Title: THERAPEUTIC OLIGONUCLEOTIDES

Abstract: Methods and compositions are provided for oligonucleotides that bind targets of interest. The targets include cells and microvesicles, such as those derived from various diseases. The oligonucleotides can be used for diagnostic and therapeutic purposes. The target of the oligonucleotides can be a therapeutic target such as Complement Component 1 Q Subcomponent (Clq) or a subunit thereof.
THERAPEUTIC OLIGONUCLEOTIDES

CROSS REFERENCE
This application claims the benefit of U.S. Provisional Patent Application Nos. 62/186,242, filed June 29, 2015; 62/198,110, filed July 28, 2015; 62/220,652, filed September 18, 2015; 62/239,226, filed October 8, 2015; 62/269,671, filed December 18, 2015; and 62/305,536, filed March 9, 2016; all of which applications are incorporated herein by reference in their entirety. This application is related to International Patent Application No. PCT/US2016/021632, filed March 9, 2016, which application is incorporated herein by reference in its entirety.

SEQUENCE LISTING SUBMITTED VIA EFS-WEB
[0001] The entire content of the following electronic submission of the sequence listing via the USPTO EFS-WEB server, as authorized and set forth in MPEP §1730 II.B.2(a), is incorporated herein by reference in its entirety for all purposes. The sequence listing is within the electronically filed text file that is identified as follows:
[0002] File Name: 825601_SeqListing_ST25.txt
[0004] Size (bytes): 883,280 bytes

BACKGROUND OF THE INVENTION
[0005] The invention relates generally to the field of aptamers capable of binding to microvesicle surface antigens, which are useful as therapeutics in and diagnostics of cancer and/or other diseases or disorders in which microvesicles implicated. The invention further relates to materials and methods for the administration of aptamers capable of binding to microvesicles. The microvesicles may be derived from cells indicative of cancer, including without limitation a breast cancer.

[0006] Aptamers are oligomeric nucleic acid molecules having specific binding affinity to molecules, which may be through interactions other than classic Watson-Crick base pairing. Unless otherwise specified, an "aptamer" as the term is used herein can refer to nucleic acid molecules that can be used to characterize a phenotype, regardless of manner of target recognition. Unless other specified, the terms "aptamer," "oligonucleotide," "polynucleotide," or the like may be used interchangeably herein.

[0007] Aptamers, like peptides generated by phage display or monoclonal antibodies ("mAbs"), are capable of specifically binding to selected targets and modulating the target's activity, e.g., through binding aptamers may block the target's ability to function. Created by an in vitro selection process from pools of random sequence oligonucleotides, aptamers have been generated for over 100 proteins including growth factors, transcription factors, enzymes, immunoglobulins, and receptors. A typical aptamer is 10-15 kDa in size (30-45 nucleotides), binds its target with sub-nanomolar affinity, and discriminates against closely related targets (e.g., aptamers will typically not bind other proteins from the same gene family). A series of structural studies have shown that aptamers are capable of using the same types of binding
interactions (e.g., hydrogen bonding, electrostatic complementarity, hydrophobic contacts, steric exclusion) that drive affinity and specificity in antibody-antigen complexes.

[0008] Aptamers have a number of desirable characteristics for use as therapeutics and diagnostics including high specificity and affinity, biological efficacy, and excellent pharmacokinetic properties. In addition, they offer specific competitive advantages over antibodies and other protein biologies, for example:

[0009] Speed and control. Aptamers are produced by an entirely in vitro process, allowing for the rapid generation of initial leads, including therapeutic leads. In vitro selection allows the specificity and affinity of the aptamer to be tightly controlled and allows the generation of leads, including leads against both toxic and non-immunogenic targets.

[0010] Toxicity and Immunogenicity. Aptamers as a class have demonstrated little or no toxicity or immunogenicity. In chronic dosing of rats or woodchucks with high levels of aptamer (10 mg/kg daily for 90 days), no toxicity is observed by any clinical, cellular, or biochemical measure. Whereas the efficacy of many monoclonal antibodies can be severely limited by immune response to antibodies themselves, it is extremely difficult to elicit antibodies to aptamers most likely because aptamers cannot be presented by T-cells via the MHC and the immune response is generally trained not to recognize nucleic acid fragments.

[0011] Administration. Whereas most currently approved antibody therapeutics are administered by intravenous infusion (typically over 2-4 hours), aptamers can be administered by subcutaneous injection (aptamer bioavailability via subcutaneous administration is >80% in monkey studies (Tucker et al., J. Chromatography B. 732: 203-212, 1999)). This difference is primarily due to the comparatively low solubility and thus large volumes necessary for most therapeutic mAbs. With good solubility (>150 mg/mL) and comparatively low molecular weight (aptamer: 10-50 kDa; antibody: 150 kDa), a weekly dose of aptamer may be delivered by injection in a volume of less than 0.5 mL. In addition, the small size of aptamers allows them to penetrate into areas of conformational constrictions that do not allow for antibodies or antibody fragments to penetrate, presenting yet another advantage of aptamer-based therapeutics or prophylaxis.

[0012] Scalability and cost. Aptamers are chemically synthesized and are readily scaled as needed to meet production demand for diagnostic or therapeutic applications. Whereas difficulties in scaling production are currently limiting the availability of some biologies and the capital cost of a large-scale protein production plant is enormous, a single large-scale oligonucleotide synthesizer can produce upwards of 100 kg/year and requires a relatively modest initial investment. The current cost of goods for aptamer synthesis at the kilogram scale is estimated at $100/g, comparable to that for highly optimized antibodies.

[0013] Stability. Aptamers are chemically robust. They are intrinsically adapted to regain activity following exposure to factors such as heat and denaturants and can be stored for extended periods (>1 yr) at room temperature as lyophilized powders.
INCORPORATION BY REFERENCE

[0014] 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.

SUMMARY OF THE INVENTION

[0015] Compositions and methods of the invention provide aptamers that bind biomarkers of interest such as microvesicle surface antigens or functional fragments of microvesicle surface antigens. In various embodiments, oligonucleotide probes of the invention are used in diagnostic, prognostic or theranostic processes to screen a biological sample for the presence or levels of biomarkers, including without limitation microvesicle surface antigens, determined to provide a relevant readout. The diagnosis may be related to a disease or disorder, e.g., a cancer. In other embodiments, oligonucleotide probes of the invention are chemically modified or composed in a pharmaceutical composition for therapeutic applications.

[0016] In an aspect, the invention provides an oligonucleotide comprising a sequence selected from any one of SEQ ID NOs. 1472, 1475, 1477, 1482, and 4151-4325. In a related aspect, the invention provides an oligonucleotide comprising a central sequence comprising a sequence according to any one of SEQ ID NOs. 1472, 1475, 1477, 1482, and 4151-4325, and a 5' region with sequence 5' - CTAGCATGACTGAGTAGCT (SEQ ID NO. 131), a 3' region with sequence 5' - CTGTCTTATACACATCTGTGACCGTCCGACGA (SEQ ID NO. 132), or both. The oligonucleotides of the invention can be capable of binding to complement protein complex Clq or a subunit thereof.

[0017] In another aspect, the invention provides an oligonucleotide comprising the sequence 5' - ACAATCC (SEQ ID NO. 4151), wherein the oligonucleotide is capable of binding to complement protein complex Clq or a subunit thereof. In still another aspect, the invention further provides an oligonucleotide comprising the sequence 5' - ACTATAGAACATCCACCGCTTTAATCAACTATCTTA (SEQ ID NO. 1472) or 5' - AACCCAAACCTCAACTCTTCACCCACGAACTAAACA (SEQ ID NO. 1482), wherein the oligonucleotide is capable of binding to complement protein complex Clq or a subunit thereof.

[0018] The Clq subunit targeted by the oligonucleotides of the invention can be the A, B or C subunit. The subunits can carry various post-translational modifications. For example, the Clq A subunit can have at least one modification selected from Table 46.

[0019] The invention further provides an oligonucleotide comprising a nucleic acid sequence or a portion thereof that is at least 50, 55, 60, 65, 70, 75, 80, 85, 86, 88, 89, 90, 95, 96, 97, 98, 99 or 100 percent homologous to an oligonucleotide sequence described above.

[0020] In another aspect, the plurality of oligonucleotides comprising at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130, 140, 150, 160, 170 or all oligonucleotide sequences described above. For example, the invention provides a plurality of oligonucleotides comprising at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 200, 300, 400, 415.
different oligonucleotide sequences comprising at least one of the sequences 5'-ACAATCC (SEQ ID NO. 4151), 5' -ACTATAGAACAATCCACCGCTTTAAATCACTATCTTTA (SEQ ID NO. 1472) or 5' -AACCCAAATCACATACACTCTCACCACCCCACAA (SEQ ID NO. 1482), wherein the oligonucleotide are capable of binding Clq or a subunit thereof.

[0021] The oligonucleotide or the plurality of oligonucleotides provided by the invention comprises at least one of DNA, RNA, PNA, and any combination thereof. The oligonucleotide or the plurality of oligonucleotides may comprise at least one functional modification selected from the group consisting of biotinylation, a non-naturally occurring nucleotide, a deletion, an insertion, an addition, and a chemical modification. In some embodiments, the chemical modification comprises at least one of C18, polyethylene glycol (PEG), PEG4, PEG6, PEG8, and PEG12.

[0022] In an aspect, the invention provides a method comprising synthesizing the at least one oligonucleotide or the plurality of oligonucleotides provided by the invention.

[0023] In another aspect, the invention provides a method comprising contacting a biological sample with the at least one oligonucleotide or the plurality of oligonucleotides provided by the invention. For example, the at least one oligonucleotide or the plurality of oligonucleotides can be as described above. In some embodiments, the method further comprises detecting a presence or level of complement protein complex Clq or a subunit thereof in the biological sample that is bound by the at least one oligonucleotide or at least one member of the plurality of oligonucleotides. Any appropriate method of the detecting the complex can be used. For example, detecting the oligonucleotide/s or other nucleic acids may comprise at least one of sequencing, amplification, hybridization, gel electrophoresis, chromatography, and any combination thereof. Any useful sequencing method can be employed, including without limitation at least one of next generation sequencing, dye termination sequencing, pyrosequencing, and any combination thereof. Detecting the Clq protein, a subunit, or a related protein (such as a protein in complex with Clq) may comprise one or more of an immunoassay, enzyme immunoassay (EIA), enzyme-linked immunosorbent assay (ELISA), enzyme-linked oligonucleotide assay (ELONA), affinity isolation, immunoprecipitation, Western blot, gel electrophoresis, or flow cytometry. One of skill will appreciate that immunoassays can employ various useful reporter systems, including without limitation enzymes, radioactive isotopes, nucleic acid reporters, fluorogenic reporters, electrochemiluminescent tags, or label-free formats.

[0024] Clq may be free in solution or part of a larger biological complex. In some cases, the Clq or subunit thereof may be associated with a microvesicle population. As desired, the method of the invention comprises isolating the microvesicle population prior to contacting the sample with the oligonucleotide or plurality of oligonucleotides, after such contacting, or both before and after. The isolating may be in whole or in part. For example, the microvesicle population may be partially isolated from other components in the sample before or after contacting the sample with the oligonucleotide or plurality of oligonucleotides. The invention may use any appropriate techniques to isolate microvesicles. Various techniques of isolating microvesicles are disclosed herein or known in the art, including without limitation
affinity purification, filtration, concentration, polymer precipitation, PEG precipitation, ultracentrifugation, a molecular crowding reagent, affinity selection, chromatography, or any combination thereof.

[0025] Any desired biological sample can be contacted with the oligonucleotide or plurality of oligonucleotides according to the invention. In various embodiments, the sample comprises a bodily fluid, tissue sample or cell culture. Any appropriate bodily fluid can be contacted, including without limitation whole blood, peripheral blood, sera, plasma, ascites, urine, cerebrospinal fluid (CSF), sputum, saliva, bone marrow, synovial fluid, aqueous humor, amniotic fluid, cerumen, breast milk, bronchoalveolar lavage fluid, semen, prostatic fluid, cowper’s fluid or pre-ejaculatory fluid, female ejaculate, sweat, fecal matter, hair oil, 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, blastocyl cavity fluid, or umbilical cord blood. In certain preferred embodiments, the bodily fluid comprises whole blood or a derivative or fraction thereof, such as sera or plasma. The biological sample may comprise purified Clq, recombinant Clq, both, or fragments thereof. For example, such Clq or subunits thereof can be spiked into the bodily fluid, tissue sample or cell culture sample.

[0026] As desired, the method of detecting the presence or level of complement protein complex Clq or a subunit thereof may be used to characterize a phenotype. The phenotype can be any appropriate phenotype, including without limitation a disease or disorder. In such cases, the characterizing may include providing, or assisting in providing, at least one of diagnostic, prognostic and theranostic information for the disease or disorder. Characterizing the phenotype may comprise comparing the presence or level to a reference. Any appropriate reference level can be used. For example, the reference can be the presence or level determined in a sample from at least one individual without the phenotype or from at least one individual with a different phenotype. As a further example, if the phenotype is a disease or disorder, the reference level may be the presence or level determined in a sample from at least one individual without the disease or disorder, or with a different state of the disease or disorder (e.g., in remission, different stage or grade, different prognosis, metastatic versus local, etc).

[0027] As noted, the sample can be from a subject suspected of having or being predisposed to a disease or disorder. The disease or disorder can be any disease or disorder that can be assessed by the subject method. For example, the disease or disorder may be a cancer, a premalignant condition, an inflammatory disease, an immune disease, an autoimmune disease or disorder, a pregnancy related disorder, a cardiovascular disease or disorder, a neurological disease or disorder, an infectious disease or pain. In an embodiment, 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; lung 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; Sezary 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; Waldenstrom macroglobulinemia; or Wilm's tumor. The premalignant condition can be Barrett's Esophagus. The autoimmune disease can 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 erythematosus (SLE), Hashimoto's Thyroiditis, Grave's disease, Ankylosing Spondylitis Sjogrens Disease, CREST syndrome, Scleroderma, Rheumatic Disease, organ rejection, Primary Sclerosing Cholangitis, or sepsis. The cardiovascular disease can be atherosclerosis, congestive heart failure, vulnerable plaque, stroke, ischemia, high blood pressure, stenosis, vessel occlusion or a
thrombotic event. The neurological disease can be Multiple Sclerosis (MS), Parkinson's Disease (PD),
Alzheimer's Disease (AD), schizophrenia, bipolar disorder, depression, autism, Prion Disease, Pick's
disease, dementia, Huntington disease (HD), Down syndrome, cerebrovascular disease, Rasmussen's
encephalitis, viral meningitis, neuropsychiatric systemic lupus erythematosus (NPSLE), amyotrophic
lateral sclerosis, Creutzfeldt-Jacob disease, Gerstmann-Straussler-Scheinker disease, transmissible
spongiform encephalopathy, ischemic reperfusion damage (e.g. stroke), brain trauma, microbial infection,
or chronic fatigue syndrome. The pain can be fibromyalgia, chronic neuropathic pain, or peripheral
neuropathic pain. The infectious disease can be a bacterial infection, viral infection, yeast infection,
Whipple's Disease, Prion Disease, cirrhosis, methicillin-resistant staphylococcus aureus, HIV, HCV,
hepatitis, syphilis, meningitis, malaria, tuberculosis, influenza.

[0028] As further described herein, the invention provides a kit comprising a reagent for carrying out the
method. Similarly, the invention provides for the use of a reagent for carrying out the method. The reagent
can be any useful reagent for carrying out the method. For example, the reagent can be at least one
oligonucleotide or a plurality of oligonucleotides as described above, or one or more primers or primer
sets for amplification or sequencing of such oligonucleotides. In some embodiments, the reagent
comprises at least one binding agent to at least one protein, a binding buffer with or without MgC\(^n\), a
sample processing reagent, a microvesicle isolation reagent, a detection reagent, a secondary detection
reagent, a wash buffer, an elution buffer, a solid support, and any combination thereof. The microvesicle
isolation reagent may comprise at least one of a concentrator unit, a filtration unit, a polymer, PEG, a size
exclusion column, a binding agent to a microvesicle antigen, and any combination thereof; and/or the
detection or secondary detection agent comprises streptavidin-horse radish peroxide (HRP), a
streptavidin-conjugated fluorophore, a streptavidin-conjugated quantum dot, and any combination thereof.

[0029] In an aspect, the invention provides a pharmaceutical composition comprising a therapeutically
effective amount of the at least one oligonucleotide or the plurality of oligonucleotides of the invention, or
a salt thereof, and a pharmaceutically acceptable carrier, diluent, or both. In a related aspect, the invention
provides a method of treating or ameliorating a disease or disorder in a subject in need thereof, comprising
administering such pharmaceutical composition to the subject. In another related aspect, the invention
provides a method of inhibiting C1q in a subject, comprising administering such pharmaceutical
composition to the subject. The pharmaceutical composition can be administered in any useful format. In
various embodiments, the administering comprises at least one of intradermal, intramuscular,
intraperitoneal, intravenous, subcutaneous, intranasal, epidural, oral, sublingual, intracerebral,
intravaginal, transdermal, rectal, by inhalation, topical administration, or any combination thereof. The
carrier or diluent can be any useful carrier or diluent, as described herein or known in the art.

[0030] In an aspect, the invention provides a multipartite construct that comprises a first segment that
binds to a first target and a second segment that binds to a second target. As desired, the first segment, the
second segment, or both, comprises a sequence selected from any one of SEQ ID Nos. 137-969 and 1072-
4325, or a region that is at least 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 96, 97, 98, 99 or 100 percent
homologous thereto.
In a preferred embodiment, the first target comprises a cell-of-origin or disease specific cellular marker, including without limitation a protein marker that may be found on a cellular surface. For example, the first target can be a cancer antigen, a tumor antigen, a cell surface antigen, a microvesicle surface antigen, a cell receptor and/or a membrane antigen. Many useful targets are provided herein. In some embodiments, the first target is selected from any one of Table 3, Table 4, Tables 18-26, or Table 45. The first target can be selected from those provided in the Section "Anti-target and multivalent oligonucleotides" herein. For example, the first target can be selected from the group consisting of CD2, CD4, CD19, CD20, CD22, CD23, CD30, CD33, CD37, CD40, CD44v6, CD52, CD56, CD70, CD74, CD79a, CD80, CD98, CD138, EGFR (Epidermal growth factor receptor), VEGF (Vascular endothelial growth factor), VEGFR1 (Vascular endothelial growth factor receptor 1), PDGFR (Platelet-derived growth factor receptor), RANKL (Receptor activator of nuclear factor kappa-B ligand), GPNMB (Transmembrane glycoprotein Neuromedin B), EphA2 (Ephrin type-A receptor 2), PSMA (Prostate-specific membrane antigen), Cripto (Cryptic family protein IB), EpCAM (Epithelial cell adhesion molecule), CTLA4 (Cytotoxic T-Lymphocyte Antigen 4), IGF- IR (Type I insulin-like growth factor receptor), GP3 (MI 3 bacteriophage), GP9 (Glycoprotein IX), CD42a, GP 40 (Glycoprotein 40kDa), GPC3 (glypican-3), GPC1 (glypican-1), TRAILR1 (Tumor necrosis factor-related apoptosis-inducing ligand receptor 1), TRAILR II (Tumor necrosis factor-related apoptosis-inducing ligand receptor II), FAS (Type II transmembrane protein), PS (phosphatidylinositol serine) lipid, Mucl (Mucin 1, PEM), Mucl8, CD146, α5β1 integrin, α4β1 integrin, αvβ integrin (Vitronectin Receptor), Chondrolectin, CAIX (Carbonic anhydrase IX), GD2 gangioside, GD3 gangioside, GM1 gangloside, Lewis Y antigen, Mesothelin, HER2 (Human Epidermal Growth factor 2), HER3, HER4, FN14 (Fibroblast Growth Factor Inducible 14), CS1 (Cell surface glycoprotein, CD2 subset 1, CRACC, SLAMF7, CD319), 41BB CD137, SIP (Siah-1 Interacting Protein), CTGF (Connective tissue growth factor), HLADR (MHC class II cell surface receptor), PD-1 (Programmed Death 1, Type I membrane protein, PD-L1 (Programmed Death Ligand 1), PD-L2 (Programmed Death Ligand 2), IL-2 (Interleukin-2), IL-8 (Interleukin-8), IL-13 (Interleukin-13), PIGF (Phosphatidylinositol-glycan biosynthesis class F protein), NRPI (Neuropilin-1), ICAM1, CD54, GC182 (Claudin 18.2), Claudin, HGF (Hepatocyte growth factor), CEA (Carcinoembryonic antigen), LTβR (lymphotoxin β receptor), Kappa Myeloma, Folate Receptor alpha, GRP78 (BIP, 78 kDa Glucose-regulated protein), A33 antigen, PSA (Prostate-specific antigen), CA 125 (Cancer antigen 125 or carbohydrate antigen 125), CA19.9, CA15.3, CA242, leptin, prolactin, osteopontin, IGF- II (Insulin-like growth factor 2), fascin, sPlgR (secreted chain of polymorphic immunoglobulin receptor), 14-3-3 eta protein, 5T4, ETA (epithelial tumor antigen), MAGE (Melanoma-associated antigen), MAPG (Melanoma-associated proteoglycan, NG2), vimentin, EPCA-1 (Early prostate cancer antigen-2), TAG-72 (Tumor-associated glycoprotein 72), factor VIII, Neprilysin (Membrane metallo-endopeptidase), 17-1 A (Epithelial cell surface antigen 17-1A), nucleolin, nucleophosmin, and any combination thereof. In some embodiments, the first target comprises CD20. In still other embodiments, the first target comprises cMet.

As described herein, the invention contemplates various configuration of the multipartite construct. See, e.g., FIGS. 23A-D herein and related discussion herein. In preferred embodiments, the first
segment comprises an oligonucleotide. Such construct may further comprise a first oligonucleotide primer region and/or a second oligonucleotide primer region surrounding the first segment.  

[0033] In some embodiments, the second target of the multipartite construct of the invention comprises an immunomodulatory molecule. For example, the immunomodulatory molecule may be selected from at least one of a member of the innate immune system, a member of the complement system, Clq, Clr, Cls, CI, C3a, C3b, C3d, C5a, C2, C4, and any combination thereof. In some embodiments, the second target comprises Clq or a subunit thereof, e.g., the A, B or C subunit. The second target may comprise any number of post-translational modifications. For example, when the immunomodulatory molecule comprises Clq, the A subunit may have at least one modification selected from Table 46.  

[0034] As noted above, the invention contemplates various configurations of the multipartite construct. In various embodiments of the invention, the second segment comprises an antibody or oligonucleotide. And in preferred embodiments, the second segment comprises an oligonucleotide. Such a second segment may further comprise a first oligonucleotide primer region and/or a second oligonucleotide primer region surrounding the second segment. The second segment may comprise an oligonucleotide as provided by the invention. For example, the oligonucleotide can be as described above, including without limitation a sequence selected from any one of SEQ ID NOs. 137-969 and 1072-4325, or a region that is at least 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 96, 97, 98, 99 or 100 percent homologous thereto.  

[0035] As desired, the multipartite construct of the invention can include a linker region between the first segment and second segment. The linker region can be chosen to achieve any desired purpose, such as modulate the distance between the first and second targets, e.g., to relieve steric hindrance or to bring the targets into close proximity. The linker can also provide certain functionalities, including without limitation an immunostimulatory sequence, an anti-proliferative sequence, a pro-apoptotic sequence, or some combination thereof. In an embodiment, the linker region comprises one or more CpG motif. In another embodiment, the linker comprises a polyG sequence. In still other embodiments, the multipartite construct of the invention comprises an immunostimulating moiety, a membrane disruptive moiety, or both.  

[0036] The multipartite construct of the invention can be modified to comprise at least one oligonucleotide chemical modification. Various useful modifications are disclosed herein or known in the art. In some embodiments, the modification is selected from the group consisting of a chemical substitution at a sugar position; a chemical substitution at a phosphate position; a chemical substitution at a base position of the nucleic acid; and any combination thereof. The modification can be selected from the group consisting of: incorporation of a modified nucleotide, 3’ capping, conjugation to an amine linker, conjugation to a high molecular weight, non-immunogenic compound, conjugation to a lipophilic compound, conjugation to a drug, conjugation to a cytotoxic moiety, labeling with a radioisotope, and any combination thereof. In a preferred embodiment, the non-immunogenic, high molecular weight compound is polyalkylene glycol, such as polyethylene glycol.  

[0037] In preferred embodiments, the multipartite construct comprises an oligonucleotide polymer. Such a construct can be flanked by a first oligonucleotide primer region and a second oligonucleotide primer.
region. In various embodiments, the first target comprises CD20 or cMET and the second segment comprises an oligonucleotide provided by the invention, e.g., an oligonucleotide that can bind Clq or a subunit thereof.

[0038] In a related aspect, the invention provides a pharmaceutical composition comprising a therapeutically effective amount of the multipartite construct described herein, or a salt thereof, and a pharmaceutically acceptable carrier, diluent or both. In a related aspect, the invention provides a method of treating or ameliorating a disease or disorder in a subject need thereof, comprising administering such pharmaceutical composition to the subject. The pharmaceutical composition can be administered in any useful format. In various embodiments, the administering comprises at least one of intradermal, intramuscular, intraperitoneal, intravenous, subcutaneous, intranasal, epidural, oral, sublingual, intracerebral, intravaginal, transdermal, rectal, by inhalation, topical administration, or any combination thereof. The carrier or diluent can be any useful carrier or diluent, as described herein or known in the art.

[0039] The first target of the multipartite construct may be determined in a personalized manner. See, e.g., FIG. 23E and related discussion herein. For example, the first target of the multipartite construct can be determined by: a) contacting a biological sample from the subject with an oligonucleotide probe library such as described herein (see, e.g., Examples 16-18, 21-22, and 26-32); b) identifying members of the probe library that bound to the biological sample upon contact in a) (see, e.g., Examples 16-18, 21-22, and 26-32); and c) determining the first target as a binding target of a member of the probe library identified in b) (see, e.g., Examples 9, 23 and 35), thereby providing a personalized medical treatment to the subject. The biological sample may comprise any useful sample from the subject, including without limitation a biopsy sample, a tissue removed during a surgical procedure, a bodily fluid, a frozen section, a cell culture, or a fraction, derivative or components of any thereof. The bodily fluid can be any useful fluid from the subject, including without peripheral blood, sera, plasma, ascites, urine, cerebrospinal fluid (CSF), sputum, saliva, bone marrow, synovial fluid, aqueous humor, amniotic fluid, cerumen, breast milk, bronchial/alveolar lavage fluid, semen (including prostatic fluid), Cowper’s fluid or pre-ejaculatory fluid, female ejaculate, sweat, fecal matter, tears, cyst fluid, pleural fluid, 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.

[0040] A method of inducing killing of a cell, comprising contacting the cell with a multipartite construct of the invention as provided above. In such a construct, the second segment may comprise an oligonucleotide provided by the invention. For example, the construct may one that binds a member of the complement system, including without limitation the C1 complex, Clq, or a subunit thereof, and can induce complement mediated lysis of the cell. In some embodiments, the cell that is killed comprises a disease or disorder.

[0041] As described above, the multipartite construct of the invention can be administered to a subject to treat a disease or disorder, or to induce cell killing. In various embodiments, the disease or disorder comprises a cancer, a premalignant condition, an inflammatory disease, an immune disease, an
autoimmune disease or disorder, a cardiovascular disease or disorder, neurological disease or disorder, infectious disease or pain. The cancer can include without limitation one of 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, medullopithelioma, 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; extrhepatic bile duct cancer; gallbladder cancer; gastric (stomach) cancer; gastrointestinal carcinoid tumor; gastrointestinal stromal tumor 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; lung cancer; malignant fibrous histiocytoma bone cancer; medulloblastoma; medullopithelioma; 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; Sezary 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; Waldenstrom macroglobulinemia; or Wilm's tumor. The premalignant
condition can include without limitation Barrett's Esophagus. The autoimmune disease can include without limitation one of 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 erythematosus (SLE), Hashimoto's Thyroiditis, Grave's disease, Ankylosing Spondylitis Sjogrens Disease, CREST syndrome, Scleroderma, Rheumatic Disease, organ rejection, Primary Sclerosing Cholangitis, or sepsis. The cardiovascular disease can include without limitation one of atherosclerosis, congestive heart failure, vulnerable plaque, stroke, ischemia, high blood pressure, stenosis, vessel occlusion or a thrombotic event. The neurological disease can include without limitation one of Multiple Sclerosis (MS), Parkinson's Disease (PD), Alzheimer's Disease (AD), schizophrenia, bipolar disorder, depression, autism, Prion Disease, Pick's disease, dementia, Huntington disease (HD), Down's syndrome, cerebrovascular disease, Rasmussen's encephalitis, viral meningitis, neuropsychiatric systemic lupus erythematosus (NPSLE), amyotrophic lateral sclerosis, Creutzfeldt-Jacob disease, Gerstmann-Straussler-Scheinker disease, transmissible spongiform encephalopathy, ischemic reperfusion damage (e.g. stroke), brain trauma, microbial infection, or chronic fatigue syndrome. The pain can include without limitation one of fibromyalgia, chronic neuropathic pain, or peripheral neuropathic pain. The infectious disease can include without limitation one of a bacterial infection, viral infection, yeast infection, Whipple's Disease, Prion Disease, cirrhosis, methicillin-resistant staphylococcus aureus, HIV, HCV, hepatitis, syphilis, meningitis, malaria, tuberculosis, or influenza.

