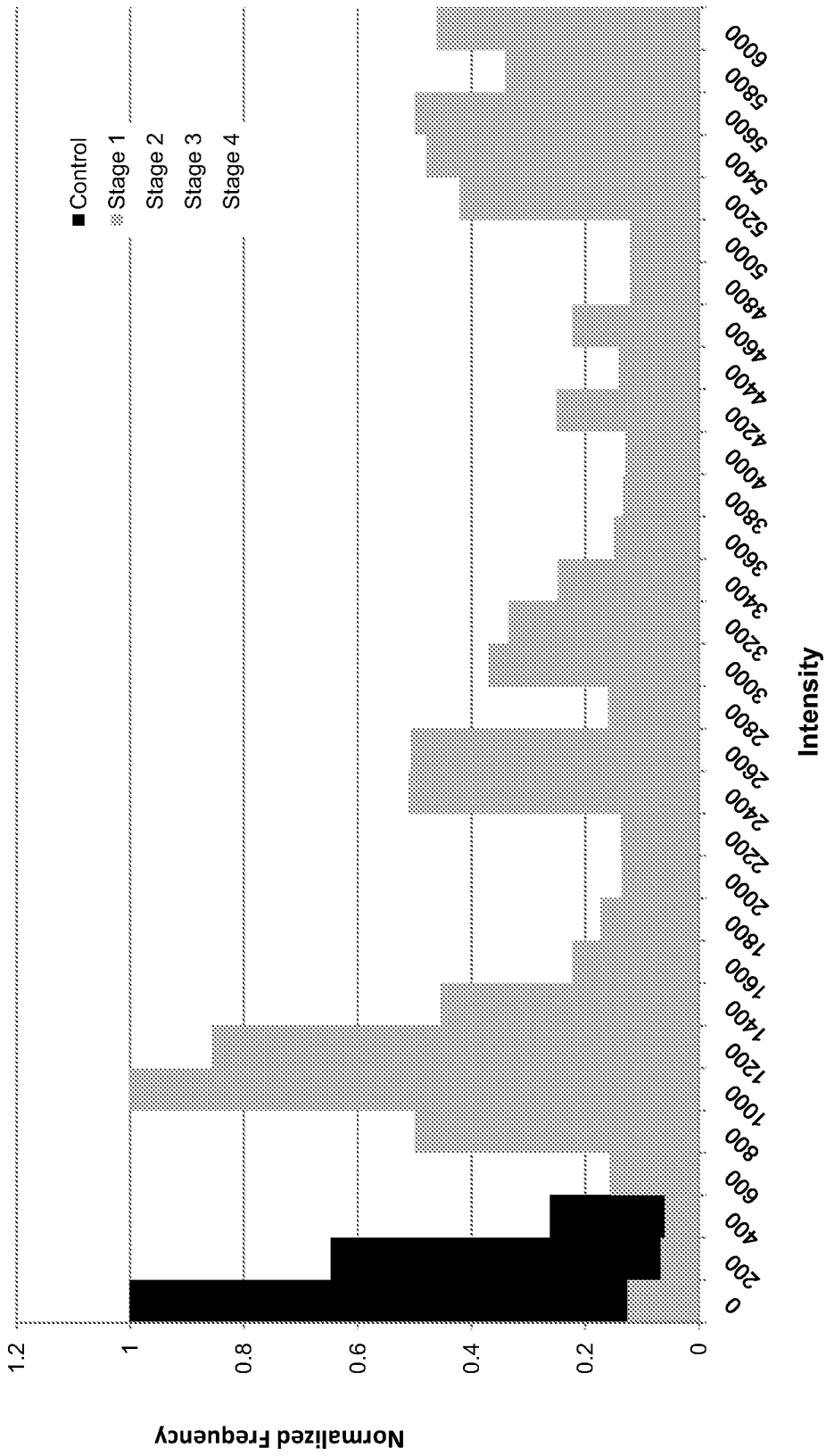

Control
Stage 1
Stage 2
Stage 3
Stage 4
Normalized Frequency
1.2
1
0.8
0.6
0.4
0.2
0
Intensity
0
200
400
600
800
1000
1200
1400
1600
1800
2000
2200
2400
2600
2800
3000
3200
3400
3600
3800
4000
4200
4400
4600
4800
5000
5200
5400
5600
5800
6000

FIG. 5B

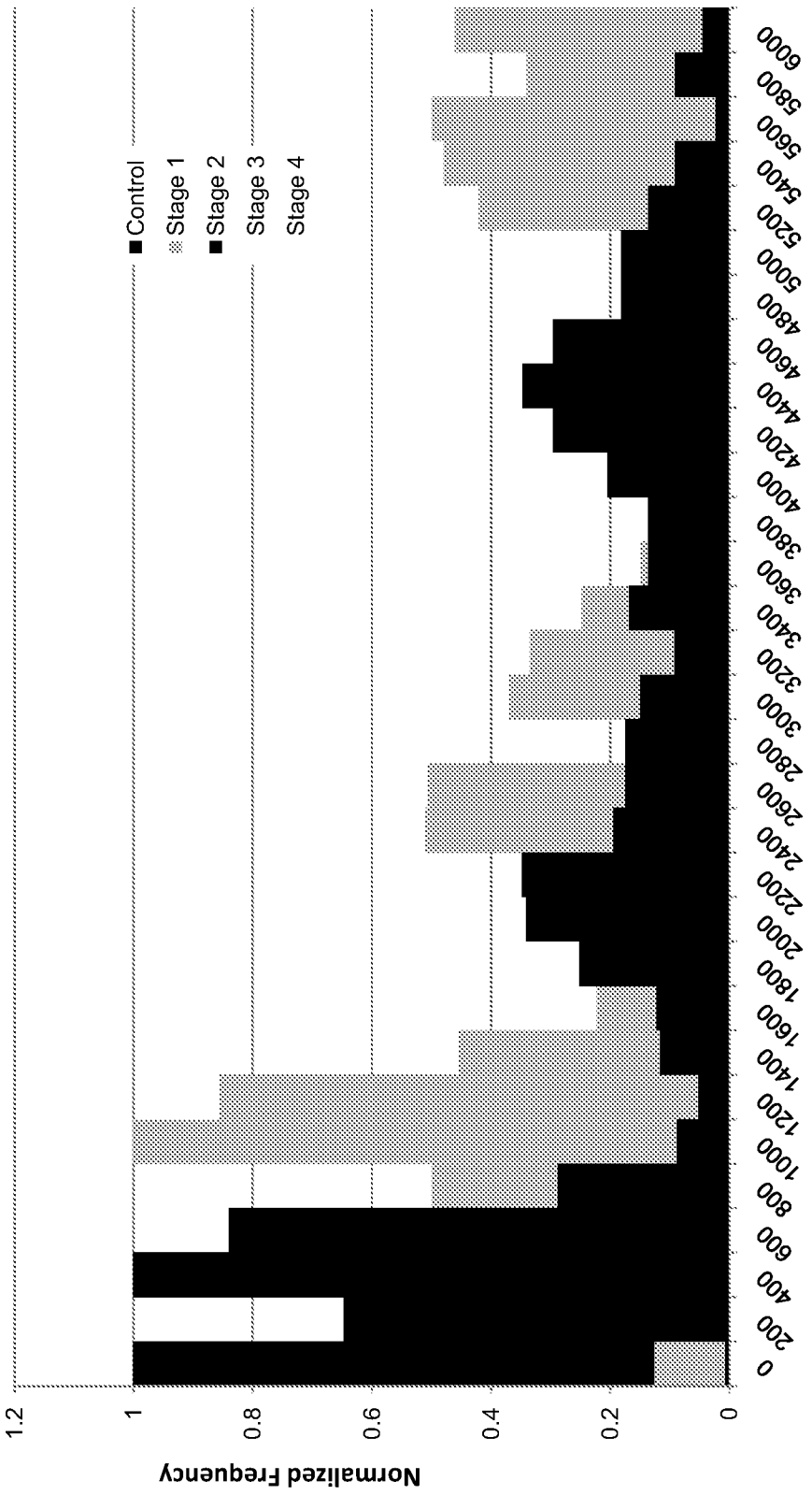

Normalized Frequency
1.2
1
0.8
0.6
0.4
0.2
0
Control
Stage 1
Stage 2
Stage 3
Stage 4
0
200
400
600
800
1000
1200
1400
1600
1800
2000
2200
2400
2600
2800
3000
3200
3400
3600
3800
4000
4200
4400
4600
4800
5000
5200
5400
5600
5800
6000
Intensity