[0042] In an aspect, the invention provides a kit comprising a multipartite construct or a pharmaceutical composition provided by the invention, e.g., as described above. In a related aspect, the invention provides a kit comprising a reagent for carrying out the methods making use of such multipartite construct or a pharmaceutical composition. Similarly, the invention provides use of a reagent for carrying out the methods making use of such multipartite construct or a pharmaceutical composition. The invention also provides use of a reagent for the manufacture of a kit or reagent for carrying out the methods making use of such multipartite construct or a pharmaceutical composition. The invention further contemplates use of a reagent for the manufacture of a medicament for carrying out the method methods making use of such multipartite construct or a pharmaceutical composition. In such kits or uses, the reagent may comprise a multipartite construct or a pharmaceutical composition provided by the invention, e.g., as described above.

[0043] In an aspect, the invention provides a method of detecting at least one microvesicle in a biological sample comprising contacting the biological sample with at least one binding agent to at least one microvesicle surface antigen and detecting the at least one microvesicle recognized by the binding agent to the at least one protein. In an embodiment, the at least one microvesicle surface antigen is selected from any one of Tables 18, 26 or 45. The at least one protein can be a complement protein or an immunoglobulin. For example, the complement protein can be Clq, a subunit thereof (e.g., A, B or C), C4b, or any combination thereof. The immunoglobulin may comprise IgM, IgG, or both. See Example 35 herein. The at least one binding agent may comprise any useful binding agent, including without
limitation a nucleic acid, DNA molecule, RNA molecule, antibody, antibody fragment, aptamer, peptoid, zDNA, peptide nucleic acid (PNA), locked nucleic acid (LNA), lectin, peptide, dendrimer, membrane protein labeling agent, chemical compound, or a combination thereof. In some embodiments, the at least one binding agent comprises at least one oligonucleotide, such as an oligonucleotide probe as provided by the invention.

[0044] The at least one binding agent can be used to capture and/or detect the at least one microvesicle. Methods of detecting biomarkers and microvesicle using binding agents are provided herein. See, e.g., FIGs. 2A-B, which figures describe sandwich assay formats. In some embodiments, the at least one binding agent used to capture the at least one microvesicle is bound to a substrate. Any useful substrate can be used, including without limitation a planar array, a column matrix, or a microbead. See, e.g., FIGs. 2A-B. In some embodiments, the at least one binding agent used to detect the at least one microvesicle is labeled. Various useful labels are provided herein or known in the art, including 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, a semiconductor nanocrystal, a nanoparticle, a quantum dot, a gold particle, a fluorophore, or a radioactive label.

[0045] In an embodiment, the detecting is used to characterize a phenotype. The phenotype can be any appropriate phenotype of interest. In some embodiments, the phenotype is a disease or disorder. The characterizing may comprise providing, or assisting in providing, diagnostic, prognostic and/or theranostic information for the disease or disorder. The characterizing may be performed by comparing a presence or level of the at least one microvesicle to a reference. The reference can be selected per the characterizing to be performed. For example, when the phenotype comprises a disease or disorder, the reference may comprise a presence or level of the at least one microvesicle in a sample from an individual or group of individuals without the disease or disorder. The comparing can be determining whether the presence or level of the microvesicle differs from that of the reference. In some embodiments, the detected at least one microvesicle is found at higher levels in a healthy sample as compared to a diseased sample. In another embodiment, the detected at least one microvesicle is found at higher levels in a diseased sample as compared to a healthy sample. When multiplex assays are performed, e.g., using a plurality of binding agents to different biomarkers, some microvesicle antigens may be observed at a higher level in the biological samples as compared to the reference whereas other microvesicle antigens may be observed at a lower level in the biological samples as compared to the reference.

[0046] The method can be used to detect the at least one microvesicle in any appropriate biological sample. For example, the biological sample may comprise a bodily fluid, tissue sample or cell culture. The bodily fluid or tissue sample can be from a subject having or suspected of having a medical condition, a disease or a disorder. Thus, the method can be used to provide a diagnostic, prognostic, or theranostic read out for the subject. Any appropriate bodily fluid can be used, including without limitation peripheral blood, sera, plasma, ascites, urine, cerebrospinal fluid (CSF), sputum, saliva, bone marrow, synovial fluid, aqueous humor, amniotic fluid, cerumen, breast milk, bronchoalveolar lavage fluid, semen, prostatic
fluid, Cowper's fluid or pre-ejaculatory fluid, female ejaculate, sweat, fecal matter, hair oil, 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, blastocyst cavity fluid, or umbilical cord blood.

[0047] The method of the invention can be used to detect or characterize any appropriate disease or disorder of interest, including without limitation Breast Cancer, Alzheimer's disease, bronchial asthma, Transitional cell carcinoma of the bladder, Giant cellular osteoblastoclastoma, Brain Tumor, Colorectal adenocarcinoma, Chronic obstructive pulmonary disease (COPD), Squamous cell carcinoma of the cervix, acute myocardial infarction (AMI) / acute heart failure, Chron's Disease, diabetes mellitus type II, Esophageal carcinoma, Squamous cell carcinoma of the larynx, Acute and chronic leukemia of the bone marrow, Lung carcinoma, Malignant lymphoma, Multiple Sclerosis, Ovarian carcinoma, Parkinson disease, Prostate adenocarcinoma, psoriasis, Rheumatoid Arthritis, Renal cell carcinoma, Squamous cell carcinoma of skin, Adenocarcinoma of the stomach, carcinoma of the thyroid gland, Testicular cancer, ulcerative colitis, or Uterine adenocarcinoma.

[0048] In some embodiments, the disease or disorder comprises a cancer, a premalignant condition, an inflammatory disease, an immune disease, an autoimmune disease or disorder, a cardiovascular disease or disorder, neurological disease or disorder, infectious disease or pain. The cancer can include without limitation one of 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, cranioaphyngioma, 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; cranioaphyngioma; 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; lung 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; Sezary 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; Waldenstrom macroglobulinemia; or Wiln's tumor. The premalignant condition can include without limitation Barrett's Esophagus. The autoimmune disease can include without limitation one of 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. The cardiovascular disease can include without limitation one of atherosclerosis, congestive heart failure, vulnerable plaque, stroke, ischemia, high blood pressure, stenosis, vessel occlusion or a thrombotic event. The neurological disease can include without limitation one of Multiple Sclerosis (MS), Parkinson's Disease (PD), Alzheimer's Disease (AD), schizophrenia, bipolar disorder, depression, autism, Prion Disease, Pick's disease, dementia, Huntington disease (HD), Down's syndrome, cerebrovascular disease, Rasmussen's encephalitis, viral meningitis, neuropsychiatric systemic lupus erythematosus (NPSLE), amyotrophic lateral sclerosis, Creutzfeldt-Jacob disease, Gerstmann-Straussler-Scheinker disease, transmissible spongiform encephalopathy, ischemic reperfusion damage (e.g. stroke), brain trauma, microbial infection, or chronic fatigue syndrome. The pain can include without limitation one of fibromyalgia, chronic neuropathic pain, or peripheral neuropathic pain. The infectious disease can include without limitation one of a bacterial infection, viral infection, yeast infection, Whipple's Disease, Prion Disease, cirrhosis, methicillin-resistant staphylococcus aureus, HIV, HCV, hepatitis, syphilis, meningitis, malaria, tuberculosis, or influenza.

[0049] In a related aspect, the invention provides a kit comprising a reagent for carrying out the methods herein. In still another related aspect, the invention provides for use of a reagent for carrying out the methods. The reagent may comprise at least one binding agent to the at least one protein. The binding
agent may be an oligonucleotide as provided herein, one or more primer for amplification or sequencing of such oligonucleotide, a binding buffer with or without MgC\textsuperscript{2+}, a sample processing reagent, a microvesicle isolation reagent, a detection reagent, a secondary detection reagent, a wash buffer, an elution buffer, a solid support, and any combination thereof. In some embodiments, the microvesicle isolation reagent comprises at least one of a concentrator unit, a filtration unit, a polymer, PEG, a size exclusion column, a binding agent to a microvesicle antigen, and any combination thereof; and/or the detection or secondary detection agent comprises streptavidin-horse radish peroxide (HRP), a streptavidin-conjugated fluorophore, a streptavidin-conjugated quantum dot, and any combination thereof.

**BRIEF DESCRIPTION OF THE DRAWINGS**

[0050] FIG. 1 illustrates a competitive assay selection strategy: the random pool of aptamer (the library) is incubated with the target protein, in this case, EpCAM. After washing and elution from the target, the eluted aptamers are again added to the target and allowed to bind. The antibody is then added to the reaction, competing with the aptamers at the epitope of the antibody. The aptamers displaced by the antibody are then collected.

[0051] FIGs. 2A-2F illustrate methods of assessing biomarkers such as microvesicle surface antigens. FIG. 2A is a schematic of a planar substrate coated with a capture agent, such as an aptamer or antibody, which captures vesicles expressing the target antigen of the capture agent. The capture agent may bind a protein expressed on the surface of vesicles shed from diseased cells ("disease vesicle"). The detection agent, which may also be an aptamer or antibody, carries a detectable label, here a fluorescent signal. The detection agent binds to the captured vesicle and provides a detectable signal via its fluorescent label. The detection agent 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 particle bead conjugated with a capture agent, which captures vesicles expressing the target antigen of the capture agent. The capture agent may bind a protein expressed on the surface of vesicles shed from diseased cells ("disease vesicle"). The detection agent, which may also be an aptamer or antibody, carries a detectable label, here a fluorescent signal. The detection agent binds to the captured vesicle and provides a detectable signal via its fluorescent label. The detection agent 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 using different combinations of capture and detection agents to the indicated biomarkers. The biomarker combinations can be detected using assays as shown in FIGs. 2A-2B. FIGs. 2D-2E present illustrative schemes for capturing and detecting vesicles to characterize a phenotype. FIG. 2F presents illustrative schemes for assessing vesicle payload to characterize a phenotype.

[0052] FIGs. 3A-B illustrates a non-limiting example of an aptamer nucleotide sequence and its secondary structure. FIG. 3A illustrates a secondary structure of a 32-mer oligonucleotide, Aptamer 4, with sequence 5'-CCCCTGAATCAGATGGGCGGGGCTCG (SEQ ID NO. 1). In the figure, the sequence is shown with 6 thymine nucleotides added to the end, which can act as a spacer to attach a biotin molecule. This particular oligo has a high binding affinity to the target, EpCAM (see Table 5). Additional candidate EpCAM binders are identified by modeling the entire database of sequenced oligos.
to the secondary structure of this oligo. FIG. 3B illustrates another 32-mer oligo with sequence 5'-ACCGGATACGGTTGAGGCGTGCTCCACTCG (SEQ ID NO. 2) that has a different secondary structure than the aptamer in FIG. 3A. This aptamer is also shown with a 6-thymine tail.

[0053] FIG. 4 illustrates a process for producing a target-specific set of aptamers using a cell subtraction method, wherein the target is a biomarker associated with a specific disease. In Step 1, a random pool of oligonucleotides are contacted with a biological sample from a normal patient. In Step 2, the oligos that did not bind in Step 1 are added to a biological sample isolated from diseased patients. The bound oligos from this step are then eluted, captured via their biotin linkage and then combined again with normal biological sample. The unbound oligos are then added again to disease-derived biological sample and isolated. This process can be repeated iteratively. The final eluted aptamers are tested against patient samples to measure the sensitivity and specificity of the set. Biological samples can include blood, including plasma or serum, or other components of the circulatory system, such as microvesicles.

[0054] FIG. 5 illustrates results from a binding assay showing the binding affinity of an exemplary aptamer (Aptamer ID BTX176881 (SEQ ID NO: 3)) to the target EpCAM protein at various target concentrations. The aptamer to be tested is fixed to a substrate using a biotin tail and is incubated with various concentrations of target (125, 250 and 500 nM). The test is performed on a surface plasmon resonance machine (SPR). The SPR machine detects association and disassociation of the aptamer and the target. Target is applied until the association and disassociation events are equal, resulting in a plateau of the curve. The equations describing the curve at each concentration can then be used to calculate the $K_D$ of the aptamer (see Table 5).

[0055] FIGs. 6A-D illustrate the use of an anti-EpCAM aptamer (Aptamer 4; SEQ ID NO. 1) to detect a microvesicle population. Vesicles in patient plasma samples were captured using bead-conjugated antibodies to the indicated microvesicle surface antigens: A) EGFR; B) PBP; C) EpCAM; D) KLK2. Fluorescently labeled Aptamer 4 was used as a detector in the microbead assay. The figure shows average median fluorescence values (MFI values) for three cancer (C1-C3) and three normal samples (N1-N3) in each plot. In each plot, the samples from left to right are ordered as: C1, C2, C3, N1, N2, N3.

[0056] FIG. 7A illustrates the sequence of EPCAM aptamer CAR003 (SEQ ID NO. 4). FIG. 7B illustrates the optimal secondary structure of CAR003 with a minimum free energy (AG) of -30.00 kcal/mol. For purposes of illustration, the aptamer is shown as an RNA aptamer (SEQ ID NO. 5) corresponding to the DNA sequence in FIG. 7A. FIG. 7C illustrates aptamer pool purification. The figure comprises an FPLC chromatogram with all product and fractions assigned in pools after checking quality on gel. FIG. 7D illustrates a SYBR GOLD stained gel with different FPLC fractions of CAR003 aptamer after synthesis. Different fractions were combined in pools based on amount of un-finished chains in order high to low (pool 1 - pool 3). The pools 1-3 correspond to those indicated in FIG. 7C. FIG. 7E-F illustrate binding of CAR003 to EPCAM protein in 25 mM HEPES with PBS-BN (FIG. 7E) or in 25 mM HEPES with 1 mM MgCl$_2$ (FIG. 7F). FIG. 7G illustrates CAR003 binding to EpCAM in the indicated salts with and without addition of bovine serum albumin (BSA). FIG. 7H illustrates the effect of denaturing on CAR003 binding to EPCAM protein. In each group of four bars, the aptamer is from left to
right: Aptamer 4, CAR003 Pool 1, CAR003 Pool 2, CAR003 Pool 3. **FIG. 71** illustrates titration of aptamers against EPCAM recombinant protein (constant input 5 µg). **FIG. 7J** illustrates a Western blot with CAR003 aptamer versus EPCAM his-tagged protein, BSA, and HSA (5 µg each). The gel was blocked 0.5% F127 and probed with ~ 50 µg/ml CAR003 biotinylated aptamer, fraction 3. The blot was visualized with NeutrAvidin-HRP followed by SuperSignal West Femto Chemiluminescent Substrate.

**[0057]** FIGs. 8A-8D illustrates methods to attach microvesicles to a substrate. **FIG. 8A** illustrates direct conjugation of a carboxylated microsphere to a vesicle surface antigen. **FIG. 8B** illustrates anchoring of a microvesicle to a microsphere via a biotin functionalized lipid anchor. **FIG. 8C** illustrates antibody binding to a vesicle surface antigen, wherein the antibody is conjugated to a carboxylated microsphere. **FIG. 8D** illustrates aptamer binding to a vesicle surface antigen, wherein the aptamer is conjugated to a carboxylated microsphere.

**[0058]** **FIG. 9** comprises a schematic for identifying a target of a selected aptamer, such as an aptamer selected by the process of the invention. The figure shows a binding agent 902, here an aptamer for purposes of illustration, tethered to a substrate 901. The binding agent 902 can be covalently attached to substrate 901. The binding agent 902 may also be non-covalently attached. For example, binding agent 902 can comprise a label which can be attracted to the substrate, such as a biotin group which can form a complex with an avidin/streptavidin molecule that is covalently attached to the substrate. The binding agent 902 binds to a surface antigen 903 of microvesicle 904. In the step signified by arrow (i), the microvesicle is disrupted while leaving the complex between the binding agent 902 and surface antigen 903 intact. Disrupted microvesicle 905 is removed, e.g., via washing or buffer exchange, in the step signified by arrow (ii). In the step signified by arrow (iii), the surface antigen 903 is released from the binding agent 902. The surface antigen 903 can be analyzed to determine its identity.

**[0059]** FIGs. 10A-1OC illustrate binding of selected aptamers against microbeads conjugated to various input sample. The aptamers were selected from an aptamer library as binding to microbeads conjugated to breast cancer-derived microvesicles. Experimental details are in the Examples herein. Each plot shows a different aptamer. The Y-axis indicates level of binding. In each group of samples, binding of 9 purified aptamer candidates is shown. The input sample is indicated on the X axis from left to right as follows: 1) Cancer Exosome: aptamer binding to microbeads conjugated to microvesicles isolated from plasma samples from breast cancer patients; 2) Cancer Non-exosome: aptamer binding to microbeads conjugated to plasma samples from breast cancer patients after removal of microvesicles by ultracentrifugation; 3) Non-Cancer Exosome: aptamer binding to microbeads conjugated to microvesicles isolated from plasma samples from normal (i.e., non-breast cancer) patients; 4) Non-Cancer Non-Exosome: aptamer binding to microbeads conjugated to plasma samples from breast cancer patients after removal of microvesicles by ultracentrifugation.

**[0060]** FIGs. 11A-11B illustrate enriching a naive aptamer library for aptamers that differentiate between breast cancer and non-cancer microvesicles in plasma samples.

**[0061]** FIGs. 12A-12G illustrate using an oligonucleotide probe library to differentiate cancer and non-cancer samples.
FIG. 13 shows protein targets of oligonucleotide probes run on a silver stained SDS-PAGE gel.

FIGs. 14A-G illustrate use of oligonucleotides that differentiate microvesicles in breast cancer plasma from normal controls.

FIG. 15 shows a heatmap of clusters of oligonucleotides enriched against aggressive versus non-aggressive breast cancer plams samples.

FIGs. 16A-B illustrate a model generated using a training (FIG. 16A) and test (FIG. 16B) set from a round of cross validation. The AUC for the test set was 0.803. Another exemplary round of cross-validation is shown in FIGs. 16C-D with training (FIG. 16C) and test (FIG. 16D) sets. The AUC for the test set was 0.678.

FIGs. 17A-E show a photo-cleavable Biotin mediated purification of an oligonucleotide library.

FIGs. 18A-C illustrate SUPRA (SsDNA by Unequal length PRimer Asymmetric PCR), a protocol for single stranded DNA (ssDNA) oligonucleotide library preparation.

FIGs. 19A-C illustrate use of aptamers in methods of characterizing a phenotype. FIG. 19A is a schematic 1900 showing an assay configuration that can be used to detect and/or quantify a target of interest. In the figure, capture aptamer 1902 is attached to substrate 1901. Target of interest 1903 is bound by capture aptamer 1902. Detection aptamer 1904 is also bound to target of interest 1903. Detection aptamer 1904 carries label 1905 which can be detected to identify target captured to substrate 1901 via capture aptamer 1902. FIG. 19B is a schematic 1910 showing use of an aptamer pool to characterize a phenotype. A pool of aptamers to a target of interest is provided 1911. The pool is contacted with a test sample to be characterized 1912. The mixture is washed to remove unbound aptamers. The remaining aptamers are disassociated and collected 1913. The collected aptamers are identified 1914 and the identity of the retained aptamers is used to characterize the phenotype 1915. FIG. 19C is a schematic 1920 showing an implementation of the method in FIG. 19B. A pool of aptamers identified as binding a microvesicle population is provided 1919. The input sample comprises microvesicles that are isolated from a test sample 1920. The pool is contacted with the isolated microvesicles to be characterized 1923. The mixture is washed to remove unbound aptamers and the remaining aptamers are disassociated and collected 1925. The collected aptamers are identified and the identity of the retained aptamers is used to characterize the phenotype 1926.

FIGs. 20A-I illustrate development and use of an oligonucleotide probe library to distinguish biological sample types.

FIGs. 21A-C illustrate enriching a naive oligonucleotide library with balanced design for oligonucleotides that differentiate between breast cancer and non-cancer microvesicles derived from plasma samples.

FIGs. 22A-D shows characterization of breast cancer samples as cancer or non-cancer using two different but related oligonucleotide probe libraries.

FIGs. 23A-E illustrate oligonucleotide constructs that recognize immunomodulatory (IMD) targets.

FIGs. 24A-C illustrate identification of oligonucleotides that recognize Clq and other targets.
DETAILED DESCRIPTION OF THE INVENTION

[0074] The details of one or more embodiments of the invention are set forth in the accompanying description below. Although any methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, the preferred methods and materials are now described. Other features, objects, and advantages of the invention will be apparent from the description. In the specification, the singular forms also include the plural unless the context clearly dictates otherwise. Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. In the case of conflict, the present Specification will control.

[0075] Disclosed herein are compositions and methods that can be used to assess a biomarker profile, which can include a presence or level of one or more biomarkers. The compositions and methods of the invention comprise the use of oligonucleotide probes (aptamers) that bind microvesicle surface antigens or a functional fragment thereof. The antigens typically comprise proteins or polypeptides but can be any useful component displayed on a microvesicle surface including nucleic acids, lipids and/or carbohydrates. In general, the oligonucleotides disclosed are synthetic nucleic acid molecules, including DNA and RNA, and variations thereof. Unless otherwise specified, the oligonucleotide probes can be synthesized in DNA or RNA format or as hybrid molecules as desired. The methods disclosed comprise diagnostic processes and techniques using one or more aptamer of the invention, to determine the level or presence of relevant microvesicle surface antigens or a functional fragment thereof. Alternatively, an oligonucleotide probe of the invention can also be used as a binding agent to capture, isolate, or enrich, a cell, cell fragment, vesicle or any other fragment or complex that comprises the antigen or functional fragments thereof.

[0076] The compositions and methods of the invention comprise individual oligonucleotides that are identified for use in assessing a biomarker profile. The invention further discloses compositions and methods of oligonucleotide pools that can be used to detect a biomarker profile in a given sample.

[0077] Oligonucleotide probes and sequences disclosed in the compositions and methods of the invention may be identified herein in the form of DNA or RNA. Unless otherwise specified, one of skill in the art will appreciate that an oligonucleotide may generally be synthesized as either form of nucleic acid and carry various chemical modifications and remain within the scope of the invention. The term aptamer may be used in the art to refer to a single oligonucleotide that binds specifically to a target of interest through mechanisms other than Watson crick base pairing, similar to binding of a monoclonal antibody to a particular antigen. Within the scope of this disclosure and unless stated explicitly or otherwise implicit in context, the terms aptamer, oligonucleotide and oligonucleotide probe, and variations thereof, may be used interchangeably to refer to an oligonucleotide capable of distinguishing biological entities of interest (e.g., biomarkers) whether or not the specific entity has been identified or whether the precise mode of binding has been determined.
An oligonucleotide probe of the invention can also be used to provide *in vitro* or *in vivo* detection or imaging, to provide any appropriate diagnostic readout (e.g., diagnostic, prognostic or theranostic).

Separately, an oligonucleotide probe of the invention can also be used for treatment or as a therapeutic to specifically target a cell, tissue or organ.

**Aptamers**

SELEX. A suitable method for generating an aptamer is with the process entitled "Systematic Evolution of Ligands by Exponential Enrichment" ("SELEX") generally described in, e.g., U.S. patent application Ser. No. 07/536,428, filed Jun. 11, 1990, now abandoned, U.S. Pat. No. 5,475,096 entitled "Nucleic Acid Ligands", and U.S. Pat. No. 5,270,163 (see also WO 91/19813) entitled "Nucleic Acid Ligands". Each SELEX-identified nucleic acid ligand, i.e., each aptamer, is a specific ligand of a given target compound or molecule. The SELEX process is based on the unique insight that nucleic acids have sufficient capacity for forming a variety of two- and three-dimensional structures and sufficient chemical versatility available within their monomers to act as ligands (i.e., form specific binding pairs) with virtually any chemical compound, whether monomeric or polymeric. Molecules of any size or composition can serve as targets.

SELEX relies as a starting point upon a large library or pool of single stranded oligonucleotides comprising randomized sequences. The oligonucleotides can be modified or unmodified DNA, RNA, or DNA/RNA hybrids. In some examples, the pool comprises 100% random or partially random oligonucleotides. In other examples, the pool comprises random or partially random oligonucleotides containing at least one fixed and/or conserved sequence incorporated within randomized sequence. In other examples, the pool comprises random or partially random oligonucleotides containing at least one fixed and/or conserved sequence at its 5’ and/or 3’ end which may comprise a sequence shared by all the molecules of the oligonucleotide pool. Fixed sequences are sequences such as hybridization sites for PCR primers, promoter sequences for RNA polymerases (e.g., T3, T4, T7, and SP6), restriction sites, or homopolymeric sequences, such as poly A or poly T tracts, catalytic cores, sites for selective binding to affinity columns, and other sequences to facilitate cloning and/or sequencing of an oligonucleotide of interest. Conserved sequences are sequences, other than the previously described fixed sequences, shared by a number of aptamers that bind to the same target.

The oligonucleotides of the pool preferably include a randomized sequence portion as well as fixed sequences necessary for efficient amplification. Typically the oligonucleotides of the starting pool contain fixed 5’ and 3’ terminal sequences which flank an internal region of 30-50 random nucleotides. The randomized nucleotides can be produced in a number of ways including chemical synthesis and size selection from randomly cleaved cellular nucleic acids. Sequence variation in test nucleic acids can also be introduced or increased by mutagenesis before or during the selection/amplification iterations.

The random sequence portion of the oligonucleotide can be of any length and can comprise ribonucleotides and/or deoxyribonucleotides and can include modified or non-natural nucleotides or nucleotide analogs. See, e.g. U.S. Pat. No. 5,958,691; U.S. Pat. No. 5,660,985; U.S. Pat. No. 5,958,691; U.S. Pat. No. 5,698,687; U.S. Pat. No. 5,817,635; U.S. Pat. No. 5,672,695, and PCT Publication WO
92/07065. Random oligonucleotides can be synthesized from phosphodiester-linked nucleotides using solid phase oligonucleotide synthesis techniques well known in the art. See, e.g., Froehler et al., Nucl. Acid Res. 14:5399-5467 (1986) and Froehler et al., Tet. Lett. 27:5575-5578 (1986). Random oligonucleotides can also be synthesized using solution phase methods such as triester synthesis methods. See, e.g., Sood et al., Nucl. Acid Res. 4:2557 (1977) and Hirose et al., Tet. Lett., 28:2449 (1978). Typical syntheses carried out on automated DNA synthesis equipment yield $10^{14}-10^{16}$ individual molecules, a number sufficient for most SELEX experiments. Sufficiently large regions of random sequence in the sequence design increases the likelihood that each synthesized molecule is likely to represent a unique sequence.

[0084] The starting library of oligonucleotides may be generated by automated chemical synthesis on a DNA synthesizer. To synthesize randomized sequences, mixtures of all four nucleotides are added at each nucleotide addition step during the synthesis process, allowing for random incorporation of nucleotides. As stated above, in one embodiment, random oligonucleotides comprise entirely random sequences; however, in other embodiments, random oligonucleotides can comprise stretches of nonrandom or partially random sequences. Partially random sequences can be created by adding the four nucleotides in different molar ratios at each addition step.

[0085] The starting library of oligonucleotides may be for example, RNA, DNA, or RNA/DNA hybrid. In those instances where an RNA library is to be used as the starting library it is typically generated by transcribing a DNA library in vitro using T7 RNA polymerase or modified T7 RNA polymerases and purified. The library is then mixed with the target under conditions favorable for binding and subjected to step-wise iterations of binding, partitioning and amplification, using the same general selection scheme, to achieve virtually any desired criterion of binding affinity and selectivity. More specifically, starting with a mixture containing the starting pool of nucleic acids, the SELEX method includes steps of: (a) contacting the mixture with the target under conditions favorable for binding; (b) partitioning unbound nucleic acids from those nucleic acids which have bound specifically to target molecules; (c) dissociating the nucleic acid-target complexes; (d) amplifying the nucleic acids dissociated from the nucleic acid-target complexes to yield a ligand-enriched mixture of nucleic acids; and (e) reiterating the steps of binding, partitioning, dissociating and amplifying through as many cycles as desired to yield highly specific, high affinity nucleic acid ligands to the target molecule. In those instances where RNA aptamers are being selected, the SELEX method further comprises the steps of: (i) reverse transcribing the nucleic acids dissociated from the nucleic acid-target complexes before amplification in step (d); and (ii) transcribing the amplified nucleic acids from step (d) before restarting the process.

[0086] Within a nucleic acid mixture containing a large number of possible sequences and structures, there is a wide range of binding affinities for a given target. A nucleic acid mixture comprising, for example, a 20 nucleotide randomized segment can have $4^{20}$ candidate possibilities. Those which have the higher affinity constants for the target are most likely to bind to the target. After partitioning, dissociation and amplification, a second nucleic acid mixture is generated, enriched for the higher binding affinity candidates. Additional rounds of selection progressively favor better ligands until the resulting nucleic
acid mixture is predominantly composed of only one or a few sequences. These can then be cloned, sequenced and individually tested for binding affinity as pure ligands or aptamers.

[0087] Cycles of selection and amplification are repeated until a desired goal is achieved. In the most general case, selection/amplification is continued until no significant improvement in binding strength is achieved on repetition of the cycle. The method is typically used to sample approximately \(10^{14}\) different nucleic acid species but may be used to sample as many as about \(10^{18}\) different nucleic acid species. Generally, nucleic acid aptamer molecules are selected in a 5 to 20 cycle procedure. In one embodiment, heterogeneity is introduced only in the initial selection stages and does not occur throughout the replicating process.

[0088] In one embodiment of SELEX, the selection process is so efficient at isolating those nucleic acid ligands that bind most strongly to the selected target, that only one cycle of selection and amplification is required. Such an efficient selection may occur, for example, in a chromatographic-type process wherein the ability of nucleic acids to associate with targets bound on a column operates in such a manner that the column is sufficiently able to allow separation and isolation of the highest affinity nucleic acid ligands.

[0089] In many cases, it is not necessarily desirable to perform the iterative steps of SELEX until a single nucleic acid ligand is identified. The target-specific nucleic acid ligand solution may include a family of nucleic acid structures or motifs that have a number of conserved sequences and a number of sequences which can be substituted or added without significantly affecting the affinity of the nucleic acid ligands to the target. By terminating the SELEX process prior to completion, it is possible to determine the sequence of a number of members of the nucleic acid ligand solution family. The invention provides for the identification of aptamer pools and uses thereof that jointly can be used to characterize a test sample. For example, the aptamer pools can be identified through rounds of positive and negative selection to identify microvesicle indicative of a disease or condition. The invention further provides use of such aptamer pools to detect and/or quantify such microvesicles in a sample, thereby allowing a diagnosis, prognosis or theranosis to be provided.

[0090] A variety of nucleic acid primary, secondary and tertiary structures are known to exist. The structures or motifs that have been shown most commonly to be involved in non-Watson-Crick type interactions are referred to as hairpin loops, symmetric and asymmetric bulges, pseudoknots and myriad combinations of the same. Almost all known cases of such motifs suggest that they can be formed in a nucleic acid sequence of no more than 30 nucleotides. For this reason, it is often preferred that SELEX procedures with contiguous randomized segments be initiated with nucleic acid sequences containing a randomized segment of between about 20 to about 50 nucleotides and in some embodiments, about 30 to about 40 nucleotides. In one example, the 5'-fixed:random:3'-fixed sequence comprises a random sequence of about 30 to about 50 nucleotides.

[0091] The core SELEX method has been modified to achieve a number of specific objectives. For example, U.S. Pat. No. 5,707,796 describes the use of SELEX in conjunction with gel electrophoresis to select nucleic acid molecules with specific structural characteristics, such as bent DNA. U.S. Pat. No. 5,763,177 describes SELEX based methods for selecting nucleic acid ligands containing photoreactive
groups capable of binding and/or photocrosslinking to and/or photoinactivating a target molecule. U.S. Pat. No. 5,567,588 and U.S. Pat. No. 5,861,254 describe SELEX based methods which achieve highly efficient partitioning between oligonucleotides having high and low affinity for a target molecule. U.S. Pat. No. 5,496,938 describes methods for obtaining improved nucleic acid ligands after the SELEX process has been performed. U.S. Pat. No. 5,705,337 describes methods for covalently linking a ligand to its target.

SELEX can also be used to obtain nucleic acid ligands that bind to more than one site on the target molecule, and to obtain nucleic acid ligands that include non-nucleic acid species that bind to specific sites on the target. SELEX provides means for isolating and identifying nucleic acid ligands which bind to any envisionable target, including large and small biomolecules such as nucleic acid-binding proteins and proteins not known to bind nucleic acids as part of their biological function as well as lipids, cofactors and other small molecules. For example, U.S. Pat. No. 5,580,737 discloses nucleic acid sequences identified through SELEX which are capable of binding with high affinity to caffeine and the closely related analog, theophylline.

Counter-SELEX is a method for improving the specificity of nucleic acid ligands to a target molecule by eliminating nucleic acid ligand sequences with cross-reactivity to one or more non-target molecules. Counter-SELEX is comprised of the steps of: (a) preparing a candidate mixture of nucleic acids; (b) contacting the candidate mixture with the target, wherein nucleic acids having an increased affinity to the target relative to the candidate mixture may be partitioned from the remainder of the candidate mixture; (c) partitioning the increased affinity nucleic acids from the remainder of the candidate mixture; (d) dissociating the increased affinity nucleic acids from the target; e) contacting the increased affinity nucleic acids with one or more non-target molecules such that nucleic acid ligands with specific affinity for the non-target molecule(s) are removed; and (f) amplifying the nucleic acids with specific affinity only to the target molecule to yield a mixture of nucleic acids enriched for nucleic acid sequences with a relatively higher affinity and specificity for binding to the target molecule. As described above for SELEX, cycles of selection and amplification are repeated until a desired goal is achieved.

One potential problem encountered in the use of nucleic acids as therapeutics and vaccines is that oligonucleotides in their phosphodiester form may be quickly degraded in body fluids by intracellular and extracellular enzymes such as endonucleases and exonucleases before the desired effect is manifest. The SELEX method thus encompasses the identification of high-affinity nucleic acid ligands containing modified nucleotides conferring improved characteristics on the ligand, such as improved in vivo stability or improved delivery characteristics. Examples of such modifications include chemical substitutions at the ribose and/or phosphate and/or base positions. SELEX identified nucleic acid ligands containing modified nucleotides are described, e.g., in U.S. Pat. No. 5,660,985, which describes oligonucleotides containing nucleotide derivatives chemically modified at the 2' position of ribose, 5' position of pyrimidines, and 8' position of purines, U.S. Pat. No. 5,756,703 which describes oligonucleotides containing various 2'-modified pyrimidines, and U.S. Pat. No. 5,580,737 which describes highly specific nucleic acid ligands.
containing one or more nucleotides modified with 2'-amino (2'~NH₂), 2'-fluoro (2'-F), and/or 2'-O-methyl (2'-OMe) substituents.

[0095] Modifications of the nucleic acid ligands contemplated in this invention include, but are not limited to, those which provide other chemical groups that incorporate additional charge, polarizability, hydrophobicity, hydrogen bonding, electrostatic interaction, and fluxionality to the nucleic acid ligand bases or to the nucleic acid ligand as a whole. Modifications to generate oligonucleotide populations which are resistant to nucleases can also include one or more substitute internucleotide linkages, altered sugars, altered bases, or combinations thereof. Such modifications include, but are not limited to, 2'-position sugar modifications, 5-position pyrimidine modifications, 8-position purine modifications, modifications at exocyclic amines, substitution of 4-thiouridine, substitution of 5-bromo or 5-ido-uracil; backbone modifications, phosphorothioate or allyl phosphate modifications, methylations, and unusual base-pairing combinations such as the isobases isocytidine and isoguanosine. Modifications can also include 3' and 5' modifications such as capping.

[0096] In one embodiment, oligonucleotides are provided in which the P(O)O group is replaced by P(O)S ("thioate"), P(S)S ("dithioate"), P(O)NR₂ ("amidate"), P(O)R, P(O)OR, CO or CH₂ ("formacetal") or 3'-amine (--NH--CH₂--CH₃--), wherein each R or R is independently H or substituted or unsubstituted alkyl. Linkage groups can be attached to adjacent nucleotides through an --0-, --N-, or --S- linkage. Not all linkages in the oligonucleotide are required to be identical. As used herein, the term phosphorothioate encompasses one or more non-bridging oxygen atoms in a phosphodiester bond replaced by one or more sulfur atoms.

[0097] In further embodiments, the oligonucleotides comprise modified sugar groups, for example, one or more of the hydroxylic groups is replaced with halogen, aliphatic groups, or functionalized as ethers or amines. In one embodiment, the 2'-position of the furanose residue is substituted by any of an O-methyl, O-alkyl, O-allyl, S-alkyl, S-allyl, or halo group. Methods of synthesis of 2'-modified sugars are described, e.g., in Sproat, et al., Nucl. Acid Res. 19:733-738 (1991); Cotton, et al., Nucl. Acid Res. 19:2629-2635 (1991); and Hobbs, et al., Biochemistry 12:5138-5145 (1973). Other modifications are known to one of ordinary skill in the art. Such modifications may be pre-SELEX process modifications or post-SELEX process modifications (modification of previously identified unmodified ligands) or may be made by incorporation into the SELEX process.

[0098] Pre-SELEX process modifications or those made by incorporation into the SELEX process yield nucleic acid ligands with both specificity for their SELEX target and improved stability, e.g., in vivo stability. Post-SELEX process modifications made to nucleic acid ligands may result in improved stability, e.g., in vivo stability without adversely affecting the binding capacity of the nucleic acid ligand.

[0099] The SELEX method encompasses combining selected oligonucleotides with other selected oligonucleotides and non-oligonucleotide functional units as described in U.S. Pat. No. 5,637,459 and U.S. Pat. No. 5,683,867. The SELEX method further encompasses combining selected nucleic acid ligands with lipophilic or non-immunogenic high molecular weight compounds in a diagnostic or therapeutic complex, as described, e.g., in U.S. Pat. No. 6,011,020, U.S. Pat. No. 6,051,698, and PCT
The identification of nucleic acid ligands to small, flexible peptides via the SELEX method has also been explored. Small peptides have flexible structures and usually exist in solution in an equilibrium of multiple conformers, and thus it was initially thought that binding affinities may be limited by the conformational entropy lost upon binding a flexible peptide. However, the feasibility of identifying nucleic acid ligands to small peptides in solution was demonstrated in U.S. Pat. No. 5,648,214. In this patent, high affinity RNA nucleic acid ligands to substance P, an 11 amino acid peptide, were identified.

The aptamers with specificity and binding affinity to the target(s) of the present invention can be selected by the SELEX N process as described herein. As part of the SELEX process, the sequences selected to bind to the target are then optionally minimized to determine the minimal sequence having the desired binding affinity. The selected sequences and/or the minimized sequences are optionally optimized by performing random or directed mutagenesis of the sequence to increase binding affinity or alternatively to determine which positions in the sequence are essential for binding activity. Additionally, selections can be performed with sequences incorporating modified nucleotides to stabilize the aptamer molecules against degradation in vivo.

2' Modified SELEX

For an aptamer to be suitable for use as a therapeutic, it is preferably inexpensive to synthesize, and safe and stable in vivo. Wild-type RNA and DNA aptamers are typically not stable in vivo because of their susceptibility to degradation by nucleases. Resistance to nuclease degradation can be greatly increased by the incorporation of modifying groups at the 2'-position.

Fluoro and amino groups have been successfully incorporated into oligonucleotide pools from which aptamers have been subsequently selected. However, these modifications greatly increase the cost of synthesis of the resultant aptamer, and may introduce safety concerns in some cases because of the possibility that the modified nucleotides could be recycled into host DNA by degradation of the modified oligonucleotides and subsequent use of the nucleotides as substrates for DNA synthesis.

Aptamers that contain 2'-0-methyl ("2'-OMe") nucleotides, as provided herein, may overcome one or more potential drawbacks. Oligonucleotides containing 2'-OMe nucleotides are nuclease-resistant and inexpensive to synthesize. Although 2'-OMe nucleotides are ubiquitous in biological systems, natural polymerases do not accept 2'-OMe NTPs as substrates under physiological conditions, thus there are no safety concerns over the recycling of 2'-OMe nucleotides into host DNA. The SELEX method used to generate 2'-modified aptamers is described, e.g., in U.S. Provisional Patent Application Ser. No. 60/430,761, filed Dec. 3, 2002, U.S. Provisional Patent Application Ser. No. 60/487,474, filed Jul. 15, 2003, U.S. Provisional Patent Application Ser. No. 60/517,039, filed Nov. 4, 2003, U.S. patent application Ser. No. 10/729,581, filed Dec. 3, 2003, and U.S. patent application Ser. No. 10/873,856, filed Jun. 21, 2004, entitled "Method for in vitro Selection of 2'-0-methyl substituted Nucleic Acids", each of which is herein incorporated by reference in its entirety.
METHODS

Biomarker Detection and Diagnostics

[00107] The aptamers of the invention can be used in various methods to assess presence or level of biomarkers in a biological sample, e.g., biological entities of interest such as proteins, nucleic acids, or microvesicles. The aptamer functions as a binding agent to assess presence or level of the cognate target molecule. Therefore, in various embodiments of the invention directed to diagnostics, prognostics or theranostics, one or more aptamers of the invention are configured in a ligand-target based assay, where one or more aptamer of the invention is contacted with a selected biological sample, where the or more aptamer associates with or binds to its target molecules. Aptamers of the invention are used to identify candidate biosignatures based on the biological samples assessed and biomarkers detected. In further embodiments, aptamers may themselves provide a biosignature for a particular condition or disease. A biosignature refers to a biomarker profile of a biological sample comprising a presence, level or other characteristic that can be assessed (including without limitation a sequence, mutation, rearrangement, translocation, deletion, epigenetic modification, methylation, post-translational modification, allele, activity, complex partners, stability, half life, and the like) of one or more biomarker of interest. Biosignatures can be used to evaluate diagnostic and/or prognostic criteria such as presence of disease, disease staging, disease monitoring, disease stratification, or surveillance for detection, metastasis or recurrence or progression of disease. For example, methods of the invention using aptamers against microvesicle surface antigen are useful for correlating a biosignature comprising microvesicle antigens to a selected condition or 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 used along with surgery (e.g., either pre-surgery or post-surgery). As an illustrative example, a biosignature of circulating biomarkers 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.

[00108] A biosignature can be used in any methods disclosed herein, e.g., 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. For example, a biosignature can be used to assess whether a subject has prostate cancer, colon cancer, or other cancer as described herein. Furthermore, a biosignature can be used to determine a stage of a disease or condition, such as colon cancer. The biosignature/biomarker profile comprising a microvesicle can include assessment of payload within the microvesicle. For example, one or more aptamer of the invention can be used to capture a microvesicle population, thereby providing readout of microvesicle antigens, and then the payload content within the captured microvesicles can be assessed, thereby providing further biomarker readout of the payload content.

[00109] A biosignature for characterizing a phenotype may comprise any number of useful criteria. As described further below, the term "phenotype" as used herein can mean any trait or characteristic that is attributed to a biosignature / biomarker profile. A phenotype can be detected or identified in part or in whole using the compositions and/or methods of the invention. In some embodiments, at least one
criterion is used for each biomarker. In some embodiments, at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90 or at least 100 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.

[00110] **Theranostics**

[00111] 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, including predictive medicine, personalized medicine, integrated medicine, pharmacodiagnostics and Dx/Rx partnering. Therapy related tests can be used to predict and assess drug response in individual subjects, i.e., to provide personalized medicine. Predicting a drug response can be determining whether a subject is a likely responder or a likely non-responder to a candidate therapeutic agent, e.g., before the subject has been exposed or otherwise treated with the treatment. Assessing a drug response can be monitoring a response to a drug, e.g., monitoring the subject's improvement or lack thereof over a time course after initiating the treatment. Therapy related tests are 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).

[00112] The compositions and methods of the invention can be used to identify or detect a biosignature associated with 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 (i.e., theranostics) are also useful in clinical diagnosis and management of many such diseases and disorders. 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.

[00113] In some embodiments, the compositions and methods of the invention provide for 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.

[00114] One or more novel biosignatures of a vesicle can also be identified. For example, one or more vesicles can be isolated from a subject that responds to a drug treatment or treatment regimen and compared to a reference, such as another subject that does not respond to the drug treatment or treatment regimen. Differences between the biosignatures can be determined and used to identify other subjects as responders or non-responders to a particular drug or treatment regimen.

[00115] In some embodiments, a biosignature is used to determine whether a particular disease or condition is resistant to a drug, in which case 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.

[00116] Biosignatures can be used in the theranosis of a cancer, such as identifying whether a subject suffering from cancer is a likely responder or non-responder to a particular cancer treatment. The subject methods can be used to theranose cancers including those listed herein, e.g., in the "Phenotypes" section below. These include without limitation lung cancer, non-small cell lung cancer 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, pancreatic cancer, brain cancer, kidney cancer, ovarian cancer, stomach cancer, melanoma, bone cancer, gastric cancer, breast cancer, glioma, glioblastoma, hepatocellular carcinoma, papillary renal carcinoma, head and neck squamous cell carcinoma, leukemia, lymphoma, myeloma, or other solid tumors.

[00117] A biosignature of circulating biomarkers, including markers associated with a component present in a biological sample (e.g., cell, cell-fragment, cell-derived extracellular vesicle), in a sample from a subject suffering from a cancer can be used select a candidate treatment for the subject. The biosignature
can be determined according to the methods of the invention presented herein. In some embodiments, the candidate treatment comprises a standard of care for the cancer. The treatment can be a cancer treatment such as radiation, surgery-, chemotherapy or a combination thereof. The cancer treatment can be a therapeutic such as anti-cancer agents and chemotherapeutic regimens. Further drug associations and rules that are used in embodiments of the invention are found in PCT/US2007/69286, filed May 18, 2007; PCT/US2009/660630, filed October 14, 2009; PCT/2010/000407, filed February 11, 2010; PCT/US12/41393, filed June 7, 2012; PCT/US2013/073184, filed December 4, 2013; PCT/US2010/54366, filed October 27, 2010; PCT/US 11/67527, filed December 28, 2011; PCT/US15/13618, filed January 29, 2015; and PCT/US 16/20657, filed March 3, 2016.

Biomarker Detection

[00118] The compositions and methods of the invention can be used to assess any useful biomarkers in a biological sample for characterizing a phenotype associated with the sample. Such biomarkers include all sorts of biological entities such as proteins, nucleic acids, lipids, carbohydrates, complexes of any thereof, and microvesicles. Various molecules associated with a microvesicle surface or enclosed within the microvesicle (referred to herein as "payload") can serve as biomarkers. The microvesicles themselves can also be used as biomarkers.

[00119] The aptamers of the invention can be used to assess levels or presence of a microvesicle population. See, e.g., FIGs. 19B-C. The aptamers of the invention can also be used to assess levels or presence of their specific target molecule. See, e.g., FIG. 19A. In addition, aptamers of the invention are used to capture or isolated a component present in a biological sample that has the aptamer's target molecule present. For example, if a given microvesicle surface antigen is present on a cell, cell fragment or cell-derived extracellular vesicle. A binding agent to the biomarker, including without limitation an aptamer provided by the invention, may be used to capture or isolate the cell, cell fragment or cell-derived extracellular vesicles. See, e.g., FIGs. 2A-B, 19A. Such captured or isolated entities may be further characterized to assess additional surface antigens or internal "payload" molecules present (i.e., nucleic acid molecules, lipids, sugars, polypeptides or functional fragments thereof, or anything else present in the cellular milieu that may be used as a biomarker), where one or more biomarkers provide a biosignature to assess a desired phenotype, such as a disease or condition. See, e.g., FIG. 2F. Therefore, aptamers of the invention are used not only to assess one or more microvesicle surface antigen of interest but are also used to separate a component present in a biological sample, where the components themselves can be further assessed to identify a candidate biosignature.

[00120] The methods of the invention can comprise multiplex analysis 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. 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, wherein 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. Detection can be with more than one biomarker, including without limitation more than one vesicle marker such as in any of Tables 3-4, and Tables 18-25.

[00121] An immunoassay based method (e.g., sandwich assay) can be used to detect a biomarker of a vesicle. An example includes ELISA. A binding agent can be bound to a well. For example, a binding agent such as an aptamer or 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 agent 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 binding agent (detection agent) against vesicle antigens of interest. Multiple capture binding agents can be used, e.g., in distinguishable addresses on an array or different wells of an immunoassay plate. The detection binding agents can be against the same antigen as the capture binding agent, or can be directed against other markers. The capture binding agent can be any useful binding agent, e.g., tethered aptamers, antibodies or lectins, and/or the detector antibodies can be similarly substituted, e.g., with detectable (e.g., labeled) aptamers, antibodies, 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, or other markers in Table 3 herein. Examples of microvesicle surface antigens are disclosed herein, e.g. in Tables 3-4 and 18-25, or are known in the art, and examples useful in methods and compositions of the invention are disclosed of International Patent Application Nos.

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 microvesicles 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, e.g., in Tables 1, 3-4 and 18-25.

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.

The methods of the invention comprise capture and detection of microvesicles of interest using any combination of useful biomarkers. For example, a microvesicle population can be captured using one or more binding agent to any desired combination of cell of origin, disease specific, or general vesicle markers. The captured microvesicles can then be detected using one or more binding agent to any desired combination of cell of origin, disease specific, or general vesicle markers. FIG. 2E represents a flow
Any one or more of a cell-of-origin biomarker (240), disease biomarkers (241), and general vesicle biomarker (242) is used to capture a microvesicle population. Thereafter, any one or more of a cell-of-origin biomarker (243), disease biomarkers (244), and general vesicle biomarker (245) is used to detect the captured microvesicle population. The biosignature of captured and detected microvesicles is then used to characterize a phenotype. The biomarkers combinations are selected to characterize the phenotype of interest and can be selected from the biomarkers and phenotypes described herein.

A microvesicle payload molecule can be assessed as a member of a biosignature panel. A payload molecule comprises any of the biological entities contained within a cell, cell fragment or 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. 2F iv, a population of vesicles is captured and/or detected (220) using one or more of cell-of-origin biomarkers (220), disease biomarkers (221), and/or 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.

In other embodiments, vesicle or cellular payload is analyzed in a population (e.g., cells or vesicles) without first capturing or detecting subpopulations of vesicles. For example, a cellular or extracellular vesicle population can be generally isolated from a sample using centrifugation, filtration, chromatography, or other techniques as described herein and known in the art. The payload of such sample components can be analyzed thereafter to detect a biosignature and characterize a phenotype. In the scheme shown in FIG. 2F v, 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.

The biomarkers used to detect a vesicle population can be selected to detect a microvesicle population of interest, e.g., a population of vesicles that provides a diagnosis, prognosis or theranosis of a selected condition or disease, including but not limited to a cancer, a premalignant condition, an inflammatory disease, an immune disease, an autoimmune disease or disorder, a cardiovascular disease or disorder, neurological disease or disorder, infectious disease or pain. See Section "Phenotypes" herein for more detail. In an embodiment, the biomarkers are selected from the group consisting of EpCam (epithelial cell adhesion molecule), CD9 (tetraspanin CD9 molecule), PCSA (prostate cell specific antigen), see Rokhlin et al., 5E10: a prostate-specific surface-reactive monoclonal antibody. Cancer Lett. 1998 131:129-36), CD63 (tetraspanin CD63 molecule), CD81 (tetraspanin CD81 molecule), PSMA (FOLH1, folate hydrolase (prostate-specific membrane antigen) 1), B7H3 (CD276 molecule), PSCA (prostate stem cell antigen), ICAM (intercellular adhesion molecule), STEAP (STEAP1, six transmembrane epithelial antigen of the prostate 1), KLK2 (kallikrein-related peptidase 2), SSX2 (synovial sarcoma, X breakpoint 2), SSX4 (synovial sarcoma, X breakpoint 4), PBP (prostatic binding protein), SPDEF (SAM pointed domain containing ets transcription factor), EGFR (epidermal growth factor receptor), and a combination thereof. One or more of these markers can provide a biosignature for a specific condition, such as to detect a cancer, including without limitation a carcinoma, a prostate cancer, a breast cancer, a lung cancer, a colorectal cancer, an ovarian cancer, melanoma, a brain cancer, or other type of cancer as disclosed herein. In an embodiment, a binding agent to one or more of these markers is used to capture a microvesicle population, and an aptamer of the invention is used to assist in detection of the capture vesicles as described herein. In other embodiments, an aptamer of the invention is used to capture a microvesicle population, and a binding agent to one or more of these markers is used to assist in detection of the capture vesicles as described herein. The binding agents can be any useful binding agent as disclosed herein or known in the art, e.g., antibodies or aptamers.

The methods of characterizing a phenotype can employ a combination of techniques to assess a component or population of components present in a biological sample of interest. For example, an aptamer of the invention can be used to assess a single cell, or a single extracellular vesicle or a population of cells or population of vesicles. A sample may be split into various aliquots, where each is analyzed separately. For example, protein content of one or more aliquot is determined and microRNA content of one or more other aliquot is determined. The protein content and microRNA content can be combined to characterize a phenotype. In another embodiment, a component present in a biological sample of interest is isolated and the payload therein is assessed (e.g., capture a population of subpopulation of vesicles using an aptamer of the invention and further assess nucleic acid or proteins present in the isolated vesicles).

In one embodiment, a population of vesicles with a given surface marker can be isolated by using a binding agent to a microvesicle surface marker. See, e.g., FIGs. 2A, 2B, 21A. The binding agent can be an aptamer that was identified to target the microvesicle surface marker using to the methods of the
invention. The isolated vesicles is assessed for additional biomarkers such as surface content or payload, which can be contemporaneous to detection of the aptamer-specific target or the assessment of additional biomarkers can be before or subsequent to aptamer-specific target detection.

[00130] A biosignature can be detected qualitatively or quantitatively by detecting a presence, level or concentration of a circulating biomarker, e.g., a microRNA, protein, vesicle or other biomarker, 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 immunoassay (ELISA; EIA), radioimmunoassay (RIA), flow cytometry, or electron microscopy (EM).

[00131] Biosignatures can be detected using aptamers of the invention that function as either as 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. The detection agent can be a binding agent, e.g., an antibody or aptamer. In other embodiments, the detection agent comprises a small molecule such as a membrane protein labeling agent. See, e.g., the membrane protein labeling agents disclosed in Alroy et al., US. Patent Publication US 2005/0158708. In an embodiment, vesicles are isolated or captured as described herein, and one or more membrane protein labeling agent is used to detect the vesicles. In many cases, the antigen or other vesicle-moiety that is recognized by the capture and detection agents are interchangeable.

[00132] In a non-limiting embodiment, a vesicle having a cell-of-origin specific antigen on its surface and a cancer-specific antigen on its surface, is captured using a binding agent that is specific to a cells-specific antigen, e.g., by tethering the capture antibody or aptamer to a substrate, and then the vesicle is detected using a binding agent to a disease-specific antigen, e.g., by labeling the binding agent used for detection with a fluorescent dye and detecting the fluorescent radiation emitted by the dye.

[00133] It will be apparent to one of skill in the art that where the target molecule for a binding agent (such as an aptamer of the invention) is informative as to assessing a condition or disease, the same
binding agent can be used to both capture a component comprising the target molecule (e.g., microvesicle surface antigen of interest) and also be modified to comprise a detectable label so as to detect the target molecule, e.g., binding agent-antigen-binding agent₂*, wherein the * signifies a detectable label; binding agent₁ and binding agent₂ may be the same binding agent or a different binding agent (e.g., same aptamer or different aptamer). In addition, binding agent₁ and binding agent₂ can be selected from wholly different categories of binding agents (e.g., antibody, aptamer, synthetic antibody, peptide-nucleic acid molecule, or any molecule that is configured to specifically bind to or associate with its target molecule). Such binding molecules can be selected solely based on their binding specificity for a target molecule.

Examples of additional biomarkers that can be incorporated into the methods and compositions of the invention are known in the art, such as those disclosed in International Patent Publication Nos. WO/2012/174282 (Int'l Appl. PCT/US20 12/0425 19 filed June 14, 2012) and WO/2013/020995 (Int'l Appl. PCT/US20 12/050030 filed August 8, 2013). The detectable signal can itself be associated with a nucleic acid molecule that hybridizes with a stretch of nucleic acids present in each oligonucleotide comprising a probing library. The stretch can be the same or different as to one or more oligonucleotides in a library. The detectable signal can comprise fluorescence agents, including color-coded barcodes which are known, such as in U.S. Patent Application Pub. No. 2014037 1088, 20130 17837, and 20120258870.

[00134] Techniques of detecting biomarkers or capturing sample components using an aptamer of 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. Additional examples of binding agents described above and useful in the compositions and methods of the invention are disclosed in International Patent Publication No. WO/2011/127219, entitled "Circulating Biomarkers for Disease" and filed April 6, 2011, which application is incorporated by reference in its entirety herein. Aptamers of the invention 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).

[00135] An aptamer of the invention or other useful binding agent may be linked directly or indirectly to a solid surface or substrate. See, e.g., FIGS. 2A-2B, 9, 21A. 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 material, 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 polypyrrole and polypyrrole); micro or nanostructured surfaces such as nucleic acid tiling arrays, nanotube, nanowire, or nanoparticulate decorated surfaces; or porous surfaces or gels such as methacylates, 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 dextran, acrylamides, gelatins or agarose. Such coatings can facilitate the use of the array with a biological sample.