FIG. 5C

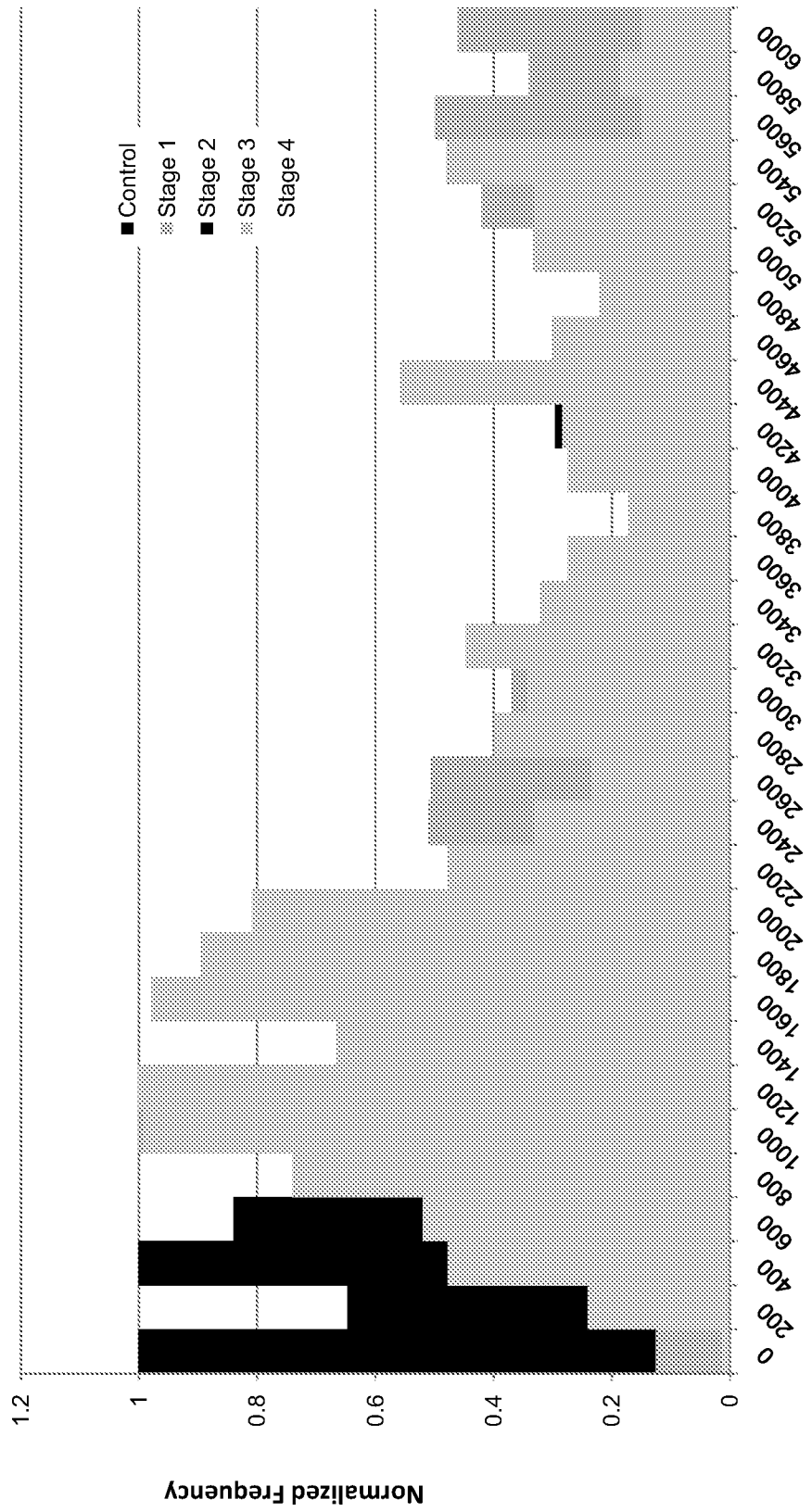

Control
Stage 1
Stage 2
Stage 3
Stage 4
Normalized Frequency
0
0.2
0.4
0.6
0.8
1
1.2
Intensity
0
200
400
600
800
1000
1200
1400
1600
1800
2000
2200
2400
2600
2800
3000
3200
3400
3600
3800
4000
4200
4400
4600
4800
5000
5200
5400
5600
5800
6000

FIG. 5D

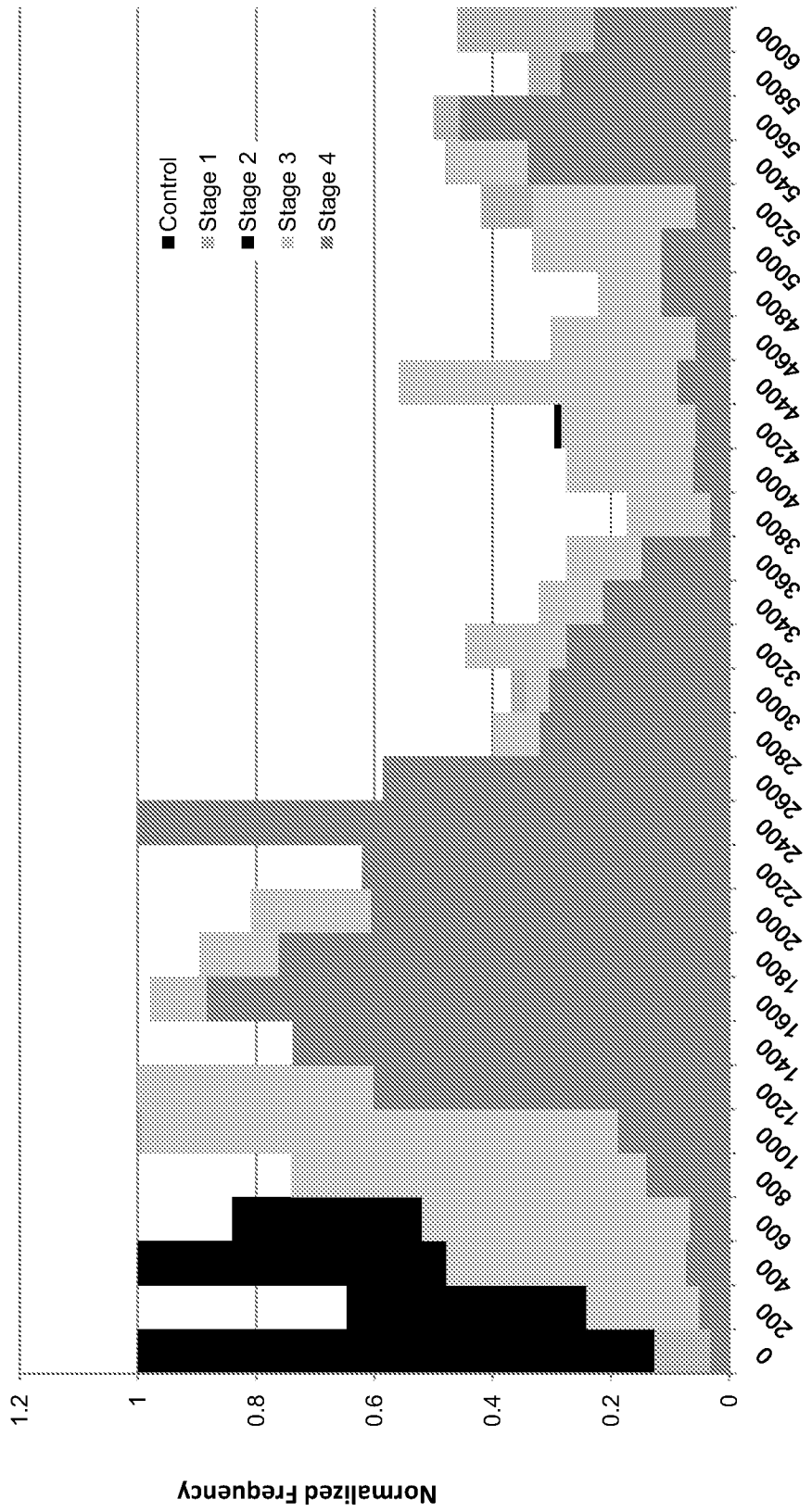

Normalized Frequency
1.2
1
0.8
0.6
0.4
0.2
0
Intensity
0
200
400
600
800
1000
1200
1400
1600
1800
2000
2200
2400
2600
2800
3000
3200
3400
3600
3800
4000
4200
4400
4600
4800
5000
5200
5400
5600
5800
6000
Control
Stage 1
Stage 2
Stage 3
Stage 4

FIG. 5E

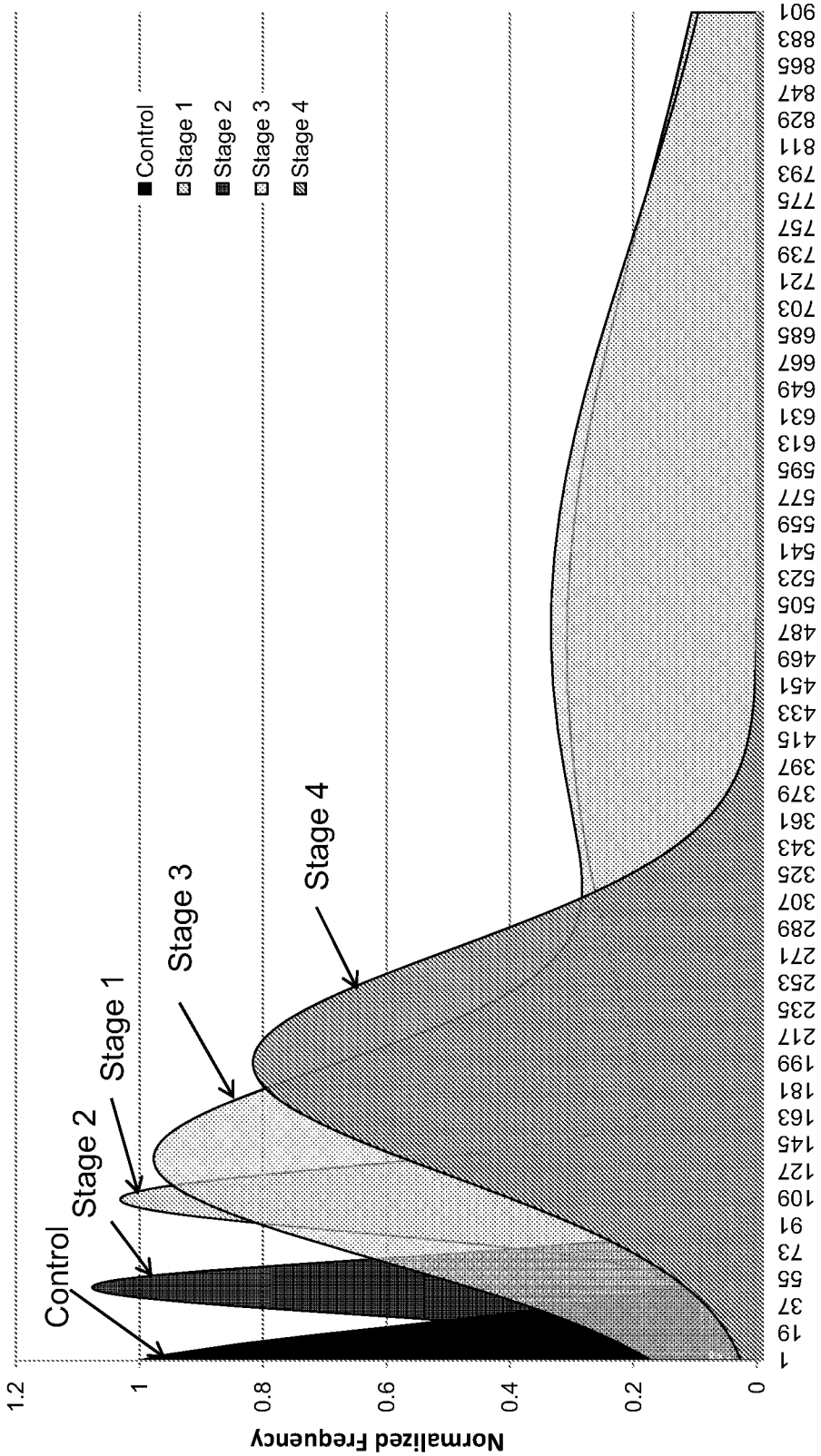

Control
Stage 1
Stage 2
Stage 3
Stage 4
Normalized Frequency
1.2
1
0.8
0.6
0.4
0.2
0
Intensity
1
19
37
55
73
91
109
127
145
163
181
199
217
235
253
271
289
307
325
343
361
379
397
415
433
451
469
487
505
523
541
559
577
595
613
631
649
667
685
703
721
739
757
775
793
811
829
847
865
883
901