[00136] As provided in the examples, below, an aptamer or other useful binding agent can be conjugated to a detectable entity or label.

[00137] 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 radionuclides (radionuclides), such as 3H, 14C, 18F, 32P, 35S, 65Cu, 68Ga, 86Y, 90Tc, 111In, 123I, 124I, 125I, 131I, 133Xe, 177Lu, 211At, or 213Bi. 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-

[00138] Using conventional techniques, an aptamer can be directly or indirectly labeled, e.g., the label is attached to the aptamer through biotin-streptavidin (e.g., synthesize a biotinylated aptamer, which is then capable of binding a streptavidin molecule that is itself conjugated to a detectable label; non-limiting example is streptavidin, phycoerythrin conjugated (SAPE)). 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-hydroxy succinimide. Succinimide groups react effectively with amino groups at pH values above 7, and preferentially between about pH 8.0 and about pH 8.5. Alternatively, an aptamer 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.

[00139] Various enzyme-substrate labels may also be used in conjunction with a composition or method of the invention. Such enzyme-substrate labels are available commercially (e.g., 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-D-galactosidase.

[00140] Aptamer(s) can be linked to a substrate such as a planar substrate. 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.

[00141] 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., aptamers or antibodies), where biomarker binding is monitored indirectly (e.g., via fluorescence).

[00142] An array or microarray that can be used to detect one or more biomarkers of a biosignature and comprising one or more aptamers of the invention 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.

[00143] In some embodiments, multiple capture molecules are disposed on an array, e.g., proteins, peptides or additional nucleic acid molecules. 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. The capture molecules can comprise one or more aptamer of the invention. In one embodiment, an array is constructed for the hybridization of a pool of aptamers. The array can then be used to identify pool members that bind a sample, thereby facilitating characterization of a phenotype. See FIGs. 19B-19C and related disclosure for further details.

[00144] 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 cm². 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.
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 desired.

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.

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.

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.

[00149] 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.

[00150] 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.

[00151] 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 known 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., UngerMetal., 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.

[00152] An array suitable for identifying a disease, condition, syndrome or physiological status can be included in a kit. A kit can include, an aptamer of the invention, including as non-limiting examples, one
or more reagents useful for preparing molecules for immobilization onto binding islands or areas of an array, reagents useful for detecting binding of a vesicle to immobilized molecules, and instructions for use.

[00153] 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.

[00154] **Particle arrays**

[00155] As an alternative to planar arrays, assays using particles, such as bead based assays are also capable of use with an aptamer of the invention. Aptamers are easily conjugated with commercially available beads. See, e.g., Srinivas et al. Anal. Chem. 2011 Oct. 21. Aptamer functionalized Microgel Particles for Protein Detection; See also, review article on aptamers as therapeutic and diagnostic agents, Brody and Gold, Rev. Mol. Biotech. 2000, 74:5-13.

[00156] 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 desired to generate custom bead arrays. Bead arrays are then incubated with the sample in a single reaction vessel to perform the assay.

[00157] Bead-based assays can also be used with one or more aptamers of the invention. A bead substrate can provide a platform for attaching one or more binding agents, including aptamer(s). For multiplexing, multiple different bead sets (e.g., Illumina, Luminex) can have different binding agents (specific to different target molecules). For example, a bead can be conjugated to an aptamer of the invention used to detect the presence (quantitatively or qualitatively) of an antigen of interest, or it can also be used to isolate a component present in a selected biological sample (e.g., cell, cell-fragment or vesicle comprising the target molecule to which the aptamer is configured to bind or associate). Any molecule of organic origin can be successfully conjugated to a polystyrene bead through use of commercially available kits.

[00158] One or more aptamers of the invention can be used with any bead based substrate, including but not limited to magnetic capture method, fluorescence activated cell sorting (FACS) or laser cytomtery. Magnetic capture methods can include, but are not limited to, the use of magnetically activated cell sorter
(MACS) microbeads or magnetic columns. Examples of bead or particle based methods that can be modified to use an aptamer of the invention include methods and bead systems described 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; 7,399,632; 8,124,015; 8,008,019; 7,955,802; 7,445,844; 7,274,316; 6,773,812; 6,623,526; 6,599,331; 6,057,107; 5,736,330; International Patent Publication No. WO/2012/174282; WO/1993/022684.

[00159] Flow Cytometry

[00160] Isolation or detection of circulating biomarkers, e.g., protein antigens, from a biological sample, or of the biomarker-comprising cells, cell fragments or vesicles may also be achieved using an aptamer of the invention in a cytometry process. As a non-limiting example, aptamers of the invention can be used in an assay comprising using a particle such as a bead or microsphere. The invention provides aptamers as binding agents, which may be conjugated to the particle. 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.

[00161] 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.

[00162] 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. Flow cytometers can have multiple lasers and fluorescence detectors, allowing multiple labels to be used to more precisely specify a target population by their phenotype. Thus, a flow cytometer, such as a multicolor flow cytometer, can be used to detect one or more vesicles with multiple fluorescent labels or colors. In some embodiments, the flow cytometer can also sort or isolate different vesicle populations, such as by size or by different markers.

[00163] The flow cytometer may have one or more lasers, such as 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more lasers. In some embodiments, the flow cytometer can detect more than one color or fluorescent label, such
as at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 different colors or fluorescent labels. For example, the flow cytometer can have at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 fluorescence detectors.

Examples of commercially available flow cytometers that can be used to detect or analyze one or more vesicles, to sort or separate different populations of vesicles, include, but are not limited to the MoFlo™ XDP Cell Sorter (Beckman Coulter, Brea, CA), MoFlo™ Legacy Cell Sorter (Beckman Coulter, Brea, CA), BD FACSAria™ Cell Sorter (BD Biosciences, San Jose, CA), BD™ LSRII (BD Biosciences, San Jose, CA), and BD FACSCalibur™ (BD Biosciences, San Jose, CA). Use of multicolor or multi-fluor cytometers can be used in multiplex analysis of vesicles, as further described below. In some embodiments, the flow cytometer can sort, and thereby collect or sort more than one population of vesicles based on one or more characteristics. For example, two populations of vesicles differ in size, such that the vesicles within each population have a similar size range and can be differentially detected or sorted. In another embodiment, two different populations of vesicles are differentially labeled.

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. In a multiplexed assay, including but not limited to a flow cytometry assay, one or more different target molecules can be assessed, wherein at least one of the target molecules is a microvesicle surface antigen assessed using an aptamer of the invention.

Microfluidics

One or more aptamer of the invention can be disposed on any useful planar or bead substrate. In one aspect of the invention one or more aptamer of the invention is disposed on a microfluidic device, thereby facilitating assessing, characterizing or isolating a component of a biological sample comprising a polypeptide antigen of interest or a functional fragment thereof. For example, the circulating antigen or a cell, cell fragment or cell-derived vesicles comprising the antigen can be assessed using one or more aptamers of the invention (alternatively along with additional binding agents). 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 a vesicle. Such systems miniaturize and compartmentalize processes that allow for binding of vesicles, detection of biosignatures, and other processes.

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 a vesicle. The selection can be based on a property of the vesicle, such as the size, shape, deformability, or biosignature of the vesicle.

[00169] In one embodiment, 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 the 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.

[00170] In some embodiments, the microfluidic device can comprise one or more channels that permit further enrichment or selection of a vesicle. 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 or vesicle population to be further enriched, such as through one or more binding agents present in the second channel.

[00171] 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 vesicle 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. In one embodiment, each population has a different biosignature. The hybridization reaction between the microsphere and vesicle 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.

[00172] Any appropriate microfluidic device can be used in the methods of the invention. 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, 7,189,568, 7,141,978, 7,138,062, 7,135,147, 7,125,711, 7,118,910, 7,118,661, 7,640,947, 7,666,361, 7,704,735; and International Patent Publication WO 2010/072410; each of which patents or applications are incorporated herein by reference in their entirety. Another example for use with methods disclosed herein is described in Chen et al., "Microfluidic isolation and transcriptome analysis of serum vesicles," Lab on a Chip, Dec. 8, 2009 DOI: 10.1039/b916199f.
Other microfluidic devices for use with the invention include devices comprising elastomeric layers, valves and pumps, including without limitation those disclosed in U.S. Patent Nos. 5,376,252, 6,408,878, 6,645,432, 6,719,868, 6,793,753, 6,899,137, 6,929,030, 7,040,338, 7,118,910, 7,144,616, 7,216,671, 7,250,128, 7,494,555, 7,501,245, 7,601,270, 7,691,333, 7,754,010, 7,837,946; U.S. Patent Application Nos. 2003/0061687, 2005/0084421, 2005/0112882, 2005/0129581, 2005/0145496, 2005/0201901, 2005/0214173, 2005/0252773, 2006/0006067; and EP Patent Nos. 0527905 and 1065378; each of which application is herein incorporated by reference. In some instances, much or all of the devices are composed of elastomeric material. Certain devices are designed to conduct thermal cycling reactions (e.g., PCR) with devices that include one or more elastomeric valves to regulate solution flow through the device. The devices can comprise arrays of reaction sites thereby allowing a plurality of reactions to be performed. Thus, the devices can be used to assess circulating microRNAs in a multiplex fashion, including microRNAs isolated from vesicles. In an embodiment, the microfluidic device comprises (a) a first plurality of flow channels formed in an elastomeric substrate; (b) a second plurality of flow channels formed in the elastomeric substrate that intersect the first plurality of flow channels to define an array of reaction sites, each reaction site located at an intersection of one of the first and second flow channels; (c) a plurality of isolation valves disposed along the first and second plurality of flow channels and spaced between the reaction sites that can be actuated to isolate a solution within each of the reaction sites from solutions at other reaction sites, wherein the isolation valves comprise one or more control channels that each overlay and intersect one or more of the flow channels; and (d) means for simultaneously actuating the valves for isolating the reaction sites from each other. Various modifications to the basic structure of the device are envisioned within the scope of the invention. MicroRNAs can be detected in each of the reaction sites by using PCR methods. For example, the method can comprise the steps of the steps of: (i) providing a microfluidic device, the microfluidic device comprising: a first fluidic channel having a first end and a second end in fluid communication with each other through the channel; a plurality of flow channels, each flow channel terminating at a terminal wall; wherein each flow channel branches from and is in fluid communication with the first fluidic channel, wherein an aqueous fluid that enters one of the flow channels from the first fluidic channel can flow out of the flow channel only through the first fluidic channel; and, an inlet in fluid communication with the first fluidic channel, the inlet for introducing a sample fluid; wherein each flow channel is associated with a valve that when closed isolates one end of the flow channel from the first fluidic channel, whereby an isolated reaction site is formed between the valve and the terminal wall; a control channel; wherein each the valve is a deflectable membrane which is deflected into the flow channel associated with the valve when an actuating force is applied to the control channel, thereby closing the valve; and wherein when the actuating force is applied to the control channel a valve in each of the flow channels is closed, so as to produce the isolated reaction site in each flow channel; (ii) introducing the sample fluid into the inlet, the sample fluid filling the flow channels; (iii) actuating the valve to separate the sample fluid into the separate portions within the flow channels; (iv) amplifying the nucleic acid in the sample fluid; (v) analyzing the portions of the sample fluid to determine whether the amplifying produced the reaction. The sample fluid can contain an
amplifiable nucleic acid target, e.g., a microRNA, and the conditions can be polymerase chain reaction (PCR) conditions, so that the reaction results in a PCR product being formed.

[00174] The microfluidic device can have one or more binding agents attached to a surface in a channel, or present in a channel. For example, the microchannel can have one or more capture agents, such as a capture agent for a tissue related antigen in Table 4, one or more general microvesicle antigen in Table 3 or a cell-of-origin or cancer related antigen in Table 4, including without limitation EpCam, CD9, CD63, CD81, B7H3, ICAM, STEAP, KLK2, SSX2, SSX4, PBP, SPDEF, and EGFR. The capture agent may be an aptamer selected by the methods of the invention. The surface of the channel can also be contacted with a blocking aptamer. In one embodiment, a microchannel surface is treated with avidin and a capture agent, such as an antibody, that is biotinylated can be injected into the channel to bind the avidin. In other embodiments, the capture agents are present in chambers or other components of a microfluidic device. The capture agents can also be attached to beads that can be manipulated to move through the microfluidic channels. In one embodiment, the capture agents are attached to magnetic beads. The beads can be manipulated using magnets.

[00175] A biological sample can be flowed into the microfluidic device, or a microchannel, at rates such as at least about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, or 50 µl per minute, such as between about 1-50, 5-40, 5-30, 3-20 or 5-15 µl per minute. One or more vesicles can be captured and directly detected in the microfluidic device. Alternatively, the captured vesicle may be released and exit the microfluidic device prior to analysis. In another embodiment, one or more captured vesicles are lysed in the microchannel and the lysate can be analyzed, e.g., to examine payload within the vesicles. Lysis buffer can be flowed through the channel and lyse the captured vesicles. For example, the lysis buffer can be flowed into the device or microchannel at rates such as at least about a, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 26, 27, 28, 29, 30, 35, 40, 45, or 50 µl per minute, such as between about 1-50, 5-40, 10-30, 5-30 or 10-35 µl per minute. The lysate can be collected and analyzed, such as performing RT-PCR, PCR, mass spectrometry, Western blotting, or other assays, to detect one or more biomarkers of the vesicle.

Phenotypes

[00176] Disclosed herein are products and processes for characterizing a phenotype using the methods and compositions of the invention. The term "phenotype" as used herein can mean any trait or characteristic that is attributed to a biomarker profile that is identified using in part or in whole the compositions and/or methods of the invention. For example, a phenotype can be a diagnostic, prognostic or theranostic determination based on a characterized biomarker profile for a sample obtained from a subject. A phenotype can be any observable characteristic or trait of, such as a disease or condition, a stage of a disease or condition, susceptibility to a disease or condition, prognosis of a disease stage or condition, a physiological state, or response / potential response to therapeutics. A phenotype can result from a subject's genetic makeup as well as the influence of environmental factors and the interactions between the two, as well as from epigenetic modifications to nucleic acid sequences.
A phenotype in a subject can be characterized by obtaining a biological sample from a subject and analyzing the sample using the compositions and/or methods of the invention. For example, characterizing a phenotype for a subject or individual can include detecting a disease or condition (including pre-symptomatic early stage detecting), determining a prognosis, diagnosis, or theranosis of a disease or condition, or determining the stage or progression of a disease or condition. Characterizing a phenotype can 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. The compositions and methods described herein allow assessment of a subject on an individual basis, which can provide benefits of more efficient and economical decisions in treatment.

In an aspect, the invention relates to the analysis of biomarkers such as microvesicles to provide a diagnosis, prognosis, and/or theranosis of a disease or condition. Theranostics includes diagnostic testing that provides the ability to affect therapy or treatment of a disease or disease 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, including predictive medicine, personalized medicine, integrated medicine, pharmacodiagnosics and Dx/Rx partnering. Therapy related tests can be used to predict and assess drug response in individual subjects, i.e., to provide personalized medicine. Predicting a drug response can be determining whether a subject is a likely responder or a likely non-responder to a candidate therapeutic agent, e.g., before the subject has been exposed or otherwise treated with the treatment. Assessing a drug response can be monitoring a response to a drug, e.g., monitoring the subject’s improvement or lack thereof over a time course after initiating the treatment. Therapy related tests are 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, analysis using the compositions and methods of the invention 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).

Thus, the compositions and methods of the invention may help predict whether a subject is likely to respond to a treatment for a disease or disorder. Characterizing a phenotype includes predicting the responder / non-responder status of the subject, wherein a responder responds to a treatment for a disease and a non-responder does not respond to the treatment. Biomarkers such as microvesicles can be analyzed in the subject and compared against that of previous subjects that were known to respond or not to a treatment. If the biomarker profile in the subject more closely aligns with that of previous subjects that were known to respond to the treatment, the subject can be characterized, or predicted, as a responder to the treatment. Similarly, if the biomarker profile in the subject more closely aligns with that of previous subjects that did not respond to the treatment, the subject can be characterized, or predicted as a non-
responder to the treatment. The treatment can be for any appropriate disease, disorder or other condition, including without limitation those disclosed herein.

[00180] In some embodiments, the phenotype comprises a disease or condition such as those listed in Tables 1 or 16. For example, the phenotype can comprise detecting the presence of or likelihood of developing a tumor, neoplasm, or cancer, or characterizing the tumor, neoplasm, or cancer (e.g., stage, grade, aggressiveness, likelihood of metastasis or recurrence, etc). Cancers that can be detected or assessed by methods or compositions described herein include, but are 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.

[00181] The phenotype can be a premalignant condition, such as actinic keratosis, atrophic gastritis, leukoplakia, erythroplasia, Lymphomatoid Granulomatosis, preleukemia, fibrosis, cervical dysplasia, uterine cervical dysplasia, xeroderma pigmentosum, Barrett's Esophagus, colorectal polyp, or other abnormal tissue growth or lesion that is likely to develop into a malignant tumor. Transformative viral infections such as HIV and HPV also present phenotypes that can be assessed according to the invention.

[00182] A 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, gawitz 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, melanocyte 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 hemangioendothelioma, malignant schwannoma, osteosarcoma, soft tissue sarcomas including: alveolar soft part sarcoma, angiosarcoma, cystosarcoma phylloides, 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 synovial sarcoma. Lymphoma and leukemia include without limitation chronic lymphocytic leukemia/small lymphocytic lymphoma, B-cell prolymphocytic leukemia, lymphoplasmacytic lymphoma (such as Waldenstrom 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 (nmzI), 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, dyserminoma, seminoma, nongerminomatous germ cell tumor, embryonal carcinoma, endodermal sinus tumor, 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, menigioma, medulloblastoma and peripheral neuroectodermal tumors, gall bladder carcinoma, bronchial carcinoma, multiple myeloma, basalioma, teratoma, retinoblastoma, choriodea melanoma, seminoma, rhabdomyosarcoma, craniopharyngeoma, osteosarcoma, chondrosarcoma, myosarcoma, liposarcoma, fibrosarcoma, Ewing sarcoma, and plasmocytoma.

[00183] 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.
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; Sezary 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; Waldenstrom macroglobulinemia; or Wilms'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.

In some embodiments, the cancer comprises an acute myeloid leukemia (AML), breast carcinoma, cholangiocarcinoma, colorectal adenocarcinoma, extrahepatic bile duct adenocarcinoma, female genital tract malignancy, gastric adenocarcinoma, gastroesophageal adenocarcinoma, gastrointestinal stromal tumors (GIST), glioblastoma, head and neck squamous carcinoma, leukemia, liver hepatocellular carcinoma, low grade glioma, lung bronchioalveolar carcinoma (BAC), lung non-small cell lung cancer (NSCLC), lung small cell cancer (SCLC), lymphoma, male genital tract malignancy, malignant solitary fibrous tumor of the pleura (MSFT), melanoma, multiple myeloma, neuroendocrine tumor, nodal diffuse large B-cell lymphoma, non epithelial ovarian cancer (non-EOC), ovarian surface epithelial carcinoma, pancreatic adenocarcinoma, pituitary carcinomas, oligodendroglioma, prostatic adenocarcinoma, retroperitoneal or peritoneal carcinoma, retroperitoneal or peritoneal sarcoma, small intestinal malignancy, soft tissue tumor, thymic carcinoma, thyroid carcinoma, or uveal melanoma. 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.

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.

The phenotype can also comprise a cardiovascular disease, such as atherosclerosis, congestive heart failure, vulnerable plaque, stroke, or ischemia. The cardiovascular disease or condition can be high blood pressure, stenosis, vessel occlusion or a thrombotic event.

The phenotype can also comprise a neurological disease, such as Multiple Sclerosis (MS), Parkinson's Disease (PD), Alzheimer's Disease (AD), schizophrenia, bipolar disorder, depression, autism, Prion Disease, Pick's disease, dementia, Huntington disease (HD), Down's syndrome, cerebrovascular disease, Rasmussen's encephalitis, viral meningitis, neuropsychiatric systemic lupus erythematosus (NPSLE), amyotrophic lateral sclerosis, Creutzfeldt-Jacob disease, Gerstmann-Straussler-Scheinker disease, transmissible spongiform encephalopathy, ischemic reperfusion damage (e.g. stroke), brain trauma, microbial infection, or chronic fatigue syndrome. The phenotype may also be a condition such as fibromyalgia, chronic neuropathic pain, or peripheral neuropathic pain.

The phenotype may also comprise an infectious disease, such as a bacterial, viral or yeast infection. For example, the disease or condition may be Whipple's Disease, Prion Disease, cirrhosis, methicillin-resistant staphylococcus aureus, HIV, hepatitis, syphilis, meningitis, malaria, tuberculosis, or influenza. Viral proteins, such as HIV or HCV-like particles can be assessed in a vesicle, to characterize a viral condition.
The phenotype can also comprise a perinatal or pregnancy related condition (e.g. preeclampsia or preterm birth), metabolic disease or condition, such as a metabolic disease or condition associated with iron metabolism. For example, hepcidin can be assayed in a vesicle to characterize an iron deficiency. The metabolic disease or condition can also be diabetes, inflammation, or a perinatal condition.

[00191] The compositions and methods of the invention can be used to characterize these and other diseases and disorders that can be assessed via biomarkers. Thus, characterizing a phenotype can be providing a diagnosis, prognosis or theranosis of one of the diseases and disorders disclosed herein.

Subject

[00192] One or more phenotypes of a subject can be determined by analyzing one or more vesicles, such as vesicles, in 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 can also include a mammal of importance due to being endangered, such as a Siberian tiger; or economic importance, such as an animal raised on a farm for consumption by humans, or an animal of social importance to humans, such as an animal kept as a pet or in a zoo. 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.

[00193] 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.

Samples

[00194] A sample used and/or assessed via the compositions and methods of the invention includes any relevant biological sample that can be used for biomarker assessment, including without limitation sections of tissues such as biopsy or tissue removed during surgical or other procedures, bodily fluids, autopsy samples, frozen sections taken for histological purposes, and cell cultures. Such samples include blood and blood fractions or products (e.g., serum, buffy coat, plasma, platelets, red blood cells, and the like), sputum, malignant effusion, cheek cells tissue, cultured cells (e.g., primary cultures, explants, and transformed cells), stool, urine, other biological or bodily fluids (e.g., prostatic fluid, gastric fluid, intestinal fluid, renal fluid, lung fluid, cerebrospinal fluid, and the like), etc. The sample can comprise biological material that is a fresh frozen & formalin fixed paraffin embedded (FFPE) block, formalin-fixed paraffin embedded, or is within an RNA preservative + formalin fixative. More than one sample of more than one type can be used for each patient.
The sample used in the methods described herein can be a formalin fixed paraffin embedded (FFPE) sample. The FFPE sample can be one or more of fixed tissue, unstained slides, bone marrow core or clot, core needle biopsy, malignant fluids and fine needle aspirate (FNA). In an embodiment, the fixed tissue comprises a tumor containing formalin fixed paraffin embedded (FFPE) block from a surgery or biopsy. In another embodiment, the unstained slides comprise unstained, charged, unbaked slides from a paraffin block. In another embodiment, bone marrow core or clot comprises a decalcified core. A formalin fixed core and/or clot can be paraffin-embedded. In still another embodiment, the core needle biopsy comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more, e.g., 3-4, paraffin embedded biopsy samples. An 18 gauge needle biopsy can be used. The malignant fluid can comprise a sufficient volume of fresh pleural/ascitic fluid to produce a 5x5x2mm cell pellet. The fluid can be formalin fixed in a paraffin block. In an embodiment, the core needle biopsy comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more, e.g., 4-6, paraffin embedded aspirates.

A sample may be processed according to techniques understood by those in the art. A sample can be without limitation fresh, frozen or fixed cells or tissue. In some embodiments, a sample comprises formalin-fixed paraffin-embedded (FFPE) tissue, fresh tissue or fresh frozen (FF) tissue. A sample can comprise cultured cells, including primary or immortalized cell lines derived from a subject sample. A sample can also refer to an extract from a sample from a subject. For example, a sample can comprise DNA, RNA or protein extracted from a tissue or a bodily fluid. Many techniques and commercial kits are available for such purposes. The fresh sample from the individual can be treated with an agent to preserve RNA prior to further processing, e.g., cell lysis and extraction. Samples can include frozen samples collected for other purposes. Samples can be associated with relevant information such as age, gender, and clinical symptoms present in the subject; source of the sample; and methods of collection and storage of the sample. A sample is typically obtained from a subject.

A biopsy comprises the process of removing a tissue sample for diagnostic or prognostic evaluation, and to the tissue specimen itself. Any biopsy technique known in the art can be applied to the molecular profiling methods of the present invention. The biopsy technique applied can depend on the tissue type to be evaluated (e.g., colon, prostate, kidney, bladder, lymph node, liver, bone marrow, blood cell, lung, breast, etc.), the size and type of the tumor (e.g., solid or suspended, blood or ascites), among other factors. Representative biopsy techniques include, but are not limited to, excisional biopsy, incisional biopsy, needle biopsy, surgical biopsy, and bone marrow biopsy. An "excisional biopsy" refers to the removal of an entire tumor mass with a small margin of normal tissue surrounding it. An "incisional biopsy" refers to the removal of a wedge of tissue that includes a cross-sectional diameter of the tumor. Molecular profiling can use a "core-needle biopsy" of the tumor mass, or a "fine-needle aspiration biopsy" which generally obtains a suspension of cells from within the tumor mass. Biopsy techniques are discussed, for example, in Harrison's Principles of Internal Medicine, Kasper, et al., eds., 16th ed., 2005, Chapter 70, and throughout Part V.

Standard molecular biology techniques known in the art and not specifically described are generally followed as in Sambrook et al, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor


[00199] The biological sample assessed using the compositions and methods of the invention can be any useful bodily or biological fluid, including but not limited to peripheral blood, sera, plasma, ascites, urine, cerebrospinal fluid (CSF), sputum, saliva, bone marrow, synovial fluid, aqueous humor, amniotic fluid, cerumen, breast milk, bronchoalveolar 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, cells, cell culture, or a cell culture supernatant. A biological sample may also include the blastocyl cavity, umbilical cord blood, or maternal circulation which may be of fetal or maternal origin. The biological sample may also be a cell culture, tissue sample or biopsy from which vesicles and other circulating biomarkers may be obtained. For example, cells of interest can be cultured and vesicles isolated from the culture. In various embodiments, biomarkers or more particularly biosignatures disclosed herein can be assessed directly from such biological samples (e.g., identification of presence or levels of nucleic acid or polypeptide biomarkers or functional fragments thereof) using various methods, such as extraction of nucleic acid molecules from blood, plasma, serum or any of the foregoing biological samples, use of protein or antibody arrays to identify polypeptide (or functional fragment) biomarker(s), as well as other array, sequencing, PCR and proteomic techniques known in the art for identification and assessment of nucleic acid and polypeptide molecules. In addition, one or more components present in such samples can be first isolated or enriched and further processed to assess the presence or levels of selected biomarkers, to assess a given biosignature (e.g., isolated microvesicles prior to profiling for protein and/or nucleic acid biomarkers).

[00200] Table 1 presents a non-limiting listing of diseases, conditions, or biological states and corresponding biological samples that may be used for analysis according to the methods of the invention.

<table>
<thead>
<tr>
<th>Illustrative Disease, Condition or Biological State</th>
<th>Illustrative Biological Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cancers/neoplasms affecting the following tissue types/bodily systems: breast, lung, ovarian, colon, rectal, prostate, pancreatic, brain, bone, connective tissue, glands, skin, lymph, nervous system, endocrine, germ cell, genitourinary, hematologic/blood, bone marrow, muscle, eye,</td>
<td>Tumor, blood, serum, plasma, cerebrospinal fluid (CSF), urine, sputum, ascites, synovial fluid, semen, nipple aspirates, saliva, bronchoalveolar lavage fluid, tears, oropharyngeal washes, feces, peritoneal fluids, pleural effusion, sweat, tears, aqueous humor, pericardial fluid, lymph, chyme,</td>
</tr>
</tbody>
</table>

-55-
<table>
<thead>
<tr>
<th><strong>Neurodegenerative/neurological disorders:</strong> Parkinson's disease, Alzheimer's Disease and multiple sclerosis, Schizophrenia, and bipolar disorder, spasticity disorders, epilepsy</th>
<th>Blood, serum, plasma, CSF, urine</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cardiovascular Disease:</strong> atherosclerosis, cardiomyopathy, endocarditis, vulnerable plaques, infection</td>
<td>Blood, serum, plasma, CSF, urine</td>
</tr>
<tr>
<td><strong>Stroke:</strong> ischemic, intracerebral hemorrhage, subarachnoid hemorrhage, transient ischemic attacks (TIA)</td>
<td>Blood, serum, plasma, CSF, urine</td>
</tr>
<tr>
<td><strong>Pain disorders:</strong> peripheral neuropathic pain and chronic neuropathic pain, and fibromyalgia,</td>
<td>Blood, serum, plasma, CSF, urine</td>
</tr>
<tr>
<td><strong>Autoimmune disease:</strong> systemic and localized diseases, rheumatic disease, Lupus, Sjogren's syndrome</td>
<td>Blood, serum, plasma, CSF, urine, synovial fluid</td>
</tr>
<tr>
<td><strong>Digestive system abnormalities:</strong> Barrett's esophagus, irritable bowel syndrome, ulcerative colitis, Crohn's disease, Diverticulosis and Diverticulitis, Celiac Disease</td>
<td>Blood, serum, plasma, CSF, urine</td>
</tr>
<tr>
<td><strong>Endocrine disorders:</strong> diabetes mellitus, various forms of Thyroiditis, adrenal disorders, pituitary disorders</td>
<td>Blood, serum, plasma, CSF, urine</td>
</tr>
<tr>
<td><strong>Diseases and disorders of the skin:</strong> psoriasis</td>
<td>Blood, serum, plasma, CSF, urine, synovial fluid, tears</td>
</tr>
<tr>
<td><strong>Urological disorders:</strong> benign prostatic hypertrophy (BPH), polycystic kidney disease, interstitial cystitis</td>
<td>Blood, serum, plasma, urine</td>
</tr>
<tr>
<td><strong>Hepatic disease/injury:</strong> Cirrhosis, induced hepatotoxicity (due to exposure to natural or synthetic chemical sources)</td>
<td>Blood, serum, plasma, urine</td>
</tr>
<tr>
<td><strong>Kidney disease/injury:</strong> acute, sub-acute, chronic conditions, Podocyte injury, focal segmental glomerulosclerosis</td>
<td>Blood, serum, plasma, urine</td>
</tr>
<tr>
<td>Endometriosis</td>
<td>Blood, serum, plasma, urine, vaginal fluids</td>
</tr>
<tr>
<td>Osteoporosis</td>
<td>Blood, serum, plasma, urine, synovial fluid</td>
</tr>
<tr>
<td>Pancreatitis</td>
<td>Blood, serum, plasma, urine, pancreatic juice</td>
</tr>
<tr>
<td>Asthma</td>
<td>Blood, serum, plasma, urine, sputum, bronchiolar lavage fluid</td>
</tr>
<tr>
<td>Allergies</td>
<td>Blood, serum, plasma, urine, sputum, bronchiolar lavage fluid</td>
</tr>
<tr>
<td>Prion-related diseases</td>
<td>Blood, serum, plasma, CSF, urine</td>
</tr>
<tr>
<td><strong>Viral Infections:</strong> HIV/AIDS</td>
<td>Blood, serum, plasma, urine</td>
</tr>
<tr>
<td>Sepsis</td>
<td>Blood, serum, plasma, urine, tears, nasal lavage</td>
</tr>
<tr>
<td>Organ rejection/transplantation</td>
<td>Blood, serum, plasma, urine, various lavage fluids</td>
</tr>
<tr>
<td><strong>Differentiating conditions:</strong> adenoma versus hyperplastic polyp, irritable bowel syndrome (IBS) versus normal, classifying Dukes stages A, B, C, and/or D of colon cancer, adenoma with low-grade hyperplasia versus high-grade hyperplasia.</td>
<td>Blood, serum, plasma, urine, sputum, feces, colonic lavage fluid</td>
</tr>
</tbody>
</table>
adenoma versus normal, colorectal cancer versus normal, IBS versus ulcerative colitis (UC) versus Crohn's disease (CD),

| Pregnancy related physiological states, conditions, or affiliated diseases: genetic risk, adverse pregnancy outcomes |
| Maternal serum, plasma, amniotic fluid, cord blood |

[00201] The methods of the invention can be used to characterize a phenotype using a blood sample or blood derivative. Blood derivatives include plasma and serum. Blood plasma is the liquid component of whole blood, and makes up approximately 55% of the total blood volume. It is composed primarily of water with small amounts of minerals, salts, ions, nutrients, and proteins in solution. In whole blood, red blood cells, leukocytes, and platelets are suspended within the plasma. Blood serum refers to blood plasma without fibrinogen or other clotting factors (i.e., whole blood minus both the cells and the clotting factors).

[00202] The biological sample may be obtained through a third party, such as a party not performing the analysis of the biomarkers, whether direct assessment of a biological sample or by profiling one or more vesicles obtained from the biological sample. 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. Alternatively, the biological sample may obtained by the same party analyzing the vesicle. In addition, biological samples be assayed, are archived (e.g., frozen) or otherwise stored in under preservative conditions.

[00203] Furthermore, a biological sample can comprise a vesicle or cell membrane fragment that is derived from a cell of origin and available extracellularly in a subject's biological fluid or extracellular milieu.

[00204] Methods of the invention can include assessing one or more vesicles, including assessing vesicle populations. A vesicle, as used herein, is a membrane vesicle that is shed from cells. Vesicles or membrane vesicles include without limitation: circulating microvesicles (cMVs), microvesicle, exosome, nanovesicle, dexosome, bleb, blebbly, 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. Furthermore, although vesicles may be produced by different cellular processes, the methods of the invention are not limited to or reliant on any one mechanism, insofar as such vesicles are present in a biological sample and are capable of being characterized by the methods disclosed herein. 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 solute 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 2:

### Table 2: Vesicle Properties

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

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

[00205] 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

-58-
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 used 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
tissue-specific gene expression and regulation. BMC Bioinformatics. 9:271; TissueDistributionDBs,
available at genome.dkfz-heidelberg.de/menu/tissue_db/index.html.

A vesicle can have a diameter of greater than about 10 nm, 20 nm, or 30 nm. A vesicle can have a
diameter of greater than 40 nm, 50 nm, 100 nm, 200 nm, 500 nm, 1000 nm, 1500 nm, 2000 nm or greater
than 10,000 nm. A vesicle can have a diameter of about 20-2000 nm, about 20-1500 nm, 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, 2000 nm, 1500 nm, 1000 nm, 800 nm, 500 nm, 200 nm, 100 nm, 50 nm,
40 nm, 30 nm, 20 nm or less than 10 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 2. Vesicles can be
assessed to measure the diameter of a single vesicle or any number of vesicles. For example, the range of
diameters of a vesicle population or an average diameter of a vesicle population can be determined.
Vesicle diameter can be assessed using methods known in the art, e.g., imaging technologies such as
electron microscopy. In an embodiment, a diameter of one or more vesicles is determined using optical
particle detection. See, e.g., U.S. Patent 7,751,053, entitled "Optical Detection and Analysis of Particles"
and issued July 6, 2010; and U.S. Patent 7,399,600, entitled "Optical Detection and Analysis of Particles"
and issued July 15, 2010.

In some embodiments, the methods of the invention comprise assessing vesicles directly such as
in a biological sample without prior isolation, purification, or concentration from the biological sample.
For example, the amount of vesicles in the sample can by itself provide a biosignature that provides a
diagnostic, prognostic or theranostic determination. Alternatively, the vesicle in the sample may be
isolated, captured, purified, or concentrated from a sample prior to analysis. As noted, isolation, capture or
purification as used herein comprises partial isolation, partial capture or partial purification apart from
other components in the sample. Vesicle 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. FIGs. 19B-C present an overview of a method of the
invention for assessing microvesicles using an aptamer pool.

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 and/or cellular 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. For example, positive detection of colon derived vesicles expressing tumor antigens can indicate that the patient has colorectal cancer. As such, methods of the invention can be used to characterize any disease or condition associated with an anatomical or cellular origin, by assessing, for example, disease-specific and cell-specific biomarkers of one or more vesicles obtained from a subject.

[00209] In another embodiment, the methods of the invention comprise assessing one or more vesicle payload to provide a phenotypic characterization. The payload with a vesicle comprises any informative biological entity that can be detected as encapsulated within the vesicle, including without limitation proteins and nucleic acids, e.g., genomic or cDNA, mRNA, or functional fragments thereof, as well as microRNAs (miRs). In addition, methods of the invention are directed to detecting vesicle surface antigens (in addition or exclusive to vesicle payload) to provide a phenotypic characterization. For example, vesicles can be characterized by using binding agents (e.g., antibodies or aptamers) that are specific to vesicle surface antigens, and the bound vesicles can be further assessed to identify one or more payload components disclosed therein. 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. The biomarkers assessed can be present or absent, increased or reduced based on the selection of the desired target sample and comparison of the target sample to the desired reference sample. Non-limiting examples of target samples include: disease; treated/not-treated; different time points, such as a in a longitudinal study; and non-limiting examples of reference sample: non-disease; normal; different time points; and sensitive or resistant to candidate treatment(s).

Microvesicle Isolation and Analysis

[00210] Sample Processing

[00211] A vesicle or a population of vesicles may be isolated, purified, concentrated or otherwise enriched prior to and/or during analysis. Unless otherwise specified, the terms "purified," "isolated," or similar as used herein in reference to vesicles or biomarker components are intended to include partial or complete purification or isolation of such components from a cell or organism. Analysis of a vesicle can include quantitating 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 can be isolated from a heterogeneous population of vesicles and quantitated. Analysis of a vesicle can also include detecting, quantitatively or qualitatively, one or more particular biomarker profile or biosignature of a vesicle, as described herein.
A vesicle can be stored and archived, such as in a bio-fluid bank and retrieved for analysis as desired. 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 can 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 vesicle for analysis.

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.

Size exclusion chromatography, such as gel permeation columns, centrifugation or density gradient centrifugation, and filtration methods can be used. For example, a vesicle 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.

Highly abundant proteins, such as albumin and immunoglobulin in blood samples, may hinder isolation of vesicles from a biological sample. For example, a vesicle can be isolated from a biological sample using a system that uses multiple antibodies that are specific to the most abundant proteins found in a biological sample, such as 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. 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 USA. 2004;101:13368-13373.

Plasma contains a large variety of proteins including albumin, immunoglobulins, and clotting proteins such as fibrinogen. About 60% of plasma protein comprises the protein albumin (e.g., human serum albumin or HSA), which contributes to osmotic pressure of plasma to assist in the transport of lipids and steroid hormones. Globulins make up about 35% of plasma proteins and are used in the transport of ions, hormones and lipids assisting in immune function. About 4% of plasma protein comprises fibrinogen which is essential in the clotting of blood and can be converted into the insoluble protein fibrin. Other types of blood proteins include: Prealbumin, Alpha 1 antitrypsin, Alpha 1 acid
glycoprotein, Alpha 1 fetoprotein, Haptoglobin, Alpha 2 macroglobulin, Ceruloplasmin, Transferrin, complement proteins C3 and C4, Beta 2 microglobulin, Beta lipoprotein, Gamma globulin proteins, C-reactive protein (CRP), Lipoproteins (chylomicrons, VLDL, LDL, HDL), other globulins (types alpha, beta and gamma), Prothrombin and Mannose-binding lectin (MBL). Any of these proteins, including classes of proteins, or derivatives thereof (such as fibrin which is derived from the cleavage of fibrinogen) can be selectively depleted from a biological sample prior to further analysis performed on the sample. Without being bound by theory, removal of such background proteins may facilitate more sensitive, accurate, or precise detection of the biomarkers of interest in the sample.

Abundant proteins in blood or blood derivatives (e.g., plasma or serum) include without limitation albumin, IgG, transferrin, fibrinogen, IgA, α2-Macro globulin, IgM, α-Antitrypsin, complement C3, haptoglobin, apolipoprotein Al, apolipoprotein A3, apolipoprotein B, α1-Acid Glycoprotein, ceruloplasmin, complement C4, C1q, IgD, prealbumin ( transthyretin), and plasminogen. Such proteins can be depleted using commercially available columns and kits. Examples of such columns comprise the Multiple Affinity Removal System from Agilent Technologies (Santa Clara, CA). This system include various cartridges designed to deplete different protein profiles, including the following cartridges with performance characteristics according to the manufacturer: Human 14, which eliminates approximately 94% of total protein (albumin, IgG, antitrypsin, IgA, transferrin, haptoglobin, fibrinogen, alpha2-macroglobulin, alpha 1-acid glycoprotein (orosomucoid), IgM, apolipoprotein Al, apolipoprotein All, complement C3 and transthyretin); Human 7, which eliminates approximately 85 - 90% of total protein (albumin, IgG, IgA, transferrin, haptoglobin, antitrypsin, and fibrinogen); Human 6, which eliminates approximately 85 - 90% of total protein (albumin, IgG, IgA, transferrin, haptoglobin, and antitrypsin); Human Albumin/IgG, which eliminates approximately 69% of total protein (albumin and IgG); and Human Albumin, which eliminates approximately 50-55% of total protein (albumin). The ProteoPrep® 20 Plasma Immunodepletion Kit from Sigma-Aldrich is intended to specifically remove the 20 most abundant proteins from human plasma or serum, which is about remove 97-98% of the total protein mass in plasma or serum (Sigma-Aldrich, St. Louis, MO). According to the manufacturer, the ProteoPrep® 20 removes: albumin, IgG, transferrin, fibrinogen, IgA, α2-Macroglobulin, IgM, α1-Antitrypsin, complement C3, haptoglobin, apolipoprotein Al, A3 and B; α1-Acid Glycoprotein, ceruloplasmin, complement C4, C1q; IgD, prealbumin, and plasminogen. Sigma-Aldrich also manufactures ProteoPrep® columns to remove albumin (HSA) and immunoglobulins (IgG). The ProteomeLab IgY-12 High Capacity Proteome Partitioning kits from Beckman Coulter (Fullerton, CA) are specifically designed to remove twelve highly abundant proteins (Albumin, IgG, Transferrin, Fibrinogen, IgA, a2-macroglobulin, IgM, a 1-Antitrypsin, Haptoglobin, Orosomucoid, Apolipoprotein A-I, Apolipoprotein A-II) from the human biological fluids such as serum and plasma. Generally, such systems rely on immunodepletion to remove the target proteins, e.g., using small ligands and/or full antibodies. The PureProteome™ Human Albumin/Immunoglobulin Depletion Kit from Millipore (EMD Millipore Corporation, Billerica, MA, USA) is a magnetic bead based kit that enables high depletion efficiency (typically >99%) of Albumin and all Immunoglobulins (i.e., IgG, IgA, IgM, IgE and IgD) from human serum or plasma samples. The
ProteoExtract® Albumin/IgG Removal Kit, also from Millipore, is designed to deplete >80% of albumin and IgG from body fluid samples. Other similar protein depletion products include without limitation the following: Aurum™Affi-Gels Blue mini kit (Bio-Rad, Hercules, CA, USA); Vivapure® anti-HSA/IgG kit (Sartorius Stedim Biotech, Goettingen, Germany), Qproteome albumin/IgG depletion kit (Qiagen, Hilden, Germany); Seppro® MIXED 12-LC20 column (GenWay Biotech, San Diego, CA, USA); Abundant Serum Protein Depletion Kit (Norgen Biotek Corp., Ontario, Canada); GBC Human Albumin/IgG/Transferrin 3 in 1 Depletion Column/Kit (Good Biotech Corp., Taiwan). These systems and similar systems can be used to remove abundant proteins from a biological sample, thereby improving the ability to detect low abundance circulating biomarkers such as proteins and vesicles.

Thromboplastin is a plasma protein aiding blood coagulation through conversion of prothrombin to thrombin. Thrombin in turn acts as a serine protease that converts soluble fibrinogen into insoluble strands of fibrin, as well as catalyzing many other coagulation-related reactions. Thus, thromboplastin is a protein that can be used to facilitate precipitation of fibrinogen/fibrin (blood clotting factors) out of plasma. In addition to or as an alternative to immunoaffinity protein removal, a blood sample can be treated with thromboplastin to deplete fibrinogen/fibrin. Thromboplastin removal can be performed in addition to or as an alternative to immunoaffinity protein removal as described above using methods known in the art. Precipitation of other proteins and/or other sample particulate can also improve detection of circulating biomarkers such as vesicles in a sample. For example, ammonium sulfate treatment as known in the art can be used to precipitate immunoglobulins and other highly abundant proteins.

In an embodiment, the invention provides a method of detecting a presence or level of one or more circulating biomarker such as a microvesicle in a biological sample, comprising: (a) providing a biological sample comprising or suspected to comprise the one or more circulating biomarker; (b) selectively depleting one or more abundant protein from the biological sample provided in step (a); (c) performing affinity selection of the one or more circulating biomarker from the sample depleted in step (b), thereby detecting the presence or level of one or more circulating biomarker. The biological sample may comprise a bodily fluid, e.g., peripheral blood, sera, plasma, ascites, urine, cerebrospinal fluid (CSF), sputum, saliva, bone marrow, synovial fluid, aqueous humor, amniotic fluid, cerumen, breast milk, bronchealveolar 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 water, pancreatic juice, lavage fluids from sinus cavities, bronchopulmonary aspirates, blastocyst cavity fluid, umbilical cord blood, or a derivative of any thereof. In some embodiments, the biological sample comprises peripheral blood, serum or plasma. Illustrative protocols and results from selectively depleting one or more abundant protein from blood plasma prior to vesicle detection can be found in Example 40 of International Patent Publication No. WO/2014/082083, filed November 26, 2013, which patent publication is incorporated by reference herein in its entirety.

An abundant protein may comprise a protein in the sample that is present in the sample at a high enough concentration to potentially interfere with downstream processing or analysis. Typically, an
abundant protein is not the target of any further analysis of the sample. The abundant protein may constitute at least 10^3, 10^4, 10^5, 0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 96, 97, 98 or at least 99% of the total protein mass in the sample. In some embodiments, the abundant protein is present at less than 10^-5% of the total protein mass in the sample, e.g., in the case of a rare target of interest. As described herein, in the case of blood or a derivative thereof, the one or more abundant protein may comprise one or more of albumin, IgG, transferrin, fibrinogen, fibrin, IgA, a2-Marcroglobulin, IgM, a-antitrypsin, complement C3, haptoglobin, apolipoprotein A1, A3 and B; a-1-Acid Glycoprotein, ceruloplasmin, complement C4, C1q, IgD, prealbumin (transthyretin), plasminogen, a derivative of any thereof, and a combination thereof. The one or more abundant protein in blood or a blood derivative may also comprise one or more of Albumin, Immunoglobulins, Fibrinogen, Prealbumin, Alpha 1 antitrypsin, Alpha 1 acid glycoprotein, Alpha 1 fetoprotein, Haptoglobin, Alpha 2 macroglobulin, Ceruloplasmin, Transferrin, complement proteins C3 and C4, Beta 2 microglobulin, Beta lipoprotein, Gamma globulin proteins, C-reactive protein (CRP), Lipoproteins (chylomicrons, VLDL, LDL, HDL), other globulins (types alpha, beta and gamma), Prothrombin, Mannose-binding lectin (MBL), a derivative of any thereof, and a combination thereof.

[00221] In some embodiments, selectively depleting the one or more abundant protein comprises contacting the biological sample with thromboplastin to initiate precipitation of fibrin. The one or more abundant protein may also be depleted by immunoaffinity, precipitation, or a combination thereof. For example, the sample can be treated with thromboplastin to precipitate fibrin, and then the sample may be passed through a column to remove HSA, IgG, and other abundant proteins as desired.

[00222] "Selectively depleting" the one or more abundant protein comprises depleting the abundant protein from the sample at a higher percentage than depletion another entity in the sample, such as another protein or microvesicle, including a target of interest for downstream processing or analysis. Selectively depleting the one or more abundant protein may comprise depleting the abundant protein at a 1.1-fold, 1.2-fold, 1.3-fold, 1.4-fold, 1.5-fold, 1.6-fold, 1.7-fold, 1.8-fold, 1.9-fold, 2-fold, 3-fold, 4-fold, 5-fold, 6-fold, 7-fold, 8-fold, 9-fold, 10-fold, 11-fold, 12-fold, 13-fold, 14-fold, 15-fold, 16-fold, 17-fold, 18-fold, 19-fold, 20-fold, 25-fold, 30-fold, 40-fold, 50-fold, 60-fold, 70-fold, 80-fold, 90-fold, 100-fold, 200-fold, 300-fold, 400-fold, 500-fold, 600-fold, 700-fold, 800-fold, 900-fold, 1000-fold, 10^4-fold, 10^5-fold, 10^6-fold, 10^7-fold, 10^8-fold, 10^9-fold, 10^10-fold, 10^11-fold, 10^12-fold, 10^13-fold, 10^14-fold, 10^15-fold, 10^16-fold, 10^17-fold, 10^18-fold, 10^19-fold, 10^20-fold, or higher rate than another entity in the sample, such as another protein or microvesicle, including a target of interest for downstream processing or analysis. In an embodiment, there is little to no observable depletion of the target of interest as compared to the depletion of the abundant protein. In some embodiments, selectively depleting the one or more abundant protein from the biological sample comprises depleting at least 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% of the one or more abundant protein.
Removal of highly abundant proteins and other non-desired entities can further be facilitated with a non-stringent size exclusion step. For example, the sample can be processed using a high molecular weight cutoff size exclusion step to preferentially enrich high molecular weight vesicles apart from lower molecular weight proteins and other entities. In some embodiments, a sample is processed with a column (e.g., a gel filtration column) or filter having a molecular weight cutoff (MWCO) of 500, 600, 700, 800, 900, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, 5000, 5500, 6000, 6500, 7000, 7500, 8000, 8500, 9000, 9500, 10000, or greater than 10000 kiloDaltons (kDa). In an embodiment, a 700 kDa filtration column is used. In such a step, the vesicles will be retained or flow more slowly than the column or filter than the lower molecular weight entities. Such columns and filters are known in the art.

Isolation or enrichment of a vesicle from a biological sample can also be enhanced by use of sonication (for example, by applying ultrasound), detergents, other membrane-activating 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 a 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.

With methods of detecting circulating biomarkers as described here, e.g., antibody affinity isolation, the consistency of the results can be optimized as desired 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. Additives can be introduced at the various steps to improve the process, e.g., to control aggregation or degradation of the biomarkers of interest.

The results can also be optimized as desirable by treating the 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 albumin (BSA), milk or StabilGuard® (a BSA-free blocking agent; Product code SG02, Surmodics, Eden Prairie, MN), 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. Blocking agents for use with the invention comprise inert proteins, e.g., milk proteins, non-fat dry milk protein, albumin, BSA, casein, or serum such as newborn calf serum (NBCS), goat serum, rabbit serum or salmon serum. The proteins can be added at a 0.1% to 10% concentration, e.g., 1%, 2%, 3%, 3.5%, 4%, 5%, 6%, 7%, 8%, 9% or 10% concentration. In some embodiments, BSA is added to 0.1% to 10% concentration, e.g., 1%, 2%, 3%, 3.5%, 4%, 5%, 6%, 7%, 8%, 9% or 10% concentration. In an embodiment, the sample is treated according to the methodology presented in U.S. Patent Application 11/632946, filed July 13, 2005, which application is incorporated herein by reference in its entirety. Commercially available blockers may be used, such as SuperBlock, StartingBlock, Protein-Free from Pierce (a division of Thermo Fisher Scientific, Rockford, IL). 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.

[00227] The methods of detecting vesicles and other circulating biomarkers 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/or F68. Such treatments may reduce the formation of large aggregates or protein or other biological debris and thus provide a more consistent detection reading.

[00228] **Filtration and Ultrafiltration**

[00229] A vesicle can be isolated from a biological sample by filtering a biological sample from a subject through a filtration module and collecting from the filtration module a retentate comprising the vesicle, thereby isolating the vesicle from the biological sample. The method can comprise filtering a biological sample from a subject through a filtration module comprising a filter (also referred to herein as a selection membrane); and collecting from the filtration module a retentate comprising the vesicle, thereby isolating the vesicle from the biological sample. For example, in one embodiment, the filter retains molecules greater than about 100 kiloDaltons. In such cases, microvesicles are generally found within the retentate of the filtration process whereas smaller entities such as proteins, protein complexes, nucleic acids, etc, pass through into the filtrate.

[00230] The method can be used when determining a biosignature of one or more microvesicle. The method can also further comprise contacting the retentate from the filtration to a plurality of substrates, wherein each substrate is coupled to one or more capture agents, and each subset of the plurality of substrates comprises a different capture agent or combination of capture agents than another subset of the plurality of substrates.
Also provided herein is a method of determining a biosignature of a vesicle in a sample comprising: filtering a biological sample from a subject with a disorder through a filtration module, collecting from the filtration module a retentate comprising one or more vesicles, and determining a biosignature of the one or more vesicles. In one embodiment, the filtration module comprises a filter that retains molecules greater than about 100 or 150 kiloDaltons.

The method disclosed herein can further comprise characterizing a phenotype in a subject by filtering a biological sample from a subject through a filtration module, collecting from the filtration module a retentate comprising one or more vesicles; detecting a biosignature of the one or more vesicles; and characterizing a phenotype in the subject based on the biosignature, wherein characterizing is with at least 70% sensitivity. In some embodiments, characterizing comprises determining an amount of one or more vesicle having the biosignature. Furthermore, the characterizing can be from about 80% to 100% sensitivity.

Also provided herein is a method for multiplex analysis of a plurality of vesicles. In some embodiments, the method comprises filtering a biological sample from a subject through a filtration module; collecting from the filtration module a retentate comprising the plurality of vesicles, applying the plurality of vesicles to a plurality of capture agents, wherein the plurality of capture agents is coupled to a plurality of substrates, and each subset of the plurality of substrates is differentially labeled from another subset of the plurality of substrates; capturing at least a subset of the plurality of vesicles; and determining a biosignature for at least a subset of the captured vesicles. In one embodiment, each substrate is coupled to one or more capture agents, and each subset of the plurality of substrates comprises a different capture agent or combination of capture agents as compared to another subset of the plurality of substrates. In some embodiments, at least a subset of the plurality of substrates is intrinsically labeled, such as comprising one or more labels. The substrate can be a particle or bead, or any combination thereof. In some embodiments, the filter retains molecules greater than 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 250, 300, 400, 500, 600, 700, 800, 900, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, 5000, 5500, 6000, 6500, 7000, 7500, 8000, 8500, 9000, 9500, 10000, or greater than 10000 kiloDaltons (kDa). In one embodiment, the filtration module comprises a filter that retains molecules greater than about 100 or 150 kiloDaltons. In one embodiment, the filtration module comprises a filter that retains molecules greater than about 9, 20, 100 or 150 kiloDaltons. In still another embodiment, the filtration module comprises a filter that retains molecules greater than about 7000 kDa.

In some embodiments, the method for multiplex analysis of a plurality of vesicles comprises filtering a biological sample from a subject through a filtration module, wherein the filtration module comprises a filter that retains molecules greater than about 100 kiloDaltons; collecting from the filtration module a retentate comprising the plurality of vesicles; applying the plurality of vesicles to a plurality of capture agents, wherein the plurality of capture agents is coupled to a microarray; capturing at least a subset of the plurality of vesicles on the microarray; and determining a biosignature for at least a subset of the captured vesicles. In some embodiments, the filter retains molecules greater than 1, 2, 3, 4, 5, 6, 7, 8,
The biological sample can be clarified prior to isolation by filtration. Clarification comprises selective removal of cellular debris and other undesirable materials. For example, cellular debris and other components that may interfere with detection of the circulating biomarkers can be removed. The clarification can be by low-speed centrifugation, such as at about 5,000x g, 4,000x g, 3,000x g, 2,000x g, 1,000x g, or less. The supernatant, or clarified biological sample, containing the vesicle can then be collected and filtered to isolate the vesicle from the clarified biological sample. In some embodiments, the biological sample is not clarified prior to isolation of a vesicle by filtration.

In some embodiments, isolation of a vesicle from a sample does not use high-speed centrifugation, such as ultracentrifugation. For example, isolation may not require the use of centrifugal speeds, such as about 100,000x g or more. In some embodiments, isolation of a vesicle from a sample uses speeds of less than 50,000 x g, 40,000 x g, 30,000 x g, 20,000 x g, 15,000 x g, 12,000 x g, or 10,000 x g.

Any number of applicable filter configurations can be used to filter a sample of interest. In some embodiments, the filtration module used to isolate the circulating biomarkers from the biological sample is a fiber-based filtration cartridge. For example, the fiber can be a hollow polymeric fiber, such as a polypropylene hollow fiber. A biological sample can be introduced into the filtration module by pumping the sample fluid, such as a biological fluid as disclosed herein, into the module with a pump device, such as a peristaltic pump. The pump flow rate can vary, such as at about 0.25, 0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5, 6, 7, 8, 9, or 10 mL/minute. The flow rate can be adjusted given the configuration, e.g., size and throughput, of the filtration module.

The filtration module can be a membrane filtration module. For example, the membrane filtration module can comprise a filter disc membrane, such as a hydrophilic polyvinylidine difluoride (PVDF) filter disc membrane housed in a stirred cell apparatus (e.g., comprising a magnetic stirrer). In some embodiments, the sample moves through the filter as a result of a pressure gradient established on either side of the filter membrane.