FIG. 5F

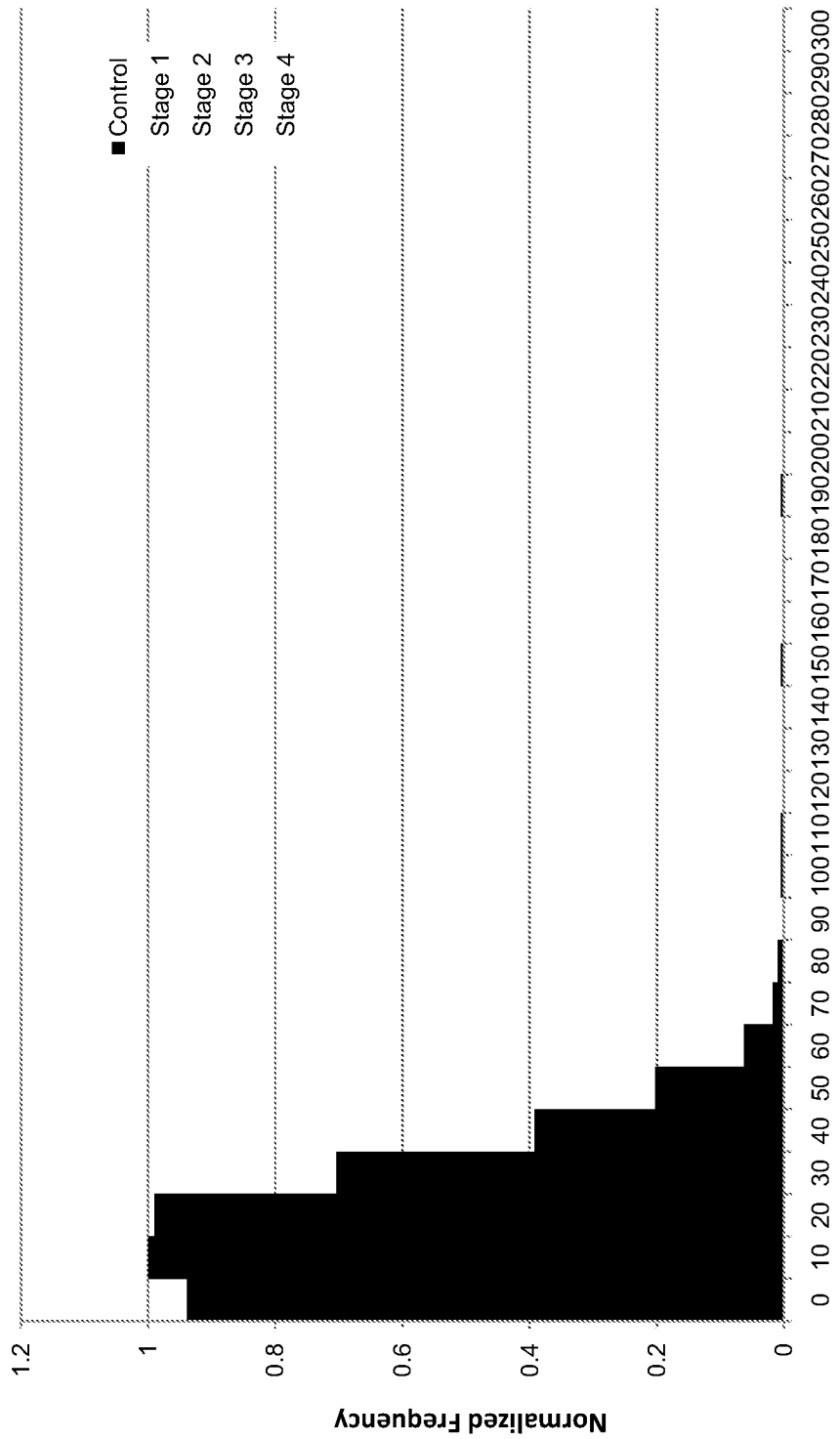

Control
Stage 1
Stage 2
Stage 3
Stage 4
Normalized Frequency
1.2
1
0.8
0.6
0.4
0.2
0
0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200 210 220 230 240 250 260 270 280 290 300
Intensity

FIG. 6A

FIG. 6B

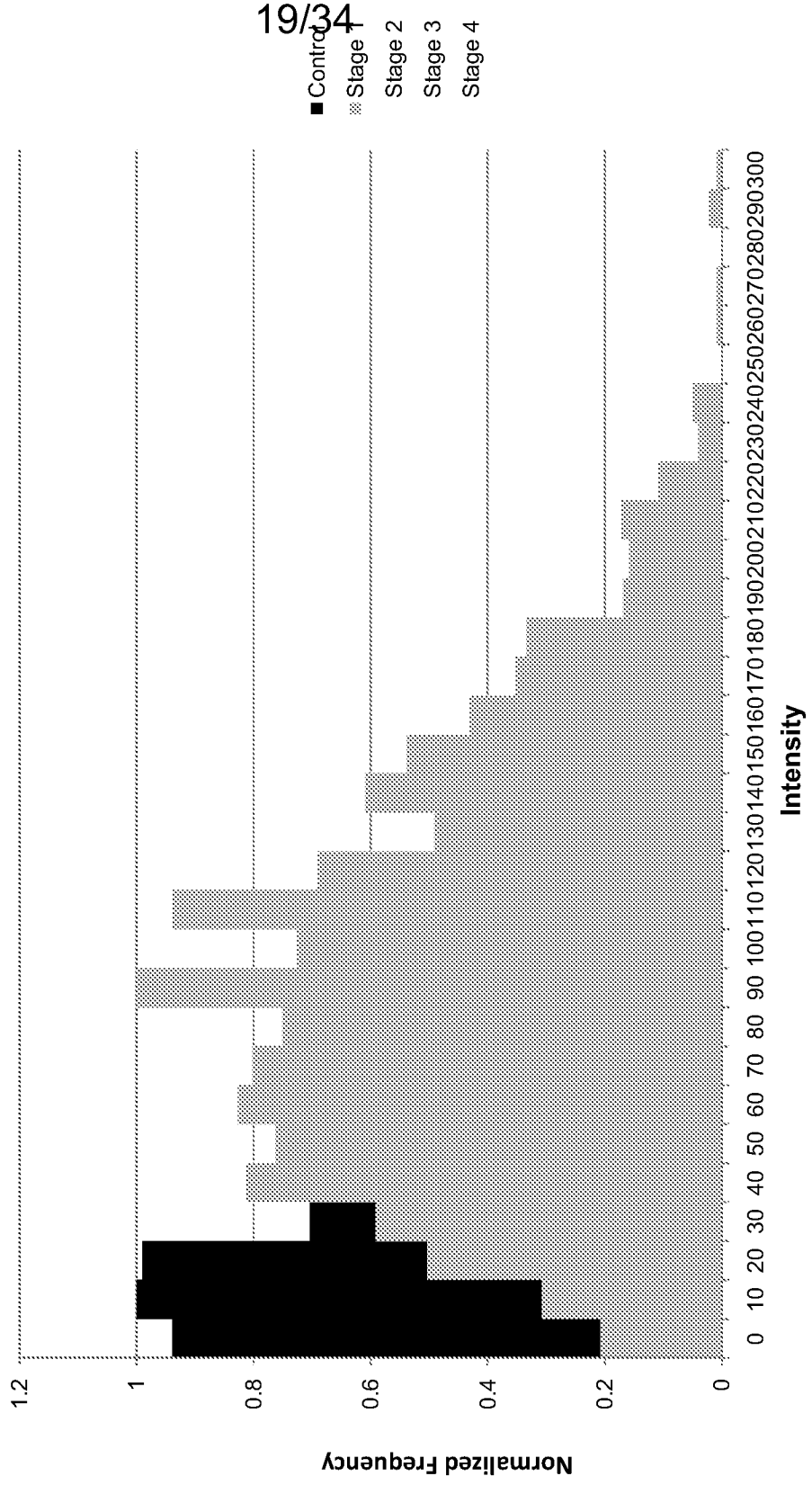
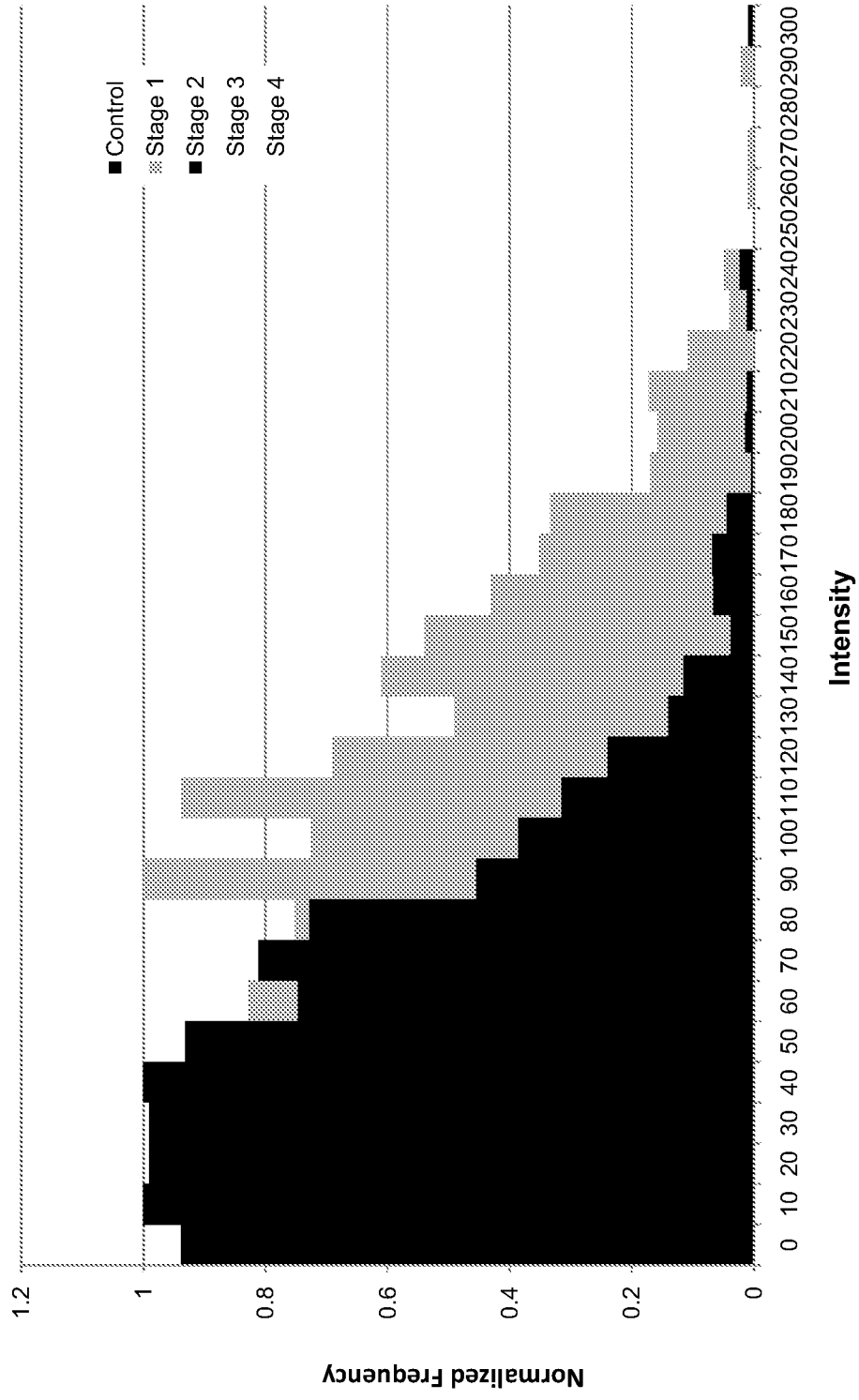