The filter can comprise a material having low hydrophobic absorptivity and/or high hydrophilic properties. For example, the filter can have an average pore size for vesicle retention and permeation of most proteins as well as a surface that is hydrophilic, thereby limiting protein adsorption. For example, the filter can comprise a material selected from the group consisting of polypropylene, PVDF, polyethylene, polyfluoroethylene, cellulose, secondary cellulose acetate, polyvinylalcohol, and ethylenevinyl alcohol
Additional materials that can be used in a filter include, but are not limited to, polysulfone and polyethersulfone.

[00240] The filtration module can have a filter that retains molecules greater than about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 250, 300, 400, 500, 600, 700, 800, or 900 kilodaltons (kDa), such as a filter that has a MWCO (molecular weight cut off) of about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 250, 300, 400, 500, 600, 700, 800, or 900 kDa, respectively. In embodiments, the filtration module has a MWCO of 1000 kDa, 1500 kDa, 2000 kDa, 2500 kDa, 3000 kDa, 3500 kDa, 4000 kDa, 4500 kDa, 5000 kDa, 5500 kDa, 6000 kDa, 6500 kDa, 7000 kDa, 7500 kDa, 8000 kDa, 8500 kDa, 9000 kDa, 9500 kDa, 10000 kDa, or greater than 10000 kDa. Ultrafiltration membranes with a range of MWCO of 9 kDa, 20 kDa and/or 150 kDa can be used. In some embodiments, the filter within the filtration module has an average pore diameter of about 0.01 µm to about 0.15 µm, and in some embodiments from about 0.05 µm to about 0.12 µm. In some embodiments, the filter has an average pore diameter of about 0.06 µm, 0.07 µm, 0.08 µm, 0.09 µm, 0.1 µm, 0.11 µm or 0.2 µm.

[00241] The filtration module can be a commercially available column, such as a column typically used for concentrating proteins or for isolating proteins (e.g., ultrafiltration). Examples include, but are not limited to, columns from Millpore (Billerica, MA), such as Amicon® centrifugal filters, or from Pierce® (Rockford, IL), such as Pierce Concentrator filter devices. Useful columns from Pierce include disposable ultrafiltration centrifugal devices with a MWCO of 9 kDa, 20 kDa and/or 150 kDa. These concentrators consist of a high-performance regenerated cellulose membrane welded to a conical device. The filters can be as described in U.S. Patents 6,269,957 or 6,357,601, both of which applications are incorporated by reference in their entirety herein.

[00242] The retentate comprising the isolated vesicle can be collected from the filtration module. The retentate can be collected by flushing the retentate from the filter. Selection of a filter composition having hydrophilic surface properties, thereby limiting protein adsorption, can be used, without being bound by theory, for easier collection of the retentate and minimize use of harsh or time-consuming collection techniques.

[00243] The collected retentate can then be used subsequent analysis, such as assessing a biosignature of one or more vesicles in the retentate, as further described herein. The analysis can be directly performed on the collected retentate. Alternatively, the collected retentate can be further concentrated or purified, prior to analysis of one or more vesicles. For example, the retentate can be further concentrated or vesicles further isolated from the retentate using size exclusion chromatography, density gradient centrifugation, differential centrifugation, immunoabsorbent capture, affinity purification, microfluidic separation, or combinations thereof, such as described herein. In some embodiments, the retentate can undergo another step of filtration. Alternatively, prior to isolation of a vesicle using a filter, the vesicle is concentrated or isolated using techniques including without limitation size exclusion chromatography, density gradient centrifugation, differential centrifugation, immunoabsorbent capture, affinity purification, microfluidic separation, or combinations thereof.
[00244] Combinations of filters can be used for concentrating and isolating biomarkers. For example, the biological sample may first be filtered through a filter having a porosity or pore size of between about 0.01 μm to about 10 μm, e.g., 0.01 μm to about 2 μm or about 0.05 μm to about 1.5 μm, and then the sample is filtered. For example, prior to filtering a biological sample through a filtration module with a filter that retains molecules greater than about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 250, 300, 400, 500, 600, 700, 800, 900, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, 5000, 5500, 6000, 6500, 7000, 7500, 8000, 8500, 9000, 9500, 10000, or greater than 10000 kiloDaltons (kDa), such as a filter that has a MWCO (molecular weight cut off) of about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 250, 300, 400, 500, 600, 700, 800, 900, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, 5000, 5500, 6000, 6500, 7000, 7500, 8000, 8500, 9000, 9500, 10000, or greater than 10000 kDa, respectively, the biological sample may first be filtered through a filter having a porosity or pore size of between about 0.01 μm to about 10 μm, e.g., 0.01 μm to about 2 μm or about 0.05 μm to about 1.5 μm. The module arises. In some embodiments, the filter has a pore size of between about 0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9 or 2.0, 3.0, 4.0, 5.0, 6.0, 7.0, 8.0, 9.0 or 10.0 μm. The filter may be a syringe filter. Thus, in one embodiment, the method comprises filtering the biological sample through a filter, such as a syringe filter, wherein the syringe filter has a porosity of greater than about 1 μm, prior to filtering the sample through a filtration module comprising a filter that retains molecules greater than about 100 or 150 kiloDaltons. In an embodiment, the filter is 1.2 μm filter and the filtration is followed by passage of the sample through a 7 ml or 20 ml concentrator column with a 150 kDa cutoff. Multiple concentrator columns may be used, e.g., in series. For example, a 7000 MWCO filtration unit can be used before a 150 MWCO unit.

[00245] The filtration module can be a component of a microfluidic device. 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 biomarkers, and other processes, such as further described herein.


[00247] A microfluidic device can also be used for isolation of a vesicle by comprising a filtration module. For example, a microfluidic device can use one more channels for isolating a vesicle from a biological sample based on size from a biological sample. A biological sample can be introduced into one or more microfluidic channels, which selectively allows the passage of vesicles. The microfluidic device can further comprise binding agents, or more than one filtration module to select vesicles based on a property of the vesicles, for example, size, shape, deformability, biomarker profile, or biosignature.
The retentate from a filtration step can be further processed before assessment of microvesicles or other biomarkers therein. In an embodiment, the retentate is diluted prior to biomarker assessment, e.g., with an appropriate diluent such as a biologically compatible buffer. In some cases, the retentate is serially diluted. In an aspect, the invention provides a method for detecting a microvesicle population from a biological sample comprising: a) concentrating the biological sample using a selection membrane having a pore size of from 0.01 µm to about 10 µm, or a molecular weight cut off (MWCO) from about 1 kDa to 10000 kDa; b) diluting a retentate from the concentration step into one or more aliquots; and c) contacting each of the one or more aliquots of retentate with one or more binding agent specific to a molecule of at least one microvesicle in the microvesicle population. In a related aspect, the invention provides a method for detecting a microvesicle population from a biological sample comprising: a) concentrating the biological sample using a selection membrane having a pore size of from 0.01 µm to about 10 µm, or a molecular weight cut off (MWCO) from about 1 kDa to 10000 kDa; and b) contacting one or more aliquots of the retentate from the concentrating step with one or more binding agent specific to a molecule of at least one microvesicle in the microvesicle population.

The selection membrane can be sized to retain the desired biomarkers in the retentate or to allow the desired biomarkers to pass through the filter into the filtrate. The filter membrane can be chosen to have a certain pore size or MWCO value. The selection membrane can have a pore size of about 0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9 or 2.0, 3.0, 4.0, 5.0, 6.0, 7.0, 8.0, 9.0 or 10.0 µm. The selection membrane can also have a MWCO of about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 250, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000 or 10000 kDa.

The retentate can be separated and/or diluted into any number of desired aliquots. For example, multiple aliquots without any dilution or the same dilution can be used to determine reproducibility. In another example, multiple aliquots at different dilutions can be used to construct a concentration curve. In an embodiment, the retentate is separated and/or diluted into at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150, 200, 250, 300, 350 or 400 aliquots. The aliquots can be at a same dilution or at different dilutions.

A dilution factor is the ratio of the final volume of a mixture (the mixture of the diluents and aliquot) divided by the initial volume of the aliquot. The retentate can be diluted into one or more aliquots at a dilution factor of about 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 250, 300, 400, 500, 600, 700, 800, 900, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, 5000, 5500, 6000, 6500, 7000, 7500, 8000, 8500, 9000, 9500, 10000, 20000, 30000, 40000, 50000, 60000, 70000, 80000, 90000 and/or 100000. For example, the retentate can be diluted into one or more aliquot at a dilution factor of about 500.

To estimate a concentration or form a curve, the retentate can be diluted into multiple aliquots. In an embodiment of the method, the retentate is diluted into one or more aliquots at a dilution factor of about 100, 250, 500, 1000, 10000 and 100000. As desired, the method can further comprise detecting an
amount of microvesicles in each aliquot of retentate, e.g., that formed a complex with the one or more binding agent. The curve can be used to determine a linear range of the amount of microvesicles in each aliquot detected versus dilution factor. A concentration of the detected microvesicles for the biological sample can be determined using the amount of microvesicles determined in one or more aliquot within the linear range. The concentration can be compared to a reference concentration, e.g., in order to characterize a phenotype as described herein.

[00253] The invention also provides a related method comprising filtering a biological sample from a subject through a filtration module and collecting a filtrate comprising the vesicle, thereby isolating the vesicle from the biological sample. In such cases cells and other large entities can be retained in the retentate while microvesicles pass through into the filtrate. It will be appreciated that strategies to retain and filter microvesicles can be used in concert. For example, a sample can be filtered with a selection membrane that allows microvesicles to pass through, thereby isolating the microvesicles from large particles (cells, complexes, etc). The filtrate comprising the microvesicle can then be filtered using a selection membrane that retains microvesicles, thereby isolating the microvesicles from smaller particles (proteins, nucleic acids, etc). The isolated microvesicles can be further assessed according to the methods of the invention, e.g., to characterize a phenotype.

[00254] **Precipitation**

[00255] Vesicles can be isolated using a polymeric precipitation method. The method can be in combination with or in place of the other isolation methods described herein. In one embodiment, the sample containing the vesicles is contacted with a formulation of polyethylene glycol (PEG). The polymeric formulation is incubated with the vesicle containing sample then precipitated by centrifugation. The PEG can bind to the vesicles and can be treated to specifically capture vesicles by addition of a capture moiety, e.g., a pegylated-binding protein such as an antibody. One of skill will appreciate that other polymers in addition to PEG can be used, e.g., PEG derivatives including methoxypolyethylene glycols, poly (ethylene oxide), and various polymers of formula HO-CH₂(CH₂-O-CH₂)n-CH₃-OH having different molecular weights. The efficiency of isolation may depend on various factors including the length of the polymer chains and concentration of polymer used. In preferred embodiments, PEG4000 or PEG 8000 may be used at a concentration of 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, or 10%, e.g., 4% or 8%.

[00256] In some embodiments of the invention, the vesicles are concentrated from a sample using the polymer precipitation method and the isolated vesicles are further separated using another approach. The second step can be used to identify a subpopulation of vesicles, e.g., that display certain biomarkers. The second separation step can comprise size exclusion, a binding agent, an antibody capture step, microbeads, as described herein.

[00257] In an embodiment, vesicles are isolated according to the ExoQuick™ and ExoQuick-TC™ kits from System Biosciences, Mountain View, CA USA. These kits use a polymer-based precipitation method to pellet vesicles. Similarly, the vesicles can be isolated using the Total Exosome Isolation (from Serum) or Total Exosome Isolation (from Cell Culture Media) kits from Invitrogen / Life Technologies
The Total Exosome Isolation reagent forces less-soluble components such as vesicles out of solution, allowing them to be collected by a short, low-speed centrifugation. The reagent is added to the biological sample, and the solution is incubated overnight at 2 °C to 8 °C. The precipitated vesicles are recovered by standard centrifugation.

[00258] **Binding Agents**

[00259] Binding agents (also referred to as binding reagents) include agents that are capable of binding a target biomarker. A binding agent can be specific for the target biomarker, meaning the agent is capable of binding a target biomarker. The target can be any useful biomarker disclosed herein, such as a biomarker on the vesicle surface. In some embodiments, the target is a single molecule, such as a single protein, so that the binding agent is specific to the single protein. In other embodiments, the target can be a group of molecules, such as a family or proteins having a similar epitope or moiety, so that the binding agent is specific to the family or group of proteins. The group of molecules can also be a class of molecules, such as protein, DNA or RNA. The binding agent can be a capture agent used to capture a vesicle by binding a component or biomarker of a vesicle. In some embodiments, a capture agent comprises an antibody or fragment thereof, or an aptamer, that binds to an antigen on a vesicle. The capture agent can be optionally coupled to a substrate and used to isolate a vesicle, as further described herein.

[00260] A binding agent is an agent that binds to a circulating biomarker, such as a vesicle or a component of a vesicle. The binding agent can be used as a capture agent and/or a detection agent. A capture agent can bind and capture a circulating biomarker, such as by binding a component or biomarker of a vesicle. For example, the capture agent can be a capture antibody or capture antigen that binds to an antigen on a vesicle. A detection agent can bind to a circulating biomarker thereby facilitating detection of the biomarker. For example, a capture agent comprising an antibody or aptamer that is sequestered to a substrate can be used to capture a vesicle in a sample, and a detection agent comprising an antibody or aptamer that carries a label can be used to detect the captured vesicle via detection of the detection agent's label. In some embodiments, a vesicle is assessed using capture and detection agents that recognize the same vesicle biomarkers. For example, a vesicle population can be captured using a tetraspanin such as by using an anti-CD9 antibody bound to a substrate, and the captured vesicles can be detected using a fluorescently labeled anti-CD9 antibody to label the captured vesicles. In other embodiments, a vesicle is assessed using capture and detection agents that recognize different vesicle biomarkers. For example, a vesicle population can be captured using a cell-specific marker such as by using an anti-PCS9 antibody bound to a substrate, and the captured vesicles can be detected using a fluorescently labeled anti-CD9 antibody to label the captured vesicles. Similarly, the vesicle population can be captured using a general vesicle marker such as by using an anti-CD9 antibody bound to a substrate, and the captured vesicles can be detected using a fluorescently labeled antibody to a cell-specific or disease specific marker to label the captured vesicles.

[00261] The biomarkers recognized by the binding agent are sometimes referred to herein as an antigen. Unless otherwise specified, antigen as used herein is meant to encompass any entity that is capable of
being bound by a binding agent, regardless of the type of binding agent or the immunogenicity of the biomarker. The antigen further encompasses a functional fragment thereof. For example, an antigen can encompass a protein biomarker capable of being bound by a binding agent, including a fragment of the protein that is capable of being bound by a binding agent.

[00262] In one embodiment, a vesicle is captured using a capture agent that binds to a biomarker on a vesicle. The capture agent can be coupled to a substrate and used to isolate a vesicle, as further described herein. In one embodiment, a capture agent is used for affinity capture or isolation of a vesicle present in a substance or sample.

[00263] A binding agent can be used after a vesicle is concentrated or isolated from a biological sample. For example, a vesicle can first be isolated from a biological sample before a vesicle with a specific biosignature is isolated or detected. The vesicle with a specific biosignature can be isolated or detected using a binding agent for the biomarker. A vesicle with the specific biomarker can be isolated or detected from a heterogeneous population of vesicles. Alternatively, a binding agent may be used on a biological sample comprising vesicles without a prior isolation or concentration step. For example, a binding agent is used to isolate or detect a vesicle with a specific biosignature directly from a biological sample.

[00264] A binding agent can be a nucleic acid, protein, or other molecule that can bind to a component of a vesicle. The binding agent can comprise DNA, RNA, monoclonal antibodies, polyclonal antibodies, Fabs, Fab', single chain antibodies, synthetic antibodies, aptamers (DNA/RNA), peptoids, zDNA, peptide nucleic acids (PNAs), locked nucleic acids (LNAs), lectins, synthetic or naturally occurring chemical compounds (including but not limited to drugs, labeling reagents), dendrimers, or a combination thereof. For example, the binding agent can be a capture antibody. In embodiments of the invention, the binding agent comprises a membrane protein labeling agent. See, e.g., the membrane protein labeling agents disclosed in Alroy et al., US. Patent Publication US 2005/0 158708. In an embodiment, vesicles are isolated or captured as described herein, and one or more membrane protein labeling agent is used to detect the vesicles.

[00265] In some instances, a single binding agent can be employed to isolate or detect a vesicle. In other instances, a combination of different binding agents may be employed 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. Furthermore, the one or more different binding agents for a vesicle can form a biosignature of a vesicle, as further described below.

[00266] Different binding agents can also be used for multiplexing. For example, isolation or detection of more than one population of vesicles can be 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. In another embodiment, an array comprising different binding agents can be 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, such as described below. 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 Fig. 2 of International Patent Publication No. WO/201 1/127219, entitled "Circulating Biomarkers for Disease" and filed April 6, 201 1, which application is incorporated by reference in its entirety herein. 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.

[00267] 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 ellipsoidesporum 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 \( \alpha-1 \rightarrow 3 \) or \( \alpha-1 \rightarrow 6 \) mannose-mannose linkages.

[00268] 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).


[00270] 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,
In some embodiments, the capture agent is an antibody to CD9, CD63, CD81, PSMA, PCSA, B7H3, EpCam, PSCA, ICAM, STEAP, or EGFR. The capture agent can also be used to identify a biomarker of a vesicle. For example, a capture agent such as an antibody to CD9 would identify CD9 as a biomarker of the vesicle. In some embodiments, a plurality of capture agents can be used, such as in multiplex analysis. The plurality of captures agents can comprise binding agents to one or more of: CD9, CD63, CD81, PSMA, PCSA, B7H3, EpCam, PSCA, ICAM, STEAP, and EGFR. In some embodiments, the plurality of capture agents comprise binding agents to CD9, CD63, CD81, PSMA, PCSA, B7H3, MFG-E8, and/or EpCam. In yet other embodiments, the plurality of capture agents comprises binding agents to CD9, CD63, CD81, PSMA, PCSA, B7H3, EpCam, PSCA, ICAM, STEAP, and/or EGFR. The plurality of capture agents comprises binding agents to TMEM211, MFG-E8, Tissue Factor (TF), and/or CD24.

The antibodies referenced 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 Fab' 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.

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, Ca- methyl amino acids, and Na-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., R₁-CH₂-NH-R₂, where R₁ and R₂ 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-life 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 a-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. The term "protein" is also intended to be used interchangeably throughout this application with the terms "polypeptides" and "peptides" unless otherwise specified.

[00274] A vesicle may be isolated, captured or detected using a binding agent. The binding agent can be an agent that binds to a vesicle "housekeeping protein," or general vesicle biomarker. The biomarker can be CD63, CD9, CD81, CD82, CD37, CD53, Rab-5b, Annexin V, MFG-E8 or other commonly observed vesicle markers include those listed in Table 3. Furthermore, any of the markers disclosed herein or in Table 3 can be selected in identifying a candidate biosignature for a disease or condition, where the one or more selected biomarkers have a direct or indirect role or function in mechanisms involved in the disease or condition.

[00275] The binding agent can also be an agent that binds to a vesicle derived from a specific cell type, such as a tumor cell (e.g. binding agent for Tissue factor, EpCam, B7H3, RAGE or CD24) or a specific cell-of-origin. The binding agent used to isolate or detect a vesicle can be a binding agent for an antigen selected from Fig. 1 of International Patent Publication No. WO/201 1/127219, entitled "Circulating Biomarkers for Disease" and filed April 6, 2011, which application is incorporated by reference in its entirety herein. The binding agent for a vesicle can also be selected from those listed in Fig. 2 of International Patent Publication No. WO/201 1/127219. The binding agent can be for an antigen such as a tetraspanin, MFG-E8, Annexin V, 5T4, B7H3, caveolin, CD63, CD9, E-Cadherin, Tissue factor, MFG-E8, TMEM211, CD24, PSCA, PCSA, PSMA, Rab-5B, STEAP, TNFR1, CD81, EpCam, CD59, CD81, ICAM, EGFR, or CD66. A binding agent for a platelet can be a glycoprotein such as GpIa-IIa, GpIIb-IIIa, GpIIIb, GpIb, or GpIX. A binding agent can be for an antigen comprisine one or more of CD9, Erb2, Erb4, CD81, Erb3, MUC16, CD63, DLL4, HLA-Drpe, B7H3, IFNAR, 5T4, PCSA, MICB, PSMA, MFG-E8, Mucl, PSA, Muc2, Unc93a, VEGFR2, EpCAM, VEGF A, TMPRSS2, RAGE, PSCA, CD40, Mucl7, IL-17-RA, and CD80. For example, the binding agent can be one or more of CD9, CD63, CD81, B7H3, PCSA, MFG-E8, MUC2, EpCam, RAGE and Mucl7. One or more binding agents, such as one or more binding agents for two or more of the antigens, can be used for isolating or detecting a vesicle. The binding agent used can be selected based on the desire of isolating or detecting a vesicle derived from a particular cell type or cell-of-origin specific vesicle. The binding agent can be to one or more vesicle marker in Table 4.

[00276] 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, polytetrafluoroethylene (PTFE, 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 polypyrrole 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.

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 is 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.

Arrays typically contain addressable moieties that can detect the presence 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.

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. For example, latex beads, such as aldehyde/sulfate beads (Interfacial Dynamics, Portland, OR) can be used.

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.

A binding agent, such as an antibody, for isolating vesicles is preferably contacted with the biological sample comprising the vesicles of interest for at least a time sufficient for the binding agent to bind to a component of the vesicle. 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.

A binding agent, such as an antibody specific to an antigen listed in Fig. 1 of International Patent Publication No. WO/2011/127219, entitled “Circulating Biomarkers for Disease” and filed April 6, 2011, which application is incorporated by reference in its entirety herein, or a binding agent listed in Fig. 2 of International Patent Publication No. WO/2011/127219, can be labeled to facilitate 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. Various protein, radioactive, fluorescent, enzymatic, and other labels are described further above.

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.

Depending on the method of isolation or detection 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-hydroxy succinimide. 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.

**Particle-based Assays**

As an alternative to planar arrays, assays using particles or microspheres, such as bead based assays, are capable of use with a binding agent. For example, antibodies or aptamers are easily conjugated with commercially available beads. See, e.g., Fan et al., Illumina universal bead arrays. Methods Enzymol. 2006 410:57-73; Srinivas et al. Anal. Chem. 2011 Oct. 21, *Aptamerfunctionalized Microgel Particles for Protein Detection*; See also, review article on aptamers as therapeutic and diagnostic agents, Brody and Gold, Rev. Mol. Biotech. 2000, 74:5-13.

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 or aptamer 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 desired to generate custom bead arrays. Bead arrays are then incubated with the sample in a single reaction vessel to perform the assay.

Various particle/bead substrates and systems useful for the methods of the invention are described further above.

**Flow Cytometry**

In various embodiments of the invention, flow cytometry, which is described in further detail above, is used to assess a microvesicle population in a biological sample. If desired, the microvesicle
population can be sorted from other particles (e.g., cell debris, protein aggregates, etc) in a sample by labeling the vesicles using one or more general vesicle marker. The general vesicle marker can be a marker in Table 3. Commonly used vesicle markers include tetraspanins such as CD9, CD63 and/or CD81. Vesicles comprising one or more tetraspanin are sometimes refered to as "Tet+" herein to indicate that the vesicles are tetraspanin-positive. The sorted microvesicles can be further assessed using methodology described herein. E.g., surface antigens on the sorted microvesicles can be detected using flow or other methods. In some embodiments, payload within the sorted microvesicles is assessed. As an illustrative example, a population of microvesicles is contacted with a labeled binding agent to a surface antigen of interest, the contacted microvesicles are sorted using flow cytometry, and payload with the microvesicles is assessed. The payload may be polypeptides, nucleic acids (e.g., mRNA or microRNA) or other biological entities as desired. Such assessment is used to characterize a phenotype as described herein, e.g., to diagnose, prognose or theranose a cancer.

[00291] In an embodiment, flow sorting is used to distinguish microvesicle populations from other biological complexes. In a non-limiting example, Ago2+/Tet+ and Ago2+/Tet- particles are detected using flow methodology to separate Ago2+ vesicles from vesicle-free Ago2+ complexes, respectively.

[00292] Multiplexing

[00293] Multiplex experiments comprise experiments that can simultaneously measure multiple analytes in a single assay. Vesicles and associated biomarkers can be assessed in a multiplex fashion. Different binding agents can be used for multiplexing different circulating biomarkers, e.g., microRNA, protein, or vesicle populations. Different biomarkers, e.g., different vesicle populations, can be isolated or detected using different binding agents. Each population in a biological sample can be labeled with a different signaling label, such as a fluorophore, quantum dot, or radioactive label, such as described above. The label can be directly conjugated to a binding agent or indirectly used to detect a binding agent that binds a vesicle. 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.

[00294] 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 circulating biomarkers 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. In some cases, hundreds or thousands of vesicles are assessed in a single assay.

[00295] 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.

[00296] 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. 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. Biomarkers bound by different capture agents can be differentially detected using different labels.

[00297] A 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 may be used to generate different detection patterns 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.

[00298] In another embodiment, multiplex analysis can be performed using a planar substrate, wherein the 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.

[00299] **Binding Agents**

[00300] A vesicle may be isolated or detected using a binding agent for a novel component of a vesicle, such as an antibody for a novel antigen specific to a vesicle of interest. Novel antigens that are specific to a vesicle of interest may be isolated or identified using different test compounds of known composition bound to a substrate, such as an array or a plurality of particles, which can allow a large amount of chemical/structural space to be adequately sampled using only a small fraction of the space. The novel
antigen identified can also serve as a biomarker for the vesicle. For example, a novel antigen identified for a cell-of-origin specific vesicle can be a useful biomarker.

[00301] The term "agent" or "reagent" as used in respect to contacting a sample can mean any entity designed to bind, hybridize, associate with or otherwise detect or facilitate detection of a target molecule, including target polypeptides, peptides, nucleic acid molecules, leptins, lipids, or any other biological entity that can be detected as described herein or as known in the art. Examples of such agents/reagents are well known in the art, and include but are not limited to universal or specific nucleic acid primers, nucleic acid probes, antibodies, aptamers, peptoid, peptide nucleic acid, locked nucleic acid, lectin, dendrimer, chemical compound, or other entities described herein or known in the art.

[00302] A binding agent can be identified by screening either a homogeneous or heterogeneous vesicle population against test compounds. Since the composition of each test compound on the substrate surface is known, this constitutes a screen for affinity elements. For example, a test compound array comprises test compounds at specific locations on the substrate addressable locations, and can be used to identify one or more binding agents for a vesicle. The test compounds can all be unrelated or related based on minor variations of a core sequence or structure. The different test compounds may include variants of a given test compound (such as polypeptide isoforms), test compounds that are structurally or compositionally unrelated, or a combination thereof.

[00303] A test compound can be a peptoid, polysaccharide, organic compound, inorganic compound, polymer, lipids, nucleic acid, polypeptide, antibody, protein, polysaccharide, or other compound. The test compound can be natural or synthetic. The test compound can comprise or consist of linear or branched heteropolymeric compounds based on any of a number of linkages or combinations of linkages (e.g., amide, ester, ether, thiol, radical additions, metal coordination, etc.), dendritic structures, cavity structures or other structures with multiple nearby sites of attachment that serve as scaffolds upon which specific additions are made. Thes test compound can be spotted on a substrate or synthesized in situ, using standard methods in the art. In addition, the test compound can be spotted or synthesized in situ in combinations in order to detect useful interactions, such as cooperative binding.

[00304] The test compound can be a polypeptide with known amino acid sequence, thus, detection of a test compound binding with a vesicle can lead to identification of a polypeptide of known amino sequence that can be used as a binding agent. For example, a homogenous population of vesicles can be applied to a spotted array on a slide containing between a few and 1,000,000 test polypeptides having a length of variable amino acids. The polypeptides can be attached to the surface through the C-terminus. The sequence of the polypeptides can be generated randomly from 19 amino acids, excluding cysteine. The binding reaction can include a non-specific competitor, such as excess bacterial proteins labeled with another dye such that the specificity ratio for each polypeptide binding target can be determined. The polypeptides with the highest specificity and binding can be selected. The identity of the polypeptide on each spot is known, and thus can be readily identified. Once the novel antigens specific to the homogeneous vesicle population, such as a cell-of-origin specific vesicle is identified, such cell-of-origin specific vesicles may subsequently be isolated using such antigens in methods described hereafter.
[00305] An array can also be used for identifying an antibody as a binding agent for a vesicle. Test antibodies can be attached to an array and screened against a heterogeneous population of vesicles to identify antibodies that can be used to isolate or identify a vesicle. A homogeneous population of vesicles such as cell-of-origin specific vesicles can also be screened with an antibody array. Other than identifying antibodies to isolate or detect a homogeneous population of vesicles, one or more protein biomarkers specific to the homogenous population can be identified. Commercially available platforms with test antibodies pre-selected or custom selection of test antibodies attached to the array can be used. For example, an antibody array from Full Moon Biosystems can be screened using prostate cancer cell derived vesicles identifying antibodies to Bel-XL, ERCC1, Keratin 15, CD81/TAPA-1, CD9, Epithelial Specific Antigen (ESA), and Mast Cell Chymase as binding agents, and the proteins identified can be used as biomarkers for the vesicles. The biomarker can be present or absent, underexpressed or overexpressed, mutated, or modified in or on a vesicle and used in characterizing a condition.

[00306] An antibody or synthetic antibody to be used as a binding agent can also be identified through a peptide array. Another method is the use of synthetic antibody generation through antibody phage display. M13 bacteriophage libraries of antibodies (e.g. Fabs) are displayed on the surfaces of phage particles as fusions to a coat protein. Each phage particle displays a unique antibody and also encapsulates a vector that contains the encoding DNA. Highly diverse libraries can be constructed and represented as phage pools, which can be used in antibody selection for binding to immobilized antigens. Antigen-binding phages are retained by the immobilized antigen, and the nonbinding phages are removed by washing. The retained phage pool can be amplified by infection of an Escherichia coli host and the amplified pool can be used for additional rounds of selection to eventually obtain a population that is dominated by antigen-binding clones. At this stage, individual phase clones can be isolated and subjected to DNA sequencing to decode the sequences of the displayed antibodies. Through the use of phase display and other methods known in the art, high affinity designer antibodies for vesicles can be generated.

[00307] Bead-based assays can also be used to identify novel binding agents to isolate or detect a vesicle. A test antibody or peptide can be conjugated to a particle. For example, a bead can be conjugated to an antibody or peptide and used to detect and quantify the proteins expressed on the surface of a population of vesicles in order to discover and specifically select for novel antibodies that can target vesicles from specific tissue or tumor types. Any molecule of organic origin can be successfully conjugated to a polystyrene bead through use of a commercially available kit according to manufacturer's instructions. Each bead set can be colored a certain detectable wavelength and each can be linked to a known antibody or peptide which can be used to specifically measure which beads are linked to exosomal proteins matching the epitope of previously conjugated antibodies or peptides. The beads can be dyed with discrete fluorescence intensities such that each bead with a different intensity has a different binding agent as described above.

[00308] For example, a purified vesicle preparation can be diluted in assay buffer to an appropriate concentration according to empirically determined dynamic range of assay. A sufficient volume of coupled beads can be prepared and approximately 1 µι of the antibody-coupled beads can be aliquoted.
into a well and adjusted to a final volume of approximately 50 µl. Once the antibody-conjugated beads have been added to a vacuum compatible plate, the beads can be washed to ensure proper binding conditions. An appropriate volume of vesicle preparation can then be added to each well being tested and the mixture incubated, such as for 15-18 hours. A sufficient volume of detection antibodies using detection antibody diluent solution can be prepared and incubated with the mixture for 1 hour or more. The beads can then be washed before the addition of detection antibody (biotin expressing) mixture composed of streptavidin phycoerythrin. The beads can then be washed and vacuum aspirated several times before analysis on a suspension array system using software provided with an instrument. The identity of antigens that can be used to selectively extract the vesicles can then be elucidated from the analysis.

Assays using imaging systems can be used to detect and quantify proteins expressed on the surface of a vesicle in order to discover and specifically select for and enrich vesicles from specific tissue, cell or tumor types. Antibodies, peptides or cells conjugated to multiple well multiplex carbon coated plates can be used. Simultaneous measurement of many analytes in a well can be achieved through the use of capture antibodies arrayed on the patterned carbon working surface. Analytes can then be detected with antibodies labeled with reagents in electrode wells with an enhanced electro-chemiluminescent plate. Any molecule of organic origin can be successfully conjugated to the carbon coated plate. Proteins expressed on the surface of vesicles can be identified from this assay and can be used as targets to specifically select for and enrich vesicles from specific tissue or tumor types.

The binding agent can also be an aptamer to a specific target. The term "specific" as used herein in regards to a binding agent can mean that an agent has a greater affinity for its target than other targets, typically with a much greater affinity, but does not require that the binding agent is absolutely specific for its target.

**Microfluidics**

The methods for isolating or identifying vesicles can be used in combination with microfluidic devices. The methods of isolating or detecting a vesicle, such as described herein, can be performed using a microfluidic device. 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 a vesicle. Such systems miniaturize and compartmentalize processes that allow for binding of vesicles, detection of biosignatures, and other processes.

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 a vesicle. The selection can be based on a property of the vesicle, such as the size, shape, deformability, or biosignature of the vesicle.

In one embodiment, 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 the 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.

[00315] In some embodiments, the microfluidic device can comprise one or more channels that permit further enrichment or selection of a vesicle. 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 or vesicle population to be further enriched, such as through one or more binding agents present in the second channel.

[00316] 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 vesicle 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. In one embodiment, each population has a different biosignature. The hybridization reaction between the microsphere and vesicle 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.

[00317] Various microfluidic devices and methods are described above.

[00318] Combined Isolation Methodology

[00319] One of skill will appreciate that various methods of sample treatment and isolating and concentrating circulating biomarkers such as vesicles can be combined as desired. For example, a biological sample can be treated to prevent aggregation, remove undesired particulate and/or deplete highly abundant proteins. The steps used can be chosen to optimize downstream analysis steps. Next, biomarkers such as vesicles can be isolated, e.g., by chromatography, centrifugation, density gradient, filtration, precipitation, or affinity techniques. Any number of the later steps can be combined, e.g., a sample could be subjected to one or more of chromatography, centrifugation, density gradient, filtration and precipitation in order to isolate or concentrate all or most microvesicles. In a subsequent step, affinity techniques, e.g., using binding agents to one or more target of interest, can be used to isolate or identify microvesicles carrying desired biomarker profiles. Microfluidic systems can be employed to perform some or all of these steps.

[00320] An exemplary yet non-limiting isolation scheme for isolating and analysis of microvesicles includes the following: Plasma or serum collection -> highly abundant protein removal -> ultrafiltration -> nanomembrane concentration -> flow cytometry or particle-based assay.
Using the methods disclosed herein or known in the art, circulating biomarkers such as vesicles can be isolated or concentrated by at least about 2-fold, 3-fold, 1-fold, 2-fold, 3-fold, 4-fold, 5-fold, 6-fold, 7-fold, 8-fold, 9-fold, 10-fold, 12-fold, 15-fold, 20-fold, 25-fold, 30-fold, 35-fold, 40-fold, 45-fold, 50-fold, 55-fold, 60-fold, 65-fold, 70-fold, 75-fold, 80-fold, 90-fold, 95-fold, 100-fold, 110-fold, 120-fold, 125-fold, 130-fold, 140-fold, 150-fold, 160-fold, 170-fold, 175-fold, 180-fold, 190-fold, 200-fold, 205-fold, 250-fold, 275-fold, 300-fold, 325-fold, 350-fold, 375-fold, 400-fold, 425-fold, 450-fold, 475-fold, 500-fold, 525-fold, 550-fold, 575-fold, 600-fold, 625-fold, 650-fold, 675-fold, 700-fold, 725-fold, 750-fold, 775-fold, 800-fold, 825-fold, 850-fold, 875-fold, 900-fold, 925-fold, 950-fold, 975-fold, 1000-fold, 1500-fold, 2000-fold, 2500-fold, 3000-fold, 4000-fold, 5000-fold, 6000-fold, 7000-fold, 8000-fold, 9000-fold, or at least 10,000-fold. In some embodiments, the vesicles are isolated or concentrated concentrated by at least 1 order of magnitude, 2 orders of magnitude, 3 orders of magnitude, 4 orders of magnitude, 5 orders of magnitude, 6 orders of magnitude, 7 orders of magnitude, 8 orders of magnitude, 9 orders of magnitude, or 10 orders of magnitude or more.

Once concentrated or isolated, the circulating biomarkers can be assessed, e.g., in order to characterize a phenotype as described herein. In some embodiments, the concentration or isolation steps themselves shed light on the phenotype of interest. For example, affinity methods can detect the presence or level of specific biomarkers of interest.

The various isolation and detection systems described herein can be used to isolate or detect circulating biomarkers such as vesicles that are informative for 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. Combinations of the isolation techniques are within the scope of the invention. In a non-limiting example, a sample can be run through a chromatography column to isolate vesicles based on a property such as size of electrophoretic motility, and the vesicles can then be passed through a microfluidic device. Binding agents can be used before, during or after these steps.

The methods and compositions of the invention can be used with microvesicles isolated or detected using such methods as described herein. In various non-limiting examples: an aptamer provided by the methods of the invention can be used as a capture and/or detector agent for a biomarker such as a protein or microvesicle; a sample such as a bodily fluid can be contacted with an oligonucleotide probe library of the invention before microvesicles in the sample are isolated using one or more technique described herein (e.g., chromatography, centrifugation, flow cytometry, filtration, affinity isolation, polymer precipitation, etc); microvesicles in a sample are isolated using one or more technique described herein (e.g., chromatography, centrifugation, flow cytometry, filtration, affinity isolation, polymer precipitation, etc) before contacting the microvesicles with an aptamer or oligonucleotide probe library of the invention. Contaminants such as highly abundant proteins can be removed in whole or in part at any appropriate step in such processes. These and various other useful iterations of such techniques for assessment of microvesicles and other biomarkers are contemplated by the invention.
Biomarkers

As described herein, the methods and compositions of the invention can be used in assays to detect the presence or level of one or more biomarker of interest. The biomarker can be any useful biomarker disclosed herein or known to those of skill in the art. In an embodiment, the biomarker comprises a protein or polypeptide. As used herein, "protein," "polypeptide" and "peptide" are used interchangeably unless stated otherwise. The biomarker can be a nucleic acid, including DNA, RNA, and various subspecies of any thereof as disclosed herein or known in the art. The biomarker can comprise a lipid. The biomarker can comprise a carbohydrate. The biomarker can also be a complex, e.g., a complex comprising protein, nucleic acids, lipids and/or carbohydrates. In some embodiments, the biomarker comprises a microvesicle. In an embodiment, the invention provides a method wherein a pool of aptamers is used to assess the presence and/or level of a population of microvesicles of interest without knowing the precise microvesicle antigen targeted by each member of the pool. See, e.g., FIGs. 19B-C. In other cases, biomarkers associated with microvesicles are assessed according to the methods of the invention. See, e.g., FIGs. 2A-F; FIG. 19A.

A biosignature comprising more than one biomarker can comprise one type of biomarker or multiple types of biomarkers. As a non-limiting example, a biosignature can comprise multiple proteins, multiple nucleic acids, multiple lipids, multiple carbohydrates, multiple biomarker complexes, multiple microvesicles, or a combination of any thereof. For example, the biosignature may comprise one or more microvesicle, one or more protein, and one or more microRNA, wherein the one or more protein and/or one or more microRNA is optionally in association with the microvesicle as a surface antigen and/or payload, as appropriate.

In some embodiments, vesicles are detected using vesicle surface antigens. A commonly expressed vesicle surface antigen can be referred to as a "housekeeping protein," or general vesicle biomarker. The biomarker can be CD63, CD9, CD81, CD82, CD37, CD53, Rab-5b, Annexin V or MFG-E8. Tetraspanins, a family of membrane proteins with four transmembrane domains, can be used as general vesicle biomarkers. The tetraspanins include CD151, CD53, CD37, CD82, CD81, CD9 and CD63. There have been over 30 tetraspanins identified in mammals, including the TSPAN1 (TSP-1), TSPAN2 (TSP-2), TSPAN3 (TSP-3), TSPAN4 (TSP-4, NAG-2), TSPAN5 (TSP-5), TSPAN6 (TSP-6), TSPAN7 (CD231, TALLA-1, A15), TSPAN8 (CO-029), TSPAN9 (NET-5), TSPAN10 (Oculospanin), TSPAN11 (CD151-like), TSPAN12 (NET-2), TSPAN13 (NET-6), TSPAN14, TSPAN15 (NET-7), TSPAN16 (TM4-B), TSPAN17, TSPAN18, TSPAN19, TSPAN20 (UPlb, UPK1B), TSPAN21 (Upla, UPK1A), TSPAN22 (RDS, PRPH2), TSPAN23 (ROM1), TSPAN24 (CD151), TSPAN25 (CD53), TSPAN26 (CD37), TSPAN27 (CD82), TSPAN28 (CD81), TSPAN29 (CD9), TSPAN30 (CD63), TSPAN31 (SAS), TSPAN32 (TSSC6), TSPAN33, and TSPAN34. Other commonly observed vesicle markers include those listed in Table 3. One or more of these proteins can be useful biomarkers for the characterizing a phenotype using the subject methods and compositions.

**Table 3: Proteins Observed in Vesicles from Multiple Cell Types**

<table>
<thead>
<tr>
<th>Class</th>
<th>Protein</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antigen Presentation</td>
<td>MHC class I, MHC class II, Integrins, Alpha 4 beta 1, Alpha M beta 2,</td>
</tr>
<tr>
<td><strong>免疫球蛋白家族</strong></td>
<td>Beta 2</td>
</tr>
<tr>
<td>-------------------</td>
<td>-------</td>
</tr>
<tr>
<td><strong>细胞表面肽酶</strong></td>
<td>ICAM1/CD54, P-selection</td>
</tr>
<tr>
<td><strong>热休克蛋白</strong></td>
<td>CD151, CD53, CD37, CD82, CD81, CD9 and CD63</td>
</tr>
<tr>
<td><strong>细胞骨架蛋白</strong></td>
<td>Hsp70, Hsp84/90</td>
</tr>
<tr>
<td><strong>膜运输和融合</strong></td>
<td>Actin, Actin-binding proteins, Tubulin</td>
</tr>
<tr>
<td><strong>信号转导</strong></td>
<td>CD63, GAPDH, CD9, CD81, ANXA2, ENOL, SDCBP, MN, MFGE8, EKR, G1, ANXA1, LAMP2, DPP4, TSG101, HSPA1A, GDI2, CLTC, LAMP1, Cd86, ANPEP, TFRC, SLC3A2, RDX, RP1B, RAB5C, RAB5B, MYH9, ICAM1, FN1, RAB11B, PIGR, LGALS3, ITGB1, EHD1, CLIC1, ATPIA1, ARF1, RAP1A, P4HB, MUC1, KRT10, HLA-A, FLOT1, CD59, Clor58, BASP1, TACSTD1, STOM</td>
</tr>
</tbody>
</table>

**其他跨膜蛋白**

Cadherins: CDH1, CDH2, CDH12, CDH3, Deomoglein, DSG1, DSG2, DSG3, DSG4, Desmocollin, DSC1, DSC2, DSC3, Protocadherins, PCDH1, PCDH10, PCDH11, PCDH11y, PCDH12, FAT, FAT2, FAT4, PCDH15, PCDH17, PCDH18, PCDH19, PCDH20, PCDH7, PCDH8, PCDH9, PCDHA1, PCDHA10, PCDHA11, PCDHA12, PCDHA13, PCDHA2, PCDHA3, PCDHA4, PCDHA5, PCDHA6, PCDHA7, PCDHA8, PCDHA9, PCDHAC1, PCDHAC2, PCDHB1, PCDHB10, PCDHB11, PCDHB12, PCDHB13, PCDHB14, PCDHB15, PCDHB16, PCDHB17, PCDHB18, PCDHB2, PCDHB3, PCDHB4, PCDHB5, PCDHB6, PCDHB7, PCDHB8, PCDHB9, PCDHG1A, PCDHG1A10, PCDHG1A11, PCDHG1A12, PCDHG1A22, PCDHG3A3, PCDHG4A4, PCDHG5A5, PCDHG5A6, PCDHG7A, PCDHG8A, PCDHG9A, PCDHGB1, PCDHGB2, PCDHGB3, PCDHGB4, PCDHGB5, PCDHGB6, PCDHGB7, PCDHGC3, PCDHGC4, PCDHGC5, CDH9 (cadherin 9, type 2 (T1-cadherin)), CDH10 (cadherin 10, type 2 (T2-cadherin)), CDH5 (VE-cadherin (vascular endothelial)), CDH6 (K-cadherin (kidney)), CDH7 (cadherin 7, type 2), CDH8 (cadherin 8, type 2), CDH11 (OB-cadherin (osteoblast)), CDH13 (T-cadherin - H-cadherin (heart)), CDH15 (M-cadherin (myotubule)), CDH16 (KSP-cadherin), CDH17 (LI cadherin (liver-intestine)), CDH18 (cadherin 18, type 2), CDH19 (cadherin 19, type 2), CDH20 (cadherin 20, type 2), CDH23 (cadherin 23, (neurosensor epithelium)), CDH10, CDH11, CDH13, CDH15, CDH16, CDH17, CDH18, CDH19, CDH20, CDH22, CDH23, CDH24, CD26, CD28, CD4, CD5, CD6, CD7, CD8, CD9, CELS1, CELS2, CELS3, CLSTN1, CLSTN2, CLSTN3, DCHS1, DCHS2, LOC389118, PCLKC, RESDA1, RET |

[00328] Any of the types of biomarkers or specific biomarkers described herein can be used and/or assessed via the subject methods and compositions, e.g., to identify a useful biosignature. Exemplary biomarkers include without limitation those in Table 4. The markers can be detected as protein, RNA or DNA as appropriate, which can be circulating freely or in a complex with other biological molecules. As appropriate, the markers in Table 4 can also be used for capture and/or detection of vesicles for characterizing phenotypes as disclosed herein. In some cases, multiple capture and/or detectors are used to enhance the characterization. The markers can be detected as protein or as mRNA, which can be circulating freely or in a complex with other biological molecules. See, e.g., FIGs. 2D-E. The markers can be detected as vesicle surface antigens and/or vesicle payload. The "Illustrative Class" indicates indications for which the markers are known markers. Those of skill will appreciate that the markers can
also be used in alternate settings in certain instances. For example, a marker which can be used to
classify one type disease may also be used to characterize another disease as appropriate. Consider a
non-limiting example of a tumor marker which can be used as a biomarker for tumors from various
lineages. The biomarker references in Tables 3 and 4 are those commonly used in the art. Gene aliases
and descriptions can be found using a variety of online databases, including GeneCards®
(www.genecards.org), HUGO Gene Nomenclature (www.genenames.org), Entrez Gene
GeneLoc (genecards.weizmann.ac.il/geneloc/), and Ensembl (www.ensembl.org). Generally, gene
symbols and names below correspond to those approved by HUGO, and protein names are those
recommended by UniProtKB/Swiss-Prot. Common alternatives are provided as well. In some cases,
biomarkers are referred to by Ensembl reference numbers, which are of the form "ENSG" followed by a
number, e.g., ENSG00000005893 which corresponds to LAMP2. In Table 4, solely for sake of brevity,
"E." is sometimes used to represent "ENSG00000005893." Where a protein name indicates a precursor, the mature protein is also implied.
Throughout the application, gene and protein symbols may be used interchangeably and the meaning can
be derived from context as necessary.

Table 4: Illustrative Biomarkers

<table>
<thead>
<tr>
<th>Illustrative Class</th>
<th>Biomarkers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drug associated targets and prognostic markers</td>
<td>ABCC1, ABCG2, ACE2, ADA, ADHIC, ADH4, AGT, AR, AREG, ASNS, BCL2, BCRP, BDCA1, beta III tubulin, BIRC5, B-RAF, BRCA1, BRCA2, CA2, caveolin, CD20, CD25, CD33, CD52, CDA, CDKN2A, CDKN1A, CDKN1B, CDK2, CDW52, CES2, CK 14, CK 17, CK S6, CK-Y, c-MET, c-Myc, COX-2, CyclinD1, DCK, DHRF, DNMT1, DNMT3A, DNMT3B, E-Cadherin, ECGF1, EGFR, EML4-ALK fusion, EPHA2, Epiregulin, ER, ERBR, ERCC1, ERCC3, EREG, ERSI, FLT1, folate receptor, FOLR1, FOLR2, FSHb, FSHPRHI, FSHR, FYN, GART, GNAI1, GNAQ, GNRH1, GNRHR1, GSTP1, HCK, HDAC1, hENT-1, Her2/Neu, HGF, HIF1A, HIG1, HSP90, HSP90AA1, HSPCA, IGF1-IR, IGFIRBP, IGFIRBP3, IGFIRBP4, IGFIRBP5, IL13RA1, IL2RA, KDR, Ki67, KIT, K-RAS, LCK, LTBP, Lymphoxygenin Beta Receptor, LYN, MET, MGMT, MLH1, MRK, MRPL1, MS4A1, MSH2, MSH5, Mye, NFKB1, NFKB2, NFKBIA, NRAS, ODC1, OGRF, pl6, P21, p27, p53, p95, PARP-1, PDGFC, PDGFR, PDGFR, PDPFR, PGP, PGR, PI3K, POLA, POLA1, PPAR, PPARC1, PR, PTEN, PTGS2, PTN512, RAF1, RARA, RSK1, RSM1, RRM2, RRM2B, RXRB, RXRG, SIK2, SPARC, SRC, SSTR1, SSTR2, SSTR3, SSTR4, SSTR5, Survivin, TKI, TLE3, TNF, TOP1, TOP2A, TOP2B, TS, TUB3B, TXN, TXNRDI, TYMS, VDR, VEGF, VEGFA, VEGFC, VHL, YES1, ZAP70</td>
</tr>
<tr>
<td>Drug associated targets and prognostic markers</td>
<td>ABL1, STK11, FGFR2, ERBB4, SMARCB1, CDKN2A, CTNNB1, FGFR1, FLT3, NOTCH1, NPM1, SRC, SMAD4, FBXW7, PTEN, TP53, AKT1, ALK, APC, CDH1, C-Met, HRAS, IDH1, JAK2, MPL, PDGFRa, PDPFRA, PDPFRA, PDPFR, PGP, PGR, PI3K, POLA, POLA1, PPAR, PPARC1, PR, PTEN, PTGS2, PTN512, RAF1, RARA, RSK1, RSM1, RRM2, RRM2B, RXRB, RXRG, SIK2, SPARC, SRC, SSTR1, SSTR2, SSTR3, SSTR4, SSTR5, Survivin, TKI, TLE3, TNF, TOP1, TOP2A, TOP2B, TS, TUB3B, TXN, TXNRDI, TYMS, VDR, VEGF, VEGFA, VEGFC, VHL, YES1, ZAP70</td>
</tr>
<tr>
<td>Drug associated targets and prognostic markers</td>
<td>ALK, AR, BRAF, CKIT, cMET, EGFR, ER, ERCC1, GNA11, HER2, IDH1, KRAS, MGMT, MGMT promoter methylation, NRAS, PDGFRa, Pgp, PIK3CA, PR, PTEN, ROS1, RRM1, SPARC, TLE3, TOP2A, TOPO1, TS, TUB3B, VHL</td>
</tr>
<tr>
<td>Drug associated targets</td>
<td>ABL1, AKT1, ALK, ARC, AR, ATM, BRAF, BRF, BRCA1, BRCA2, CDH1, cKIT, cMET, CSF1R, CTNNB1, EGFR, EGFR (H-score), EGFRvIII, ER, ERBB2 (HER2), ERBB4, ERCC1, FBXW7, FGFR1, FGFR2, FLT3, GNA11, GNAQ, GNAS, HER2, HNF1A, HRAS, IDH1, IDH2, JAK2, JAK3, KDR (VEGFR2), KRAS, MGMT, MGMT Promoter Methylation, microsatellite instability (MSI), MLH1, MPL, MSH2, MSH6, NOTCH1, NPM1, NRAS, PD-1, PDGFRa, PD-L1, Pgp, PIK3CA, PMS2, PR, PTEN, PTN511, RB1, RET, ROS1, RRM1, SMAD4, SMARCB1, SOM, SPARC, STK11, TLE3, TOP2A, TOPO1, TP53, TS, TUB3B, VHL</td>
</tr>
<tr>
<td>Disease</td>
<td>Targets</td>
</tr>
<tr>
<td>-------------------------</td>
<td>-------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Breast cancer</td>
<td>lpl9q co-deletion, ABL1, AKTI, ALK, APC, AR, ARAF, ATM, BAP1, BRAF, BRCA1, BRCA2, CDH1, CHEK1, CHEK2, cKIT, cMET, CSF1R, CTNNB1, DDR2, EGFR, EGFRvIII, ER, ERBB2 (HER2), ERBB3, ERBB4, ERCC1, FBXW7, FGFR1, FGFR2, FLT3, GNAI1, GNAQ, GNAT, H3K36me3, HER2, HNFA, HRS, IDH1, IDH2, JAK2, JAK3, KDR (VEGFR2), KRAS, MDMT, MGMT, MGMT Methylation, Microsatellite instability, MLH1, MPL, MSH2, MSH6, NFI, NOTCH1, NPM1, NRAS, NY-ESO-1, PD-1, PDGFRα, PD-L1, Pdgf, PIK3CA, PMS2, PR, PTEN, PTNP11, RAF1, RIB1, RET, RO5, RO5I, RRM1, SMAD4, SMARCB1, SMO, SPARC, STK11, TLE3, TOP2A, TOPO1, TP53, TRKA, TS, TUBB3, VHL, WT1</td>
</tr>
<tr>
<td>Breast cancer</td>
<td>ABL1, AKTI, ALK, APC, AR, ATM, BRAF, BAP1, BRAF, BRCA1, BRCA2, CDH1, cKIT, cMET, CSF1R, CTNNB1, EGFR, EGFR (H-score), EGFRVIII, ER, ERBB2 (HER2), ERBB4, ERCC1, FBXW7, FGFR1, FGFR2, FLT3, GNAI1, GNAQ, GNAT, HER2, HNFA, HRAS, IDH1, IDH2, JAK2, JAK3, KDR (VEGFR2), KRAS, MDMT, MGMT Methylation, microsatellite instability (MSI), MLH1, MSH1, MSH2, MSH6, NOTCH1, NPM1, NRAS, PD-1, PDGFRα, PD-L1, Pdgf, PIK3CA, PMS2, PR, PTEN, PTNP11, RAF1, RIB1, RET, RO5, RO5I, SMAD4, SMARCB1, SMO, SPARC, STK11, TLE3, TOP2A, TOPO1, TP53, TS, TUBB3, VHL</td>
</tr>
<tr>
<td>Ovarian cancer</td>
<td>lpl9q, ALK, ALK (2p23), Androgen Receptor, BRCA, cMET, EGFR, EGFR, EGFRVIII, ER, ERCC1, Her2, Her2/Neu, MGMT, MGMT Promoter Methylation, microsatellite instability (MSI), MLH1, MSH2, MSH6, PD-1, PD-L1, PMS2, PR, PTEN, RO5, RO5I, TLE3, TOP2A, TOP2A, TOPO1, TS, TUBB3</td>
</tr>
<tr>
<td>Colon cancer</td>
<td>5-aminosalicylic acid (5-ASA) efficacy</td>
</tr>
<tr>
<td>Melanoma cancer</td>
<td>µ-protocadherin, KLF4, CEBPa</td>
</tr>
<tr>
<td>Cancer treatment</td>
<td>AR, AREG (Amphiregulin), BRAF, BRCA1, cKIT, cMET, EGFR, EGFR w/T790M, EML4-ALK, ER, ERBB3, ERBB4, ERCC1, EREG, GNAL1, GNAQ, hENT-1, Her2, Her2/Neu 20 insert, IGF1R, Ki67, KRAS, MGMT, MGMT methylation, microsatellite instability (MSI), MLH1, MSH2, MSH6, PD-1, PD-L1, PMS2, PR, PTEN, RO5, RO5I, TLE3, TOP2A, TS, TUBB3, VEGFR2</td>
</tr>
<tr>
<td>Cancer treatment</td>
<td>AR, AREG, BRAF, BRCA1, cKIT, cMET, EGFR, EGFR w/T790M, EML4-ALK, ER, ERBB3, ERBB4, ERCC1, EREG, GNAL1, GNAQ, Her2, Her2/Neu 20 insert, IGF1R, Ki67, KRAS, MGMT-Me, MSH2, MSH6, PD-1, Pdg, PIK3CA, PR, PTEN, RO5, RO5I, TLE3, TOP2A, TS, TUBB3, VEGFR2</td>
</tr>
<tr>
<td>Colon cancer</td>
<td>AREG, BRAF, EGFR, EML4-ALK, ERCC1, EREG, KRAS, MSH2, MSH6, PIK3CA, PR, PTEN, TS, VEGFR2</td>
</tr>
<tr>
<td>Colon cancer</td>
<td>AREG, BRAF, EGFR, EML4-ALK, ERCC1, EREG, KRAS, MSH2, MSH6, PIK3CA, PR, PTEN, TS, VEGFR2</td>
</tr>
<tr>
<td>Ovarian cancer</td>
<td>BRAF, cKIT, ERBB3, ERBB4, ERCC1, GNAI1, GNAQ, MGMT, MGMT Methylation, MSH2, MSH6, PIK3CA, TUBB3, VEGFR2</td>
</tr>
<tr>
<td>Ovarian cancer</td>
<td>BRAF, cKIT, ERBB3, ERBB4, ERCC1, GNAI1, GNAQ, MGMT-Me, MSH2, MSH6, PIK3CA, TUBB3, VEGFR2</td>
</tr>
<tr>
<td>Breast cancer</td>
<td>BRCA1, cMET, EML4-ALK, ER, ERBB3, ERCC1, hENT-1, HER2, IGF1R, Pdgf(MDR1), PIK3CA, PR, PTEN, RO5I, TLE3, TOP2A, TS</td>
</tr>
<tr>
<td>Breast cancer</td>
<td>BRCA1, cMET, EML4-ALK (translocation), ER, ERBB3, ERCC1, HER2, PIK3CA, PR, PTEN, RO5I, TLE3, TS</td>
</tr>
<tr>
<td>Breast cancer</td>
<td>BRAF, BRCA1, EGFR, EGFR T790M, EML4-ALK, ER, ERBB3, ERCC1, HER2, Ki67, Pdgf (MDR1), PIK3CA, PR, PTEN, RO5, RO5I Translocation, TLE3, TOP2A, TS</td>
</tr>
<tr>
<td>Breast cancer</td>
<td>BRAF, BRCA1, EGFR w/T790M, EML4-ALK, ER, ERBB3, ERCC1, HER2, Ki67, KRAS, PIK3CA, PR, PTEN, RO5I Translocation, TLE3, TOP2A, TS</td>
</tr>
<tr>
<td>NSCLC cancer</td>
<td>BRAF, BRCA1, cMET, EGFR, EGFR w/T790M, EML4-ALK, ERCC1, HER2/Neu 20 insert, KRAS, MSH2, PIK3CA, PR, PTEN, RO5I (trans), RRMI, TLE3, TS, VEGFR2</td>
</tr>
</tbody>
</table>
NSCLC cancer
treatment
associated markers
BRAF, cMET, EGFR, EGFR w/T790M, EML4-ALK, ERCC1, Her2 Exon 20 insert, KRAS, MSH2, PIK3CA, Pten, ROS1 translocation, RRMI, TLE3, TS