Control
Stage 1
Stage 2
Stage 3
Stage 4
Normalized Frequency
1.2
1
0.8
0.6
0.4
0.2
0
0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200 210 220 230 240 250 260 270 280 290 300
Intensity

FIG. 6C


Control
Stage 1
Stage 2
Stage 3
Stage 4
Normalized Frequency
1.2
1
0.8
0.6
0.4
0.2
0
0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200 210 220 230 240 250 260 270 280 290 300
Intensity

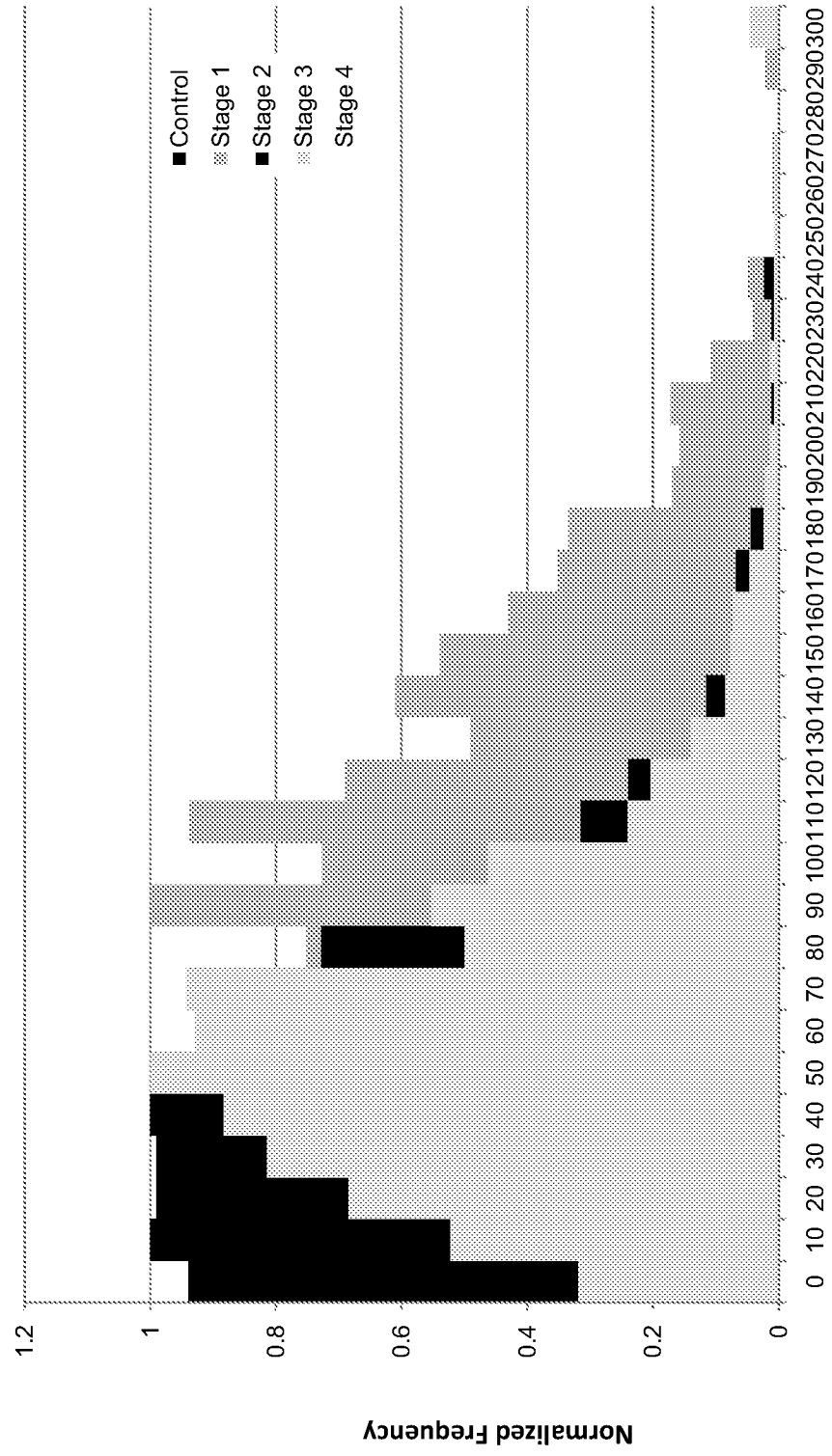

Control
Stage 1
Stage 2
Stage 3
Stage 4
Normalized Frequency
1.2
1
0.8
0.6
0.4
0.2
0
0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200 210 220 230 240 250 260 270 280 290 300
Intensity

FIG. 6D

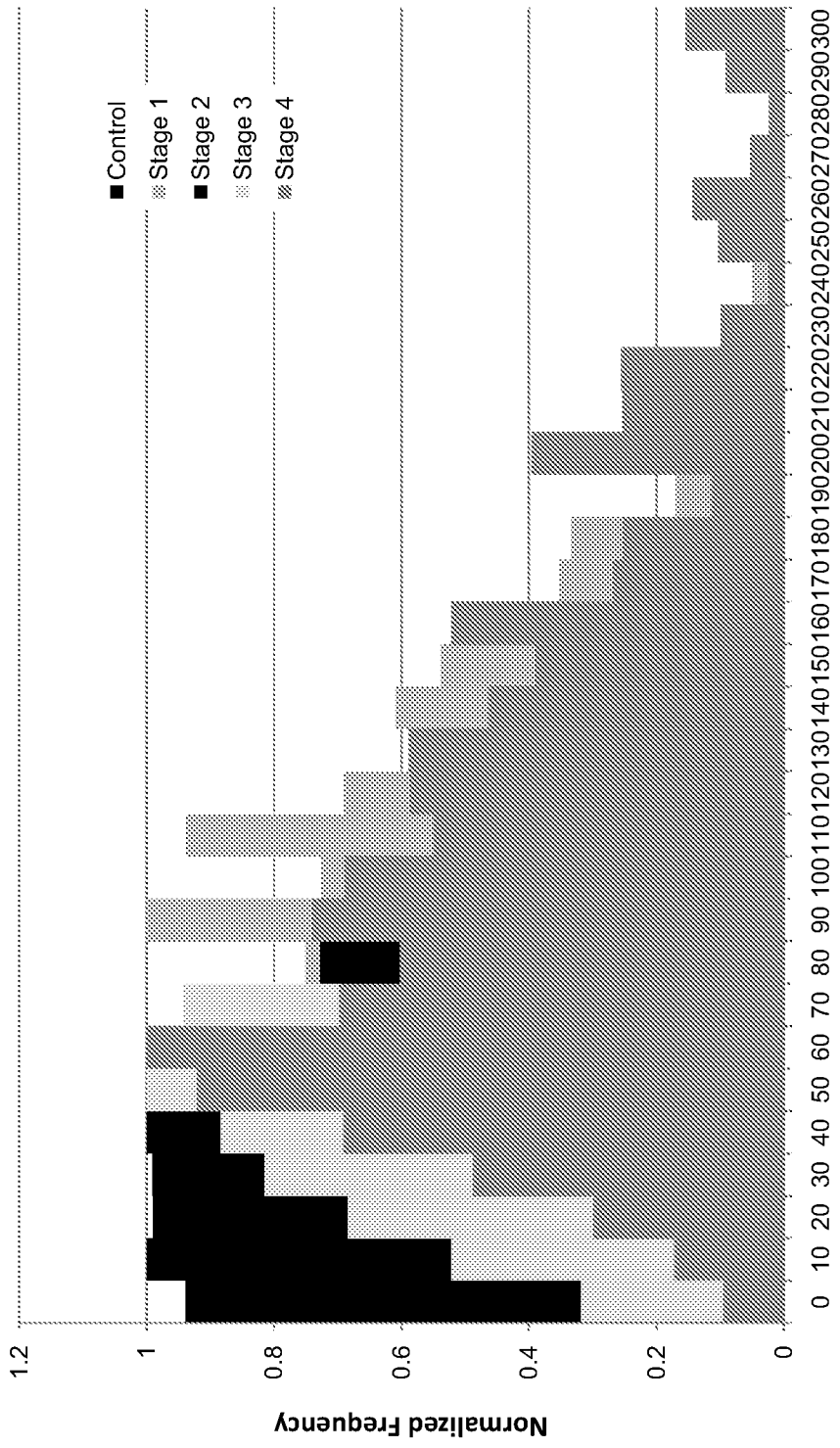

Control
Stage 1
Stage 2
Stage 3
Stage 4
Normalized Frequency
1.2
1
0.8
0.6
0.4
0.2
0
0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200 210 220 230 240 250 260 270 280 290 300
Intensity