Mutated in cancers
AKT1, ALK, APC, ATM, BRAF, CDH1, CDKN2A, c-KIT, C-Met, CSF1R, CTNNB1, EGFR, ERBB2, ERBB4, FBXW7, FGFR1, FGFR2, FGFR3, FLT3, GNA11, GNAQ, GNAS, HNF1A, HRAS, ID1H, JAK2, JAK3, KDR, KRAS, MLH1, MPL, NOTCH1, NPM1, NRAS, PDGFRα, PIK3CA, Pten, PTEN11, RB1, RET, SMAD4, SMARC1B1, SMO, SRC, STK11, TP53, VHL

Mutated in cancers
ALK, BRAF, BRCA1, BRCA2, EGFR, ERBB2, GNA11, GNAQ, ID1H, ID2H, KIT, KRAS, MET, NRAS, PDGFRα, PIK3CA, Pten, RET, SRC, TP53

Mutated in cancers
AKT1, HRAS, GNAS, MEK1, MEK2, ERK1, ERK2, ERBB3, CDKN2A, PDGFRβ, IFGIR, FGFR1, FGFR2, FGFR3, ERBB4, SMO, DDR2, GRB1, PTCH1, SHH, PDI, UGT1A1, BIM, ESRI, MLL, AR, CDK4, SMAD4

Mutated in cancers
ABL, APC, ATM, CDH1, CSF1R1, CTNNB1, FBXW7, FLT3, HNF1A, JAK2, JAK3, KDR, MLH1, MPL, NOTCH1, PCNA, PTEN11, RB1, SMARC1B1, STK11, VHL

Mutated in cancers
ABL1, AKT1, AKT2, AKT3, ALK, APC, AR, ARAF, ARFRP1, ARID1A, ARID2, ASXL1, ATM, ATR, ATRX, AURKA, AURKB, AXL, BAP1, BARD1, BCL2, BCL2L2, BCL6, BCROR, BCROR1, BLM, BRAF, BRCA1, BRCA2, BRIP1, BTK, CARD11, CBFB, CBL, CCND1, CCND2, CCND3, CCNE1, CD79A, CD79B, CD73, CDH1, CDK2, CDK4, CDK6, CKDN1B, CDKN2A, CDKN2B, CDKN2C, CEPPA, CHEK1, CHEK2, CIC, CREBBP, CRKL, CRLF2, CSF1R, CTCF, CTNNB1, CTNNB1, DAXX, DDR2, DNM3A, DOT1L, EGFR, EMT3, (C1QRo5), GNA11, GNA13, GNAQ, GNAS, GPR124, GRN2A, GSK3B, HGF, HRAS, ID1H, ID2H, IGF1R, IKBKE, IKZF1, IL7R, INHBA, JAK4, IRS2, JAK1, JAK2, JAK3, JUN, KAT6A (MYST3), KDM5A, KDM5C, KDM6A, KDR, KEP1A1, KIT, KLH, KRAS, LRP1B, MAPK1, MAPK2, MAPK4, MAPK3, MCL1, MDM2, MDM4, MED12, MEF2B, MEN1, MET, MITF, MLH1, MLL1, MLL2, MPR1, MREIIA, MSH2, MSH6, MTO, MUTYH, MYC, MYCL1, MYCN, MYD88, NFI, NFE2L2, NFKBIA, NKG2A, NTC1, NOTCH1, NOTCH2, NPM1, NRAS, NTRKI, NTRK2, NTRK3, NUP93, PAK3, PALB2, PAX5, PBRM1, PDGFRα, PDGFRβ, PDK1, PIK3CA, PIK3CG, PIK3R1, PIK3R2, PPP2R1A, PRDM1, PRKAR1A, PRKDC, PTCH1, PTEN, PTEN11, RAD50, RAD51, RAF1, RARA, RBL1, RET, RICTOR, RNF43, RPTOR, RUNX1, SETD2, SF3B1, SMAD2, SMAD4, SMARC4, SMARC1B1, SMd, SOCS1, SOX10, SOX2, SPEN, SPOC1, SRC, STAG2, STAG4, STK11, SUFU, TET2, TGFB2, TNAIP3, TNFRSF4, TOP1, TP53, TSC1, TSC2, TSHR, VHL, WISP3, WT1, XPO1, ZNF217, ZNF703

Gene rearrangement in cancer
ALK, BCR, BCL2, BRAF, EGFR, ETV1, ETV4, ETV5, ETV6, EWSR1, MLL, MYC, NTRKI, PDGFRα, RAI, RARA, RET, ROS1, TMPRSS2

Cancer Related
ABL1, ACE2, ADA, ADH1C, ADH4, AGT, AKT1, AKT2, AKT3, ALK, APC, AR, ARAF, AREG, ARFRP1, ARID1A, ARID2, ASNS, ASXL1, ATM, ATR, ATRX, AURKA, AURKB, AXL, BAP1, BARD1, BCL2, BCL2L2, BCL6, BCROR, BCROR1, BCR, BRC (survivin), BLM, BRAF, BRCA1, BRCA2, BRIP1, BTK, CA2, CARD11, CAV, CBFB, CBL, CCND1, CCND2, CCND3, CCNE1, CD33, CD52 (CDW52), CD79A, CD79B, CD73, CDH1, CDK12, CDK2, CDK4, CDK6, CKDN1B, CDKN2A, CDKN2B, CDKN2C, CEPPA, CHEK1, CHEK2, CIC, CREBBP, CRKL, CRLF2, CSF1R, CTCF, CTNNB1, DAXX, DCK, DDR2, DHFR, DNM3, DNM3A, DNM3B, DOT1L, EGFR, EMSY (C1QRo5), EP300, EPHA2, EPHA3, EPHA5, EPB1, ERBB2, ERBB3, ERBB4, ERBB4 (type?), ERCC3, EREG, Ereg, ESRI, ETV1, ETV4, ETV5, ETV6, EWSR1, EZH2, FAM123B (WTX), FAM46C, FANCA, FANCC, FANCD2, FANCE, FANCF, FANCG, FANCL, FBXW7, FGFR1, FGFR10, FGFR4, FGFR19, FGFR23, FGFR3, FGFR4, FGFR1, FGFR2, FGFR3, FGFR4, FLT1, FLT3, FLT4, FOLR1, FOLR2, FOXL2, FSHB, FSHPRH1, FSHR, GART, GATA1, GATA2, GATA3, GI3D (C1QRo5), GNA11, GNA13, GNAQ, GNAS, GNRH1, GNRH1, GPR124, GRN2A, GSK3B, GSTP1, HAC1, HGF, HIG1, HNFA1, HRAS, HSPCA (HSP90), ID1H, ID2H, IGF1R, IKBKE, IKZF1, IL13RA1, IL2, IL2RA (CD25), IL7R, INHBA, IGF4, IRS2, JAK1, JAK2, JAK3, JUN, KAT6A (MYST3), KDM5A, KDM5C, KDM6A, KDR (VEGFR2), KEAP1, KIT, KLH, KRAS, LCK, LRP1B, LT, LTBR, MAP2K1, MAP2K2, MAP2K4, MAP3K1, MAPK, MCL1, MDM2, MDM4, MED12, MEF2B, MEN1, MET, MM一方面, MMG, MITF, MLH1, MLL2, MPR1, MREIIA, MSH2, MSH6, MTO, MUTYH, MYC, MYCL1, MYCN, MYD88, NFI, NFE2L2, NFKBIA, NFKBII, NGF, NKG2A, NTC1, NOTCH1, NOTCH2, NPM1, NRAS, NTRKI, NTRK2, NTRK3, NUP93, PAK3, PALB2, PAX5, PBRM1, PDGFRα, PDGFRβ, PDK1, PIK3CA, PIK3CG, PIK3R1, PIK3R2, PPP2R1A, PRDM1, PRKAR1A, PRKDC, PTCH1, PTEN, PTEN11, RAD50, RAD51, RAF1, RARA, RBL1, RET, RICTOR, RNF43, RPTOR, RUNX1, SETD2, SF3B1, SMAD2, SMAD4, SMARC4, SMARC1B1, SMd, SOCS1, SOX10, SOX2, SPOC1, SRC, STAG2, STAG4, STK11, SUFU, TET2, TGFB2, TNAIP3, TNFRSF4, TOP1, TP53, TSC1, TSC2, TSHR, VHL, WISP3, WT1, XPO1, ZNF217, ZNF703

-92-
PTEN, PTK2, PTPN11, PTPRC, QKI, RABEP1, RAC1, RAD21, RAD50, RAD51, RAD51B, RAF1, RALGDS, RANBP17, RANBP2, RANKL, RAPID5I, RARA, RBA, RB1, RBM10, RBM15, RECQL4, REL, RET, RHOH, RICTOR, RM12, RNF213, RNF43, ROS1, RPL10, RPL20, RPL5, RPN1, RPS6KB1, RPTOR, RUNX1, RUNXI1, SBD5, SDC4, SDHA, SDHAF2, SDHB, SDHC, SDHD, SEPT5, SEPT6, SEPT9, SET, SETBP1, SETD2, SF3B1, SFQ, SH2B3, SH3GL1, SLAMF7, SLC34A2, SLC4A3, SLI2, SLMD2, SMAD3, SMAD4, SMARC4, SMARCB1, SMARC1, SMOK, SNAIP1, SNX29, SOCS1, SOX10, SOX2, SOX9, SPP1, SPO, SPTA1, SRC, SRGAP3, SRFS2, SRFS3, SS18, SS18L1, SSX1, SSX2, SSX4, STAG2, STAT3, STAT4, STAT5B, STEAP1, STIL, STK14, SUFU, SUZ12, SYK, TAF15, TAL1, TAI2, TBL1XR1, TBX3, TCEAI, TCF12, TCF3, TCF7L2, TCLA, TERC, TERT, TET1, TET2, TFE3, TFE6, TFG, TFF, TFR, TGFBR2, THAP3, TIE2, TLX1, TLX3, TPMRSS2, TNAI3, TNFRSF14, TNFRSF17, TOP1, TOP2A, TP53, TPM3, TPR, TRAF7, TRIM26, TRIM27, TRIM33, TRIP11, TRAP, TSC1, TSC2, TSHR, TTL, U2AF1, UBA1, UBR5, USP6, VEGER, VEGFB, VEGFR, VHL, VTI1A, WAS, WEI1, WHSCI1, WHSCI2, WIFI, WISP3, WNT11, WNT2B, WNT3, WNT4, WNT5A, WNT6, WNT7B, WRN, WT1, WWT1, XPA, XPC, XPOI, YWHAE, ZAK, ZBTB16, ZBTB2, ZMYM2, ZNF217, ZNF331, ZNF521, ZNF570, ZR52R

| Gene fusions and mutations in cancer | AKT3, ALK, ARHGA26, AXL, BRAF, BRD3/4, EGFR, ERG, ERK1/2, ET1V1/4/5/6, EWSRI, FGFR1, FGFR2, FGFR3, FGR, INS, MAML2, MAST1/2, MET, MSMB, MUSK, MYB, NOTCH1/2, NRG1, NTRK1/2/3, NUMBL, NUTM1, PDGFRAB/P, PIK3CA, PKN1, PPARG, PKCAB, RAF1, RELA, RET, ROS1, RSP0/2/3, TERT, TFE3, TFE6, THADA, TPMRSS2 |

Genes fusions and

ABL1 fusion to (ETV6, NUP214, RCSD1, RANBP2, SNX2, orZMIPI); ABL2 fusion to
mutations in cancer: (PAG1 or RCSD1); CSF1 R fusion to (SSBP2); PDGFRB fusion to (EBF1, SSBP2, TNIP1 or ZEB2); CRLF2 fusion to (P2RY8); JAK2 fusion to (ATF7IP, BCR, ETV6, PAX5, PPIBPI, SSBP2, STRN3, TERF2, or TPR); EPO fusion to (IGH or IGK); IL2RB fusion to (MYH9); NTRK3 fusion to (ETV6); PTK2B fusion to (KDM6A or STAG2); TSLP fusion to (IQGAP2); TYK2 fusion to (MYB).

Cytohesins

cytohesin-1 (CYTH1), cytohesin-2 (CYTH2; ARNO), cytohesin-3 (CYTH3; Grpl; ARN03), cytohesin-4 (CYTH4)

Cancer/Angio

Erb 2, Erb 3, Erb 4, UNC93a, B7H3, MUC1, MUC2, MUC16, MUC17, 5T4, RAGE, VEGF A, VEGFR2, FLT1, DLL4, Epcam

Tissue (Breast)

BIG H3, GCDFP-15, PR(B), GPR 30, CYFRA 21, BRCA 1, BRCA 2, ESR 1, ESR2

Tissue (Prostate)

PSMA, PCSA, PSCA, PSA, TMPRSS2

Inflammation/Immune

MFG-E8, IFNAR, CD40, CD80, MICB, HLA-DRb, IL-17-Ra

Common vesicle markers

HSPA8, CD63, Actb, GAPDH, CD9, CD81, ANXA2, HSP90AA1, ENO1, YWHAZ, PDCD6IP, CFL1, SDCBP, PKN2, MNF, MFGES8, EZR, YWHAQ, PGK1, EEF1A1, PPIA, GLC1F, GK, ANXA6, ANXA1, ALDOA, ACTG1, TPI1, LAMP2, HSP90AB1, DPP4, YWHAQ, TSG101, PFN1, LDLB, HSPA1B, HSPA1A, GSN1, GSN2, CLTC, ANXA5, YWHAQ, TUBA1A, THBS1, PRDX1, LDHA, LAMP1, CLU, CD86

Common vesicle membrane markers

CD63, GAPDH, CD9, CD81, ANXA2, ENO1, SDCBP, MSN, MFGES8, EZR, GK, ANXA1, LAMP2, DPP4, TSG101, HSPA1A, CD2, CLTC, LAMP1, CD86, ANPEP, TFCR, SLC3A2, RDX, RAP1B, RABC, RAB5B, MYH9, ICAM1, FN1, RAB11B, PIGR, LGALS3, ITGB1, EHD1, CLIC1, ATPIAI, ARFI, RAPIA, P4HB, MUC1, KRTIO, HLA-A, FLOT1, CD59, Clorf58, BASP1, TACSTD1, STOM

Common vesicle markers

MHC class I, MHC class II, Integrins, Alpha 4 beta 1, Alpha M beta 2, Beta 2, ICAM1/CD54, P-selection, Dipeptidylpeptidase IV/CD26, Aminopeptidase n/CD13, CD151, CD153, CD37, CD82, CD81, CD9, HSP90ab1, Hsp90/90

Actin, Actin-binding proteins, Tubulin, Annexin I, Annexin II, Annexin IV, Annexin V, Annexin VI, RAB7/RAP1B/RADGDI1, G2alpha14/3-3, CBL/LCK, CD63, GAPDH, CD9, CD81, ANXA2, ENO1, SDCBP, MNF, MFGES8, EZR, GK, ANXA1, LAMP2, DPP4, TSG101, HSPA1A, CD2, CLTC, LAMP1, CD86, ANPEP, TFCR, SLC3A2, RDX, RAP1B, RABC, RAB5B, MYH9, ICAM1, FN1, RAB11B, PIGR, LGALS3, ITGB1, EHD1, CLIC1, ATPIAI, ARFI, RAPIA, P4HB, MUC1, KRTIO, HLA-A, FLOT1, CD59, Clorf58, BASP1, TACSTD1, STOM

Vesicle markers

A33, a33 n5, AFP, ALA, ALIX, ALP, AnnexinV, APC, ASCA, ASPH (246-260), ASPH (666-680), ASPH (A-10), ASPH (DOIP), ASPH (D03), ASPH (G-20), ASPH (H-300), AURKA, AURKB, B7H3, B7H4, BCA-225, BCNP, BDNF, BRCA, CA125 (MUC6), CA19-9, C-Bir, CDL1, CD1O, CD174 (Lewis y), CD24, CD44, CD46, CD59 (MEM-43), CD63, CD66e CEA, CD73, CD81, CD9, CD2a, CDAC1 la2, CEa, C-Erb2, C-erbB2, CRMP-2, CRP, CXCL12, CYFRA21-1, D4LL, DR3, ERpG, Epcam, EphA2, EphA7 (H-77), ER, ErbB4, EZH2, FASL, FRT, FRT c.123, GDF51, GPCR, GPR30, Gro-alpha, HAP, HBD 1, HBD2, HER 3 (ErbB3), HSP, HSP70, hVEGFR2, ic3B, IL 6 Unc, IL-1B, IL6 Unc, IL6R, IL8, IL-8 INSIG-2, KLK2, LCC1, LAMN, LDH, MHC-C, MAPK4, MART-1, M-CAP, M-CSF, MGF-E8, MCI, MIF, MIS RII, MMG, MMP26, MMP7, MMP9, MS4A1, MUC1, MUCI sequl, MUCI sequlA, MUC17, MUC2, Ncam, NGAL, NPRG/NPPF2, OPG, OPN, p53, p53, PA2G4, PBP, PCSA, PDGFRB, PGP9.5, PIMI, PR (B), PRL, PSA, PSMA, PSME3, PTEN, R5-CD9 Tube 1, Reg IV, RUNX2, SCRNI1, seprase, SERPINC3, SPARC, SPB, SPDE5, SRNT, STAT 3, STEAP1, TF (FL-295), TFF3, TMG2, TIMP-1, TIMP2, TIMP2, TME211, TMRPS2, TNF-alpha, Trail-R2, Trail-R4, TrkB, TROP2, Tsg 101, TWEAK, UNC93A, VEGF A, YPSMA-1

Vesicle markers

NSE, TRIM29, CD63, CD151, ASPH, LAMP2, TSPAN1, SNAIL1, CD45, CKS1, NSE, FSHR, OPN, FTHI, PGP9, ANNEXIN 1, SPD, CD81, EPCAM, PTHIR, CEA, CYTO 7, CCL2, SPA, KRAS, TWIST1, AURKB, MMP9, P27, MMP1, HLA, HIF, CEA/CAM, CENPH, BTUB, INTG b4, EGFR, NACC1, CYTO 18, NAP2, CYTO 19, ANNEXIN V, TGM2, ERBB2, BRCA1, B7H3, SFTP1, PNT, NCM, MS4A1, P53, INGA3, MUC2, SPA, OPN, CD63, CD9, MUC1, UNCR3, PAN ADH, HCG, TIMP, PSMA, GPCR, RACK1, PSCA, VEGF, BMP, CD61, CRP, PRO, GPR, B7H3, MUC1, M2PK, CD9, PCSA, PSMA

Vesicle markers

TFF3, MS4A1, EphA2, GAL3, EGFR, N-gal, PCSA, CD63, MUC1, TGM2, CD81, DR3, MACC-1, TrkB, CD24, TIMP-1, A33, CD66 CEA, PRL, MMP7, TME211, SCRNI1, TROP2, TWEAK, CDACCI, UNC93A, APC, C-Erb, CDIO, BDNF, FRT, GPR30, P53, SPR, OPN, MUC2, GRO-1, tsg 101, GDF15

Vesicle markers

CD9, Erb2, Erb4, CD81, Erb3, MUC16, CD63, DLL4, HLA-Drpe, B7H3, IFNAR, 5T4, PCSA, MICB, PSMA, MFG-E8, Mac1, PSA, Mac2, Unc93a, VEGFR2, Epcam, VEGF A,
Benign Prostate Hyperplasia (BPH)
TMPRSS2, RAGE, PSCA, CD40, Muc17, IL-17A, CD80

Metastatic Prostate Cancer
BCMA, CEACAM-1, HVEM, IL-1 R4, IL-10 Rb, Trappin-2, p53, hsa-miR-329, hsa-miR-30a, hsa-miR-335, hsa-miR-152, hsa-miR-151-5p, hsa-miR-200a, hsa-miR-145, hsa-miR-29a, hsa-miR-106b, hsa-miR-595, hsa-miR-142-5p, hsa-miR-99a, hsa-miR-20b, hsa-miR-373, hsa-miR-502-5p, hsa-miR-29b, hsa-miR-142-3p, hsa-miR-663, hsa-miR-423-5p, hsa-miR-15a, hsa-miR-888, hsa-miR-361-3p, hsa-miR-365, hsa-miR-10b, hsa-miR-199a-3p, hsa-miR-181a, hsa-miR-19a, hsa-miR-125b, hsa-miR-7a, hsa-miR-671-5p, hsa-miR-7c, hsa-miR-1979, hsa-miR-103

Prostate Cancer
MFG-E8, FLNA, PTGFRN, EHD1, BASP1, p53, hsa-miR-18a*, hsa-let-7b, hsa-miR-130b, hsa-miR-582-3p, hsa-miR-574-3p, hsa-miR-452, miR-7c, miR-181a, miR-335, miR-331-3p, hsa-miR-196a, hsa-miR-1979, hsa-miR-335, miR-502-5p, miR-574-3p, miR-885-3p

Prostate Cancer
FOXO1A, SOX9, CLNS1A, PTGDS, XPO1, LETMD1, RAD23B, ABCC3, APC, CHES1, EDNRA, FRZB, HSPG2, TMPRSS2, ETV1 fusion

Prostate Cancer
A33, a33 n5, AFP, ALA, ALIX, ALP, AnnexinV, APC, ASCA, ASPH (246-260), ASPH (666-680), ASPH (A-10), ASPH (DOP), ASPH (D03), ASPH (G-20), ASPH (H-300), AURKA, AURKB, B73H, B7H4, BCA-225, BCNP, BDNF, BRCA, CA125 (MUC6), CA19-9, C-Bir, CDL1, CDIO, CD174 (Lewis y), CD24, CD44, CD46, CD59 (MEM-43), CD63, CD66c, CEA, CD73, CD81, CDA, CDAC1 la2, CEA, C-erbB, C-erbB2, CRMP-2, CRP, CXCL12, CYFRA21-1, DLL4, DR3, EGFR, EpCam, EphA2, EphA5 (H-77), ER, Erbb4, EZH2, FASL, FRT, FRT c.123, GDF15, GPCR, GPR30, Gro-alpha, HAP, HBD 1, HB2D, HER 3 (Erbb3), HSP, HSP70, hVEGFR2, ic3b, IL 6 Unc, IL-1B, IL6 Unc, IL6R, INS, IL-8, INSGN-2, KLK2, L1CAM, LANM, LDH, MACC-1, MAPK4, MART-1, MCP-1, M-CSF, MEG-F8, MIC1, MIF, MIS RII, MMG, MMP2, MMP7, MMP9, MS4A1, MUC1, MUC1 seqa, MUC1 seqb, MUC17, MUC2, Ncam, NGAL, NPGP/NPFF2, OPG, OPN, p53, p53, PA2G4, PBP, PSCA, PDGFRB, PGFP9, PIM1, P (B), PRL, PSA, PSMA, PSM1, PTEN, R5-CD9 Tube 1, Reg IV, RUNX2, SCRNI1, seprase, SERPINB3, SPARC, SPB, SPDEF, SRVN, STAT 3, STEAP1, TF (FL-295), TFF3, TG2M, TIMP-1, TIMP2, TIMP2, TMEM211, TMPRSS2, TNF-alpha, Trail-R2, Trail-R4, TRK, TROP2, Tsg 101, TWEAK, UNC93A, VEGF A, YPSMA-1

Prostate Cancer Vesicle Markers
ST4, ACTG1, ADAM10, ADAM15, ALDOA, ANXA2, ANXA6, APOA1, ATPA1, BASP1, Clor58, C20orl54, C5B, CAPZAI, CA1, CD40, CD1A, CD2A, CD59, CD9, CD9, CFL1, CEP, CHMP4B, CLTC, COT1L, CTNN1D, CTTS, CYCS, DPP4, EEF1A1, EHD1, ENO1, F1R, F2, F5, FAM125A, FNBP1L, FOLH1, GAPDH, GLBI, GPX3, HIST1H1C, HIST1H2AB, HSP90AB1, HSP1A1, HSPA8, IGFS8, ITGB1, ITIH3, JUP, LDHA, LDHB, LUM, LYZ, MFGES8, MGAM, MMP9, MYH2, MYL6B, NME1, NME2, NBPBC1, PABPC4, PACSIN2, PCB2, PDCD6IP, PRDX2, PSA, PSMA, PSMA1, PSMA2, PSMA4, PSMA6, PSMA7, PSMB1, PSMB2, PSMB3, PSMB4, PSMB5, PSMB6, PTGFRN, RPS27A, SDCBP, SERINC5, SH3GL1, SLC3A2, SMPDL3B, SNX9, TACSTD1, TCN2, THBS1, TPI, TSG101, TUBB, VDAC2, VPS37B, YWHAQ, YWHAQ, YWHAZ

Prostate Cancer Vesicle Markers
FLNA, DCRN, HER 3 (ErbB3), VCAN, CD9, GAL3, CDADC1, GM-CSF, EGFR, RANK, CSA, PSMA, Chickenylg, B7H3, PCSA, CD63, CD3, MUC1, TGM2, CD81, S100A4, MEG-F8, Integrin, NK-2R(C-21), PSA, CD24, TIMP-1, IL6 Unc, PB, PIM1, CA-19-9, Trail-R4, MMP9, PRL, EphA2, TWEAK, NY-ESO-1, Mammaglobin, UNC93A, A33, AURKB, CD41, XAGE-1, SPDEF, AMACR, seprase/FAP, NGAL, CXCL12, FRT, CD66e
Prostate Cancer

Markers

CEA, SIM2 (C-15), C-Bir, STEAP, PSIP1/LEDGF, MUC17, hVEGFR2, ERG, MUC2, ADAM10, ASPH (A-10), CA125, Gro-alpha, Tsg 101, SXX2, Trail-R4

Prostate Vesicle

Markers

NTSE (CD73), A33, ABL2, ADAM10, AFP, ALA, ALIX, ALPL, AMACR, Apo J/CLU, ASCA, ASPH (A-10), ASPH (D01P), AURKB, B7H3, B7H4, BCNP, BDFN, CA125 (MUC16), CA-19-9, C-Bir (Flagellin), CD10, CD151, CD24, CD3, CD41, CD44, CD46, CD59(MEM-43), CD63, CD66e CEAC, CD81, CD9, CDADCD1, C-erbB2, CRMP-2, CRP, CSA, CXCL12, CXCR3, CYFRA21-1, DCRN, DDX-1, DLL4, EGF, EpCAM, Epha2, ERG, EZH2, FASL, FLNA, FRT, GAL3, GATA2, GM-CSF, Gro-alpha, HAP, HER3 (ErbB3), HSP70, HSPB1, hVEGFR2, i3cB, IL-1B, IL6 R, IL6 Unc, IL7 R, IGF1R, JK-127, JK-18, INSIG-2, Integrin, KLR2, Label, LAMN, Mammmoglobin, M-CSF, MGF-E, MIF, MIS RI, MMP7, MS4A1, MUC1, MUC7, MUC2, Ncam, NDUFB7, NGAL, NK-2R(C-21), NY-ESO-1, p53, PBP, PCA1, PDGFRB, PIM1, PRL, PSA, PSIP1/LEDGF, PSMA, RAGE, RANK, Reg IV, RUNX2, S100-A4, seprase/FAP, SERPINB3, SIM2 (C-15), SPARC, SP, SPDEF, SPP1, SXX2, SXX4, STEAP, STEAP4, TFF