FIG. 6E

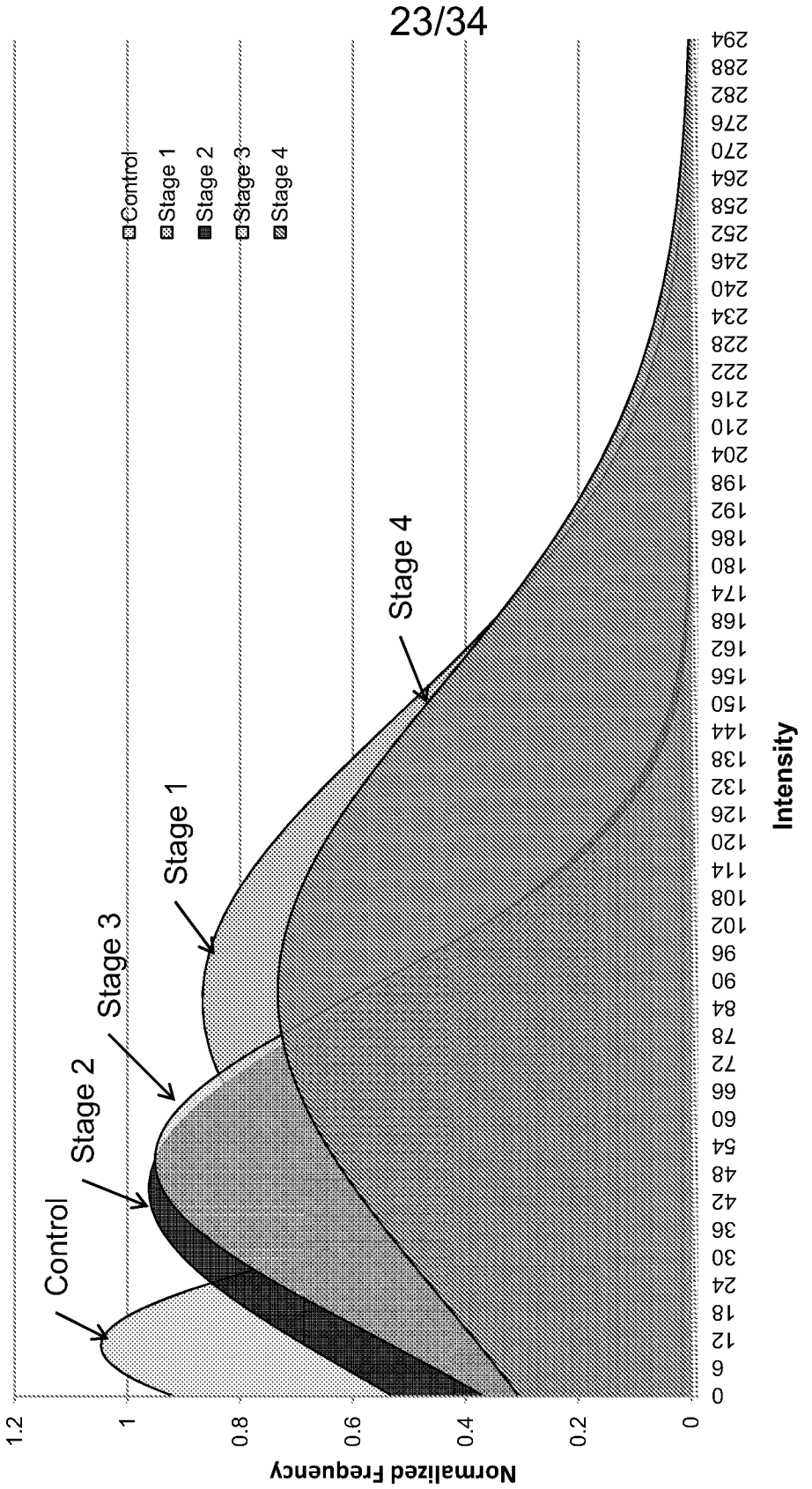

Control
Stage 1
Stage 2
Stage 3
Stage 4
Normalized Frequency
Intensity
Control
Stage 2
Stage 3
Stage 1
Stage 4

FIG. 6F

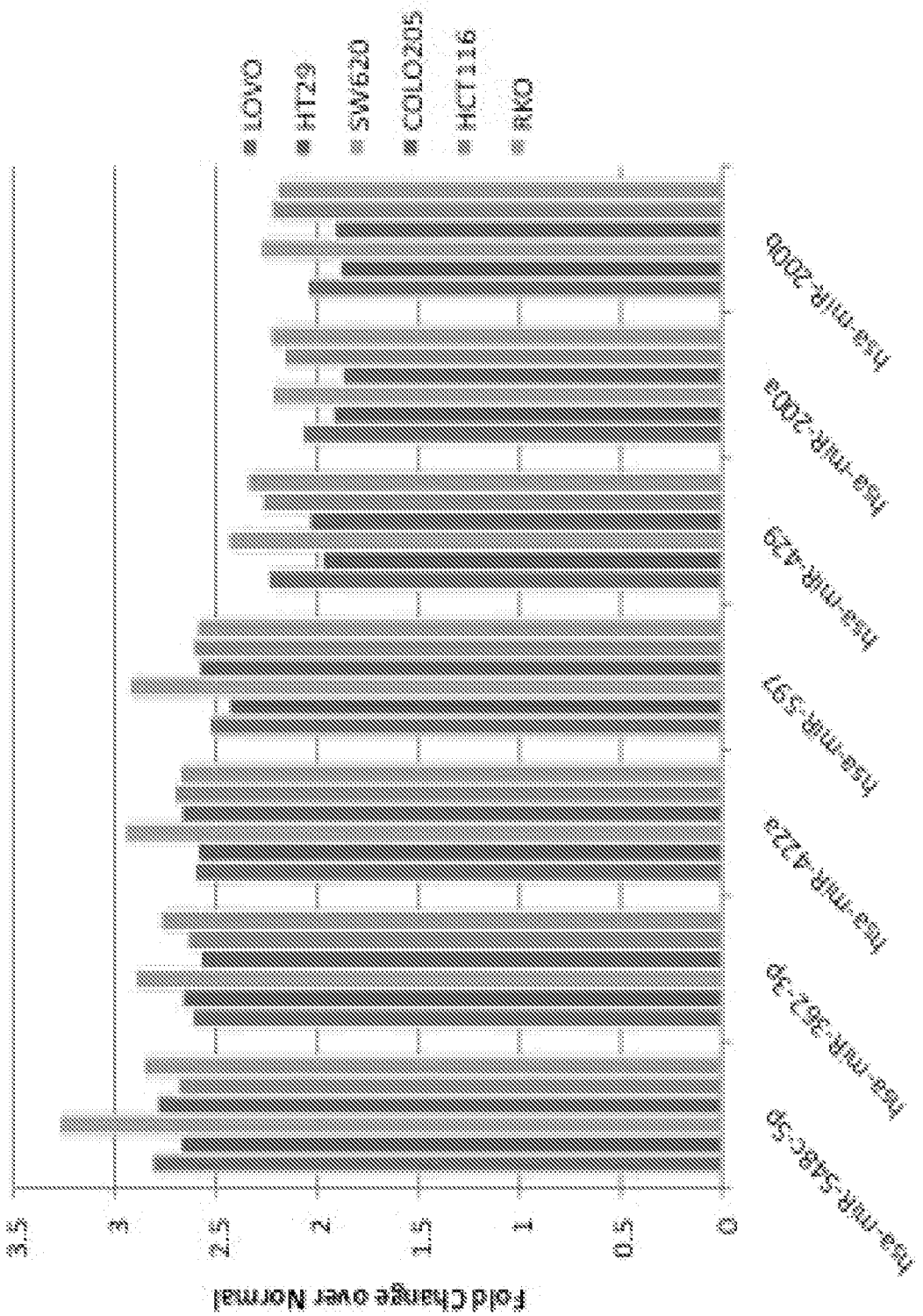

LOVO
HT29
SW620
COLO205
HCT116
RKO
Fold Change over Normal
3.5
3
2.5
2
1.5
1
0.5
0
hsa-miR-548c-5p
hsa-miR-362-3p
hsa-miR-423
hsa-miR-597
hsa-miR-200a
hsa-miR-200b

FIG. 7

Detector vs Capture	Sensitivity	Specificity	Confidence
Epcam vs CD63	95%	ND	99%
Epcam vs CD9	90%	ND	99%
CD63 vs CD63	100%	ND	99%
CD9 vs CD63	100%	ND	99%
CD66 vs CD9	85%	ND	99%

FIG. 8A

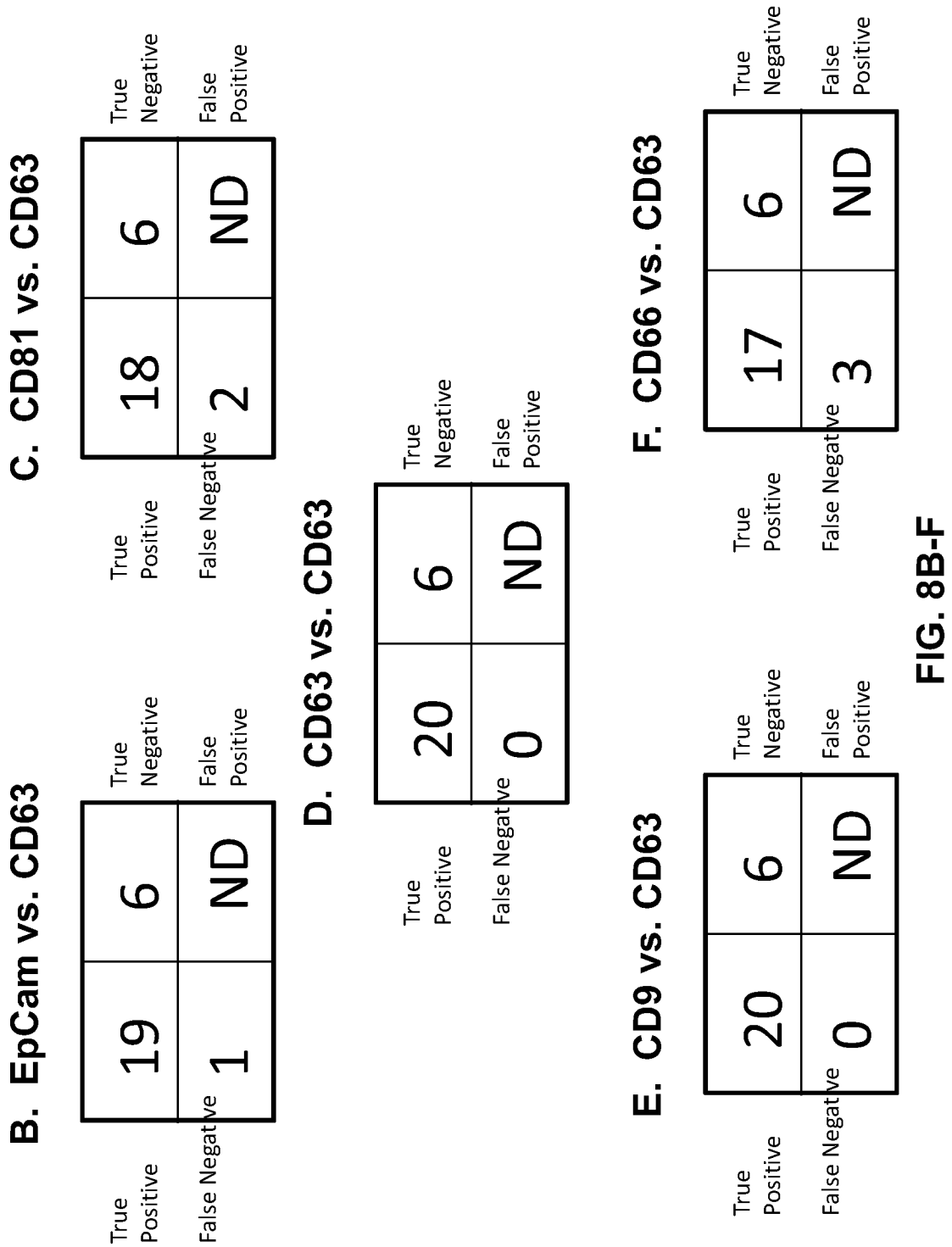

B. EpCam vs. CD63
True Positive
True Negative
False Negative
False Positive
19
6
1
ND
C. CD81 vs. CD63
True Positive
True Negative
False Negative
False Positive
18
6
2
ND
D. CD63 vs. CD63
True Positive
True Negative
False Negative
False Positive
20
6
0
ND
E. CD9 vs. CD63
True Positive
True Negative
False Negative
False Positive
20
6
0
ND
F. CD66 vs. CD63
True Positive
True Negative
False Negative
False Positive
17
6
3
ND

FIG. 8B-F

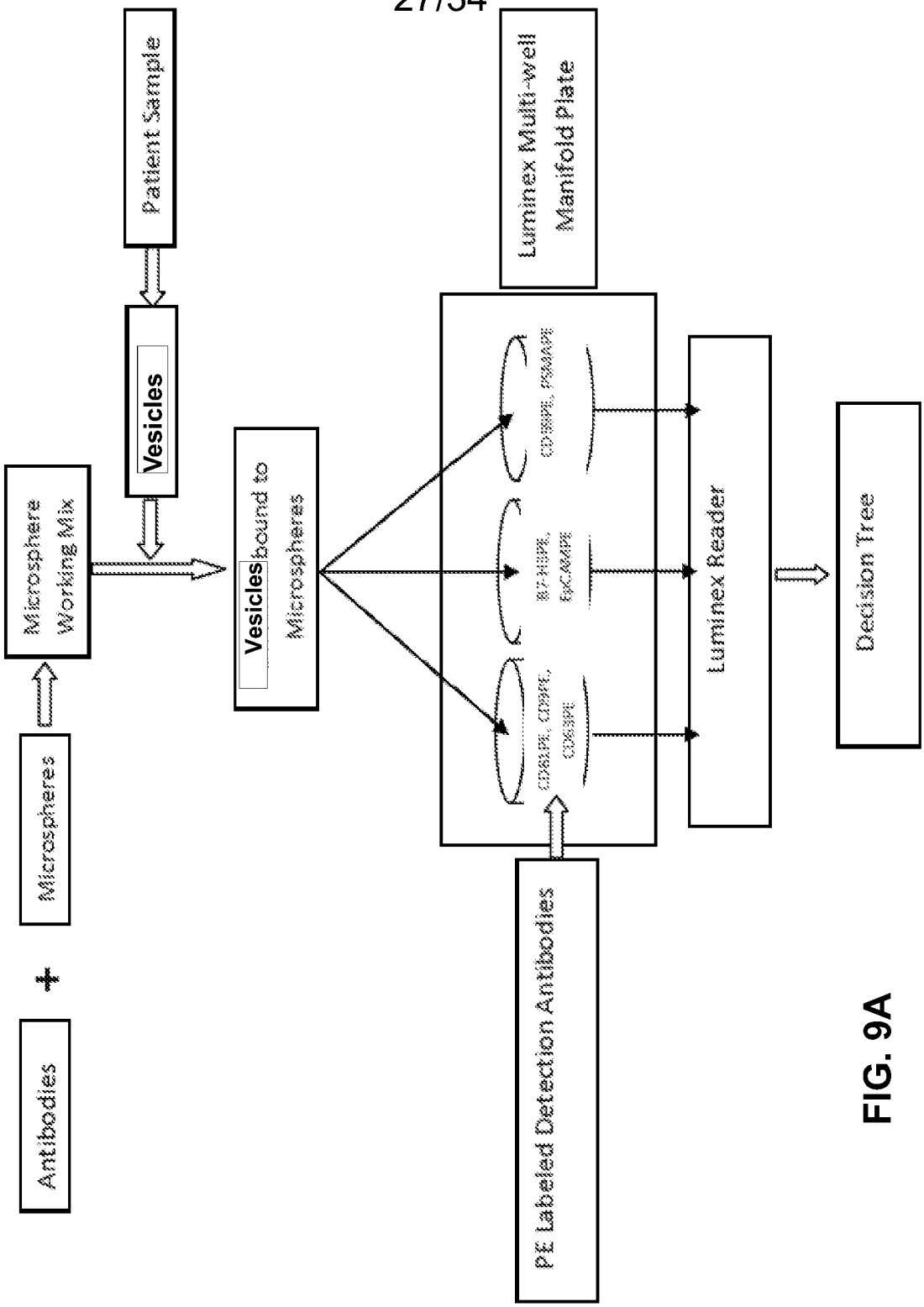

Patient Sample
Vesicles
Luminex Multi-well Manifold Plate
Microsphere Working Mix
Vesicles bound to Microspheres
Luminex Reader
Decision Tree
Microspheres
Antibodies
+
PE Labeled Detection Antibodies

FIG. 9A

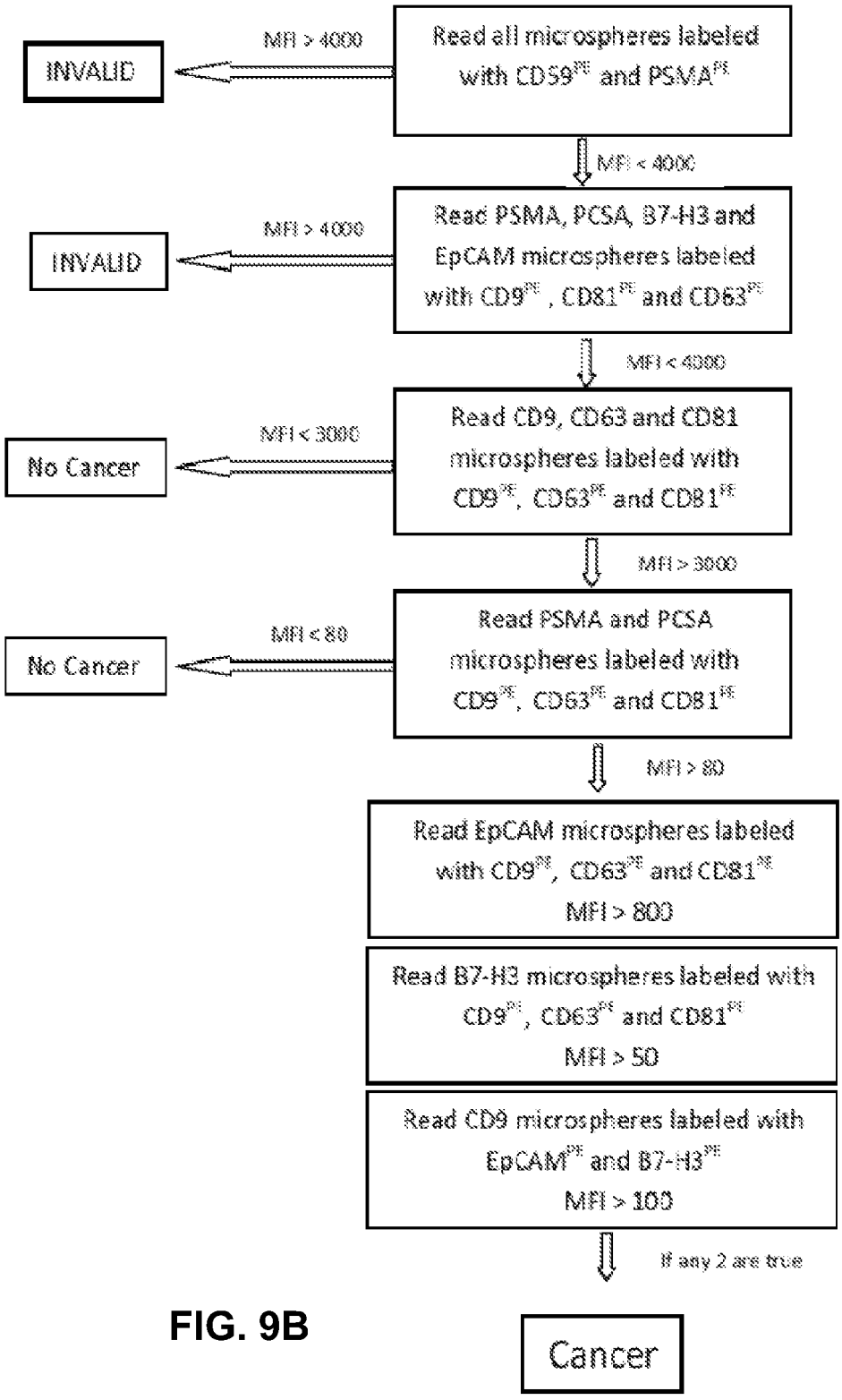

Read all microspheres labeled with CD59PE and PSMAPE
MFI > 4000
INVALID
MFI < 4000
Read PSMA, PCSA, B7-H3 and EpCAM microspheres labeled with CD9PE , CD81PE and CD63PE
MFI > 4000
INVALID
MFI < 4000
Read CD9, CD63 and CD81 microspheres labeled with CD9PE, CD63PE and CD81PE
MFI < 3000
No Cancer
MFI > 3000
Read PSMA and PCSA microspheres labeled with CD9PE, CD63PE and CD81PE
MFI < 80
No Cancer
MFI > 80
Read EpCAM microspheres labeled with CD9PE, CD63PE and CD81PE
MFI > 800
Read B7-H3 microspheres labeled with CD9PE, CD63PE and CD81PE
MFI > 50
Read CD9 microspheres labeled with EpCAMPE and B7-H3PE
MFI > 100
If any 2 are true
Cancer

FIG. 9B

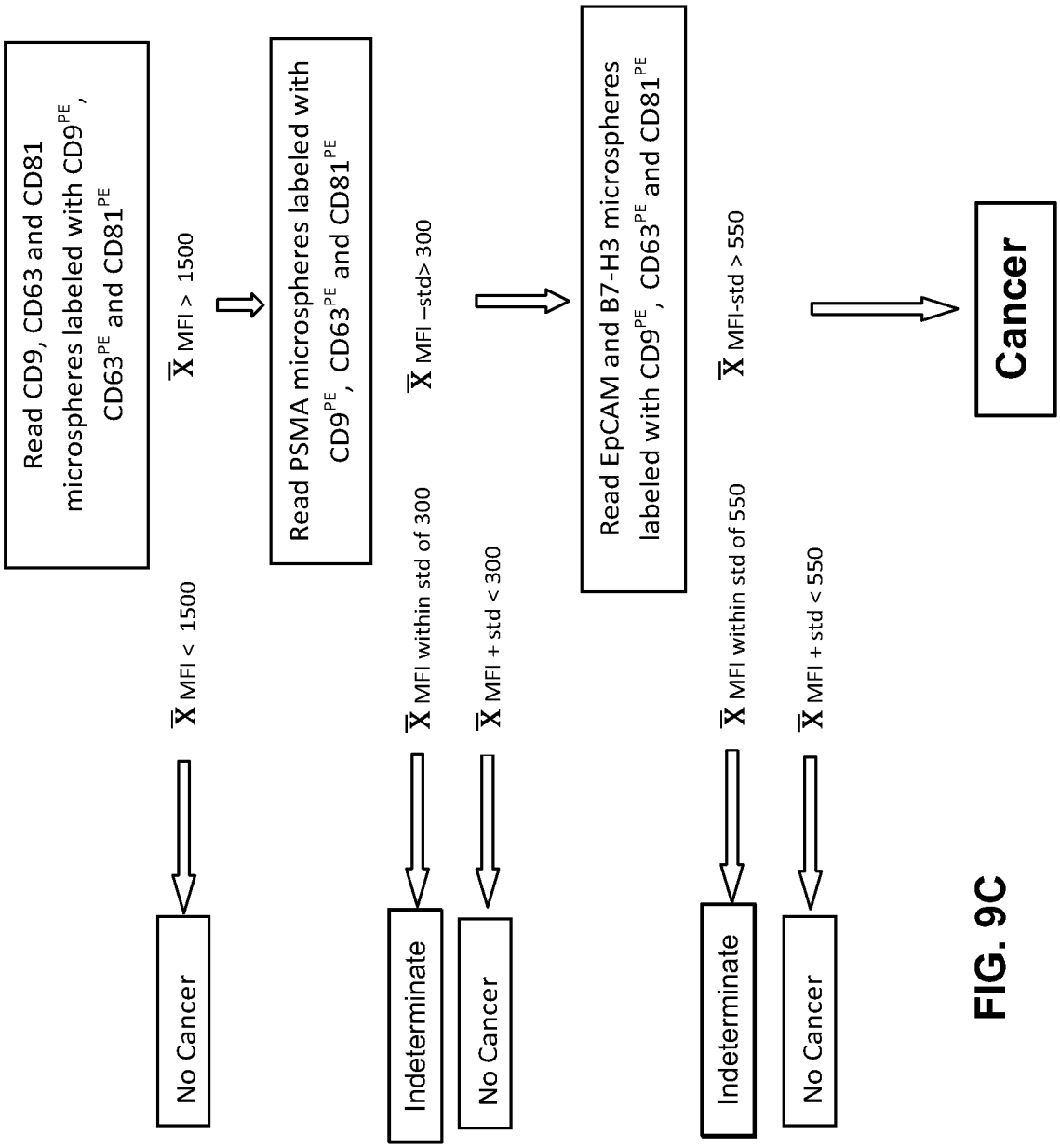

Read CD9, CD63 and CD81 microspheres labeled with CD9PE, CD63PE and CD81PE
X̄ MFI < 1500
No Cancer
X̄ MFI > 1500
Read PSMA microspheres labeled with CD9PE, CD63PE and CD81PE
X̄ MFI within std of 300
Indeterminate
X̄ MFI –std> 300
X̄ MFI + std < 300
No Cancer
Read EpCAM and B7-H3 microspheres labeled with CD9PE, CD63PE and CD81PE
X̄ MFI within std of 550
Indeterminate
X̄ MFI-std > 550
X̄ MFI + std < 550
No Cancer
Cancer

FIG. 9C

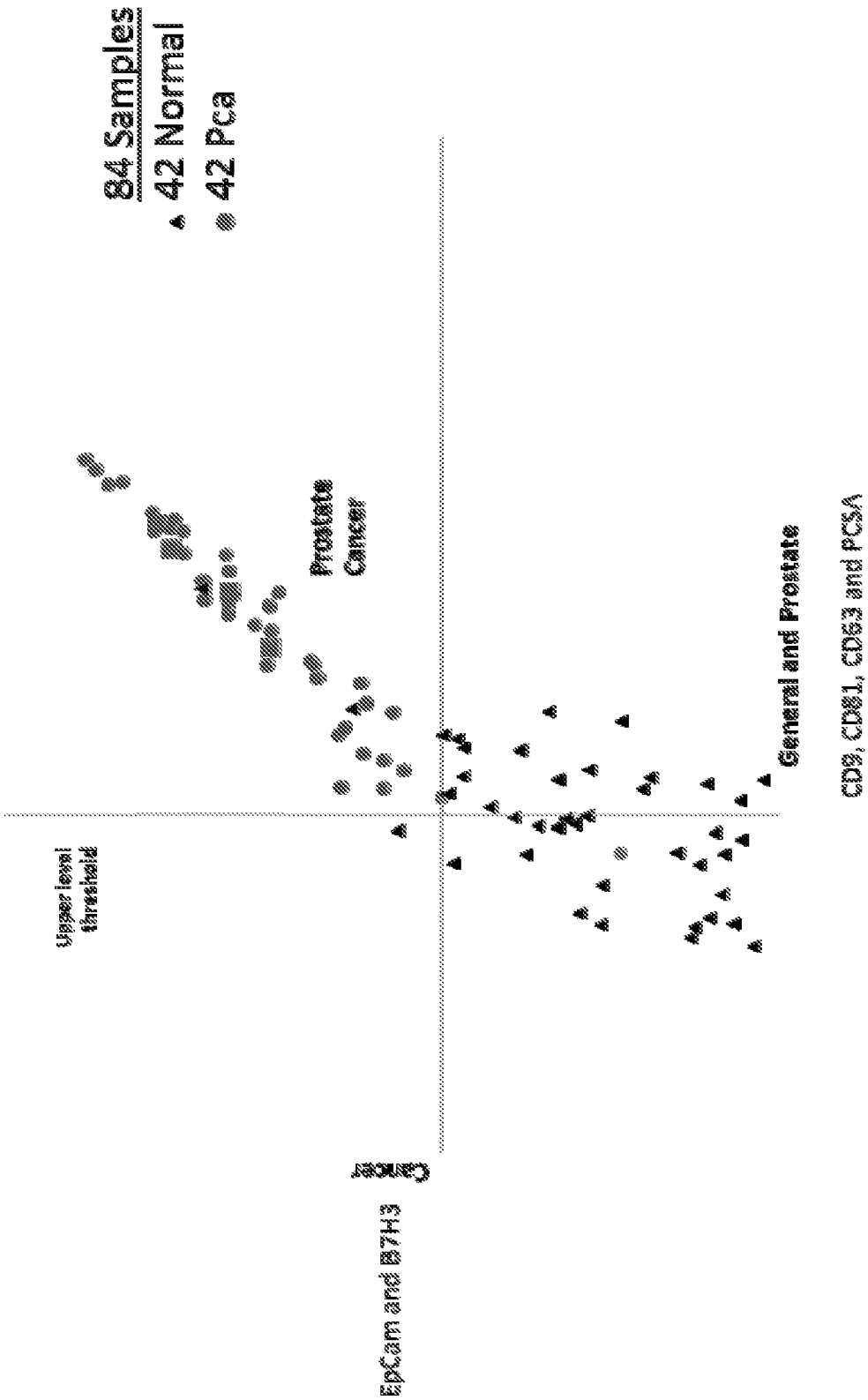

84 Samples
42 Normal
42 Pca
Prostate Cancer
General and Prostate
CD9, CD81, CD63 and PCSA
EpCam and B7H3
Upper level threshold

FIG. 10A

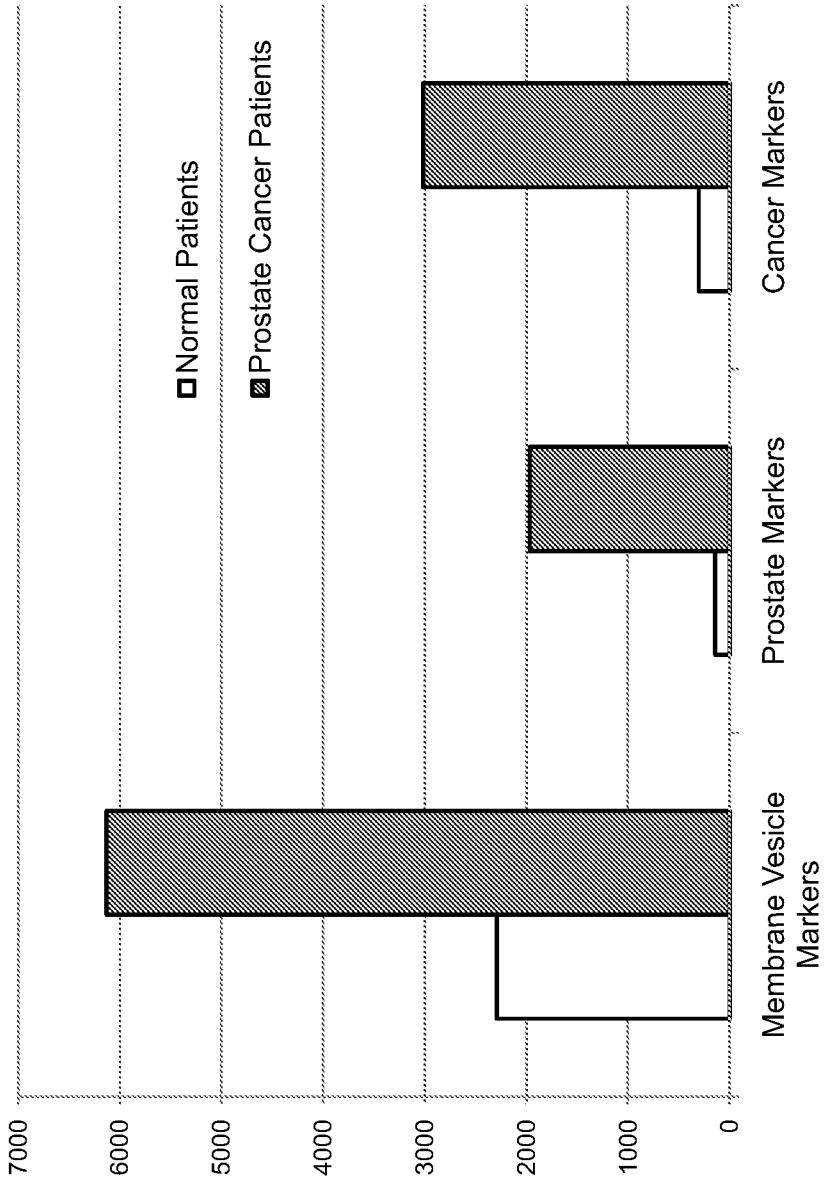

Normal Patients
Prostate Cancer Patients
7000
6000
5000
4000
3000
2000
1000
0
Membrane Vesicle Markers
Prostate Markers
Cancer Markers

FIG. 10B

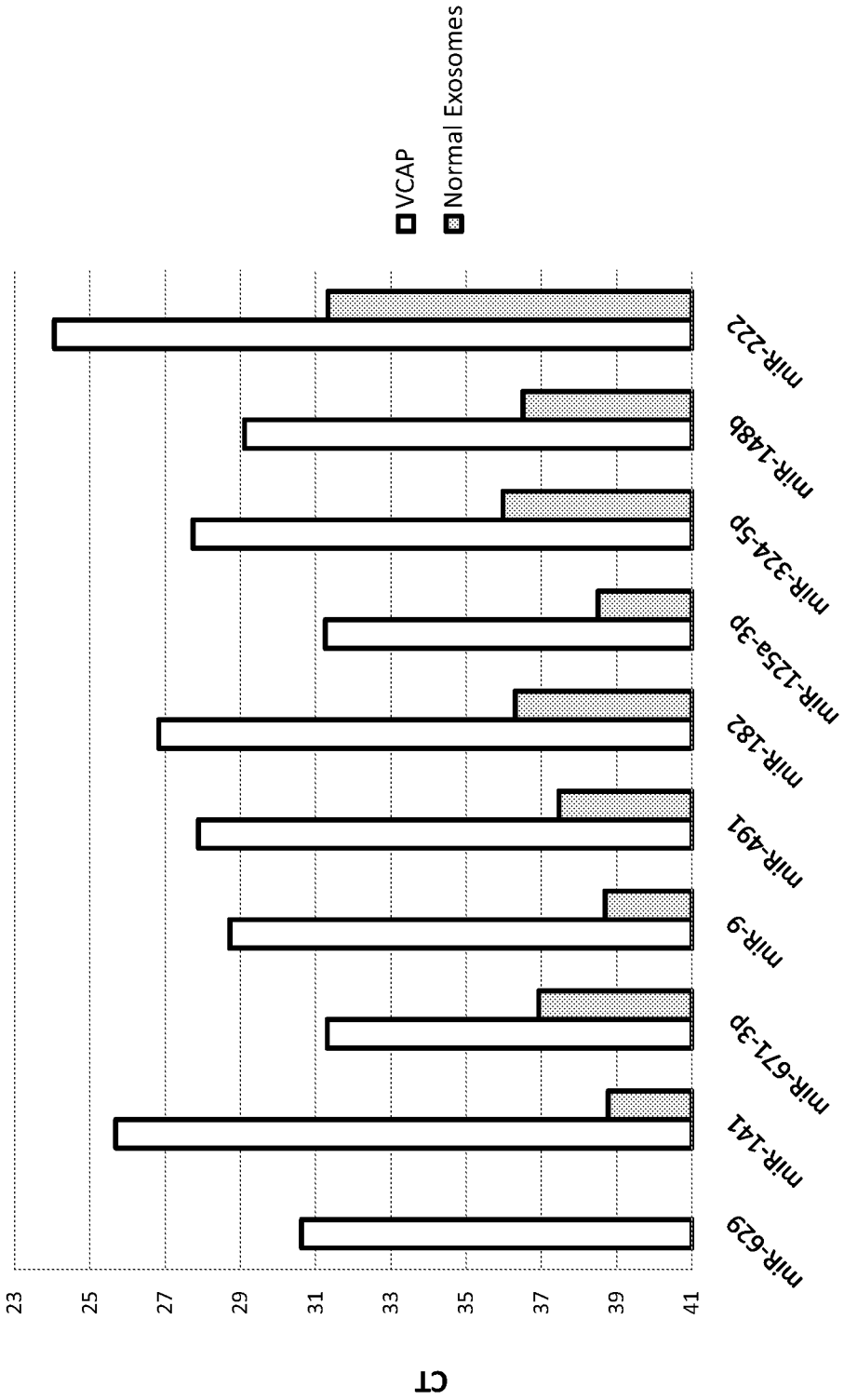

VCAP
Normal Exosomes
miR-222
miR-148b
miR-324-5p
miR-125a-3p
miR-182
miR-491
miR-9
miR-671-3p
miR-141
miR-629
23
25
27
29
31
33
35
37
39
41
CT

FIG. 11

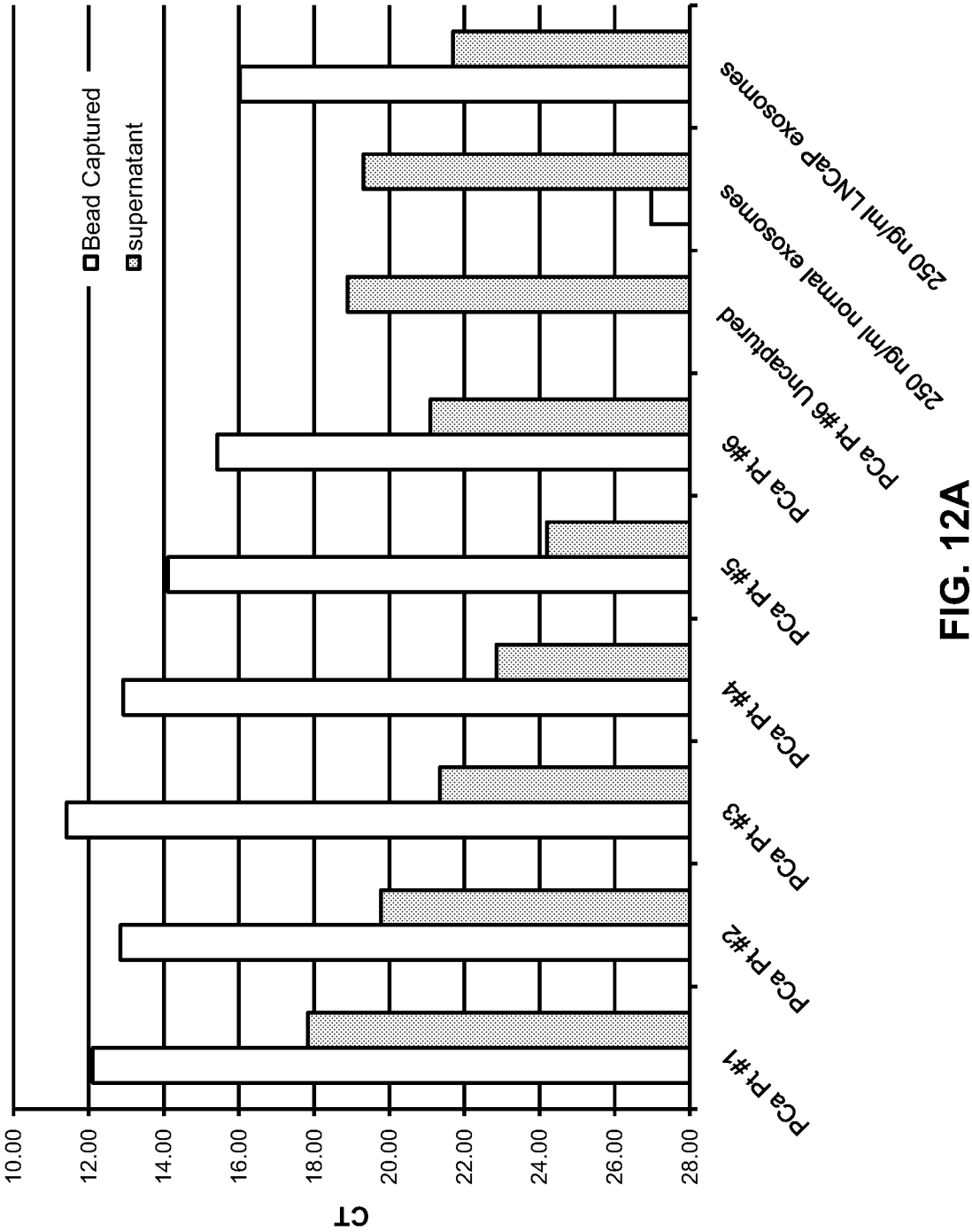

Bead Captured
supernatant
10.00
12.00
14.00
16.00
18.00
20.00
22.00
24.00
26.00
28.00
CT
PCa Pt #1
PCa Pt #2
PCa Pt #3
PCa Pt #4
PCa Pt #5
PCa Pt #6
PCa Pt #6 Uncaptured
250 ng/ml normal exosomes
250 ng/ml LNCaP exosomes

FIG. 12A

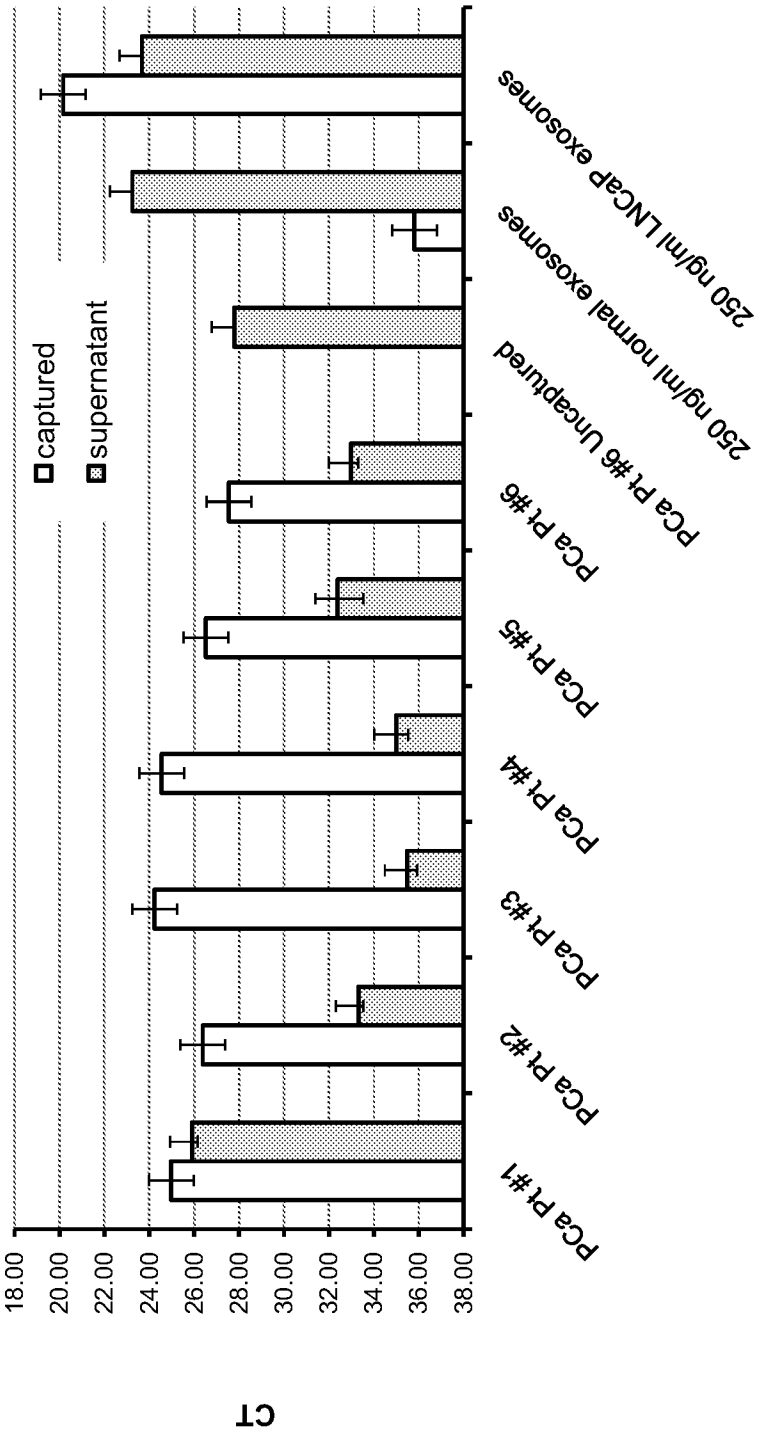

captured
supernatant
18.00
20.00
22.00
24.00
26.00
28.00
30.00
32.00
34.00
36.00
38.00
CT
PCa Pt #1
PCa Pt #2
PCa Pt #3
PCa Pt #4
PCa Pt #5
PCa Pt #6
PCa Pt #6 Uncaptured
250 ng/ml normal exosomes
250 ng/ml LNCap exosomes

FIG. 12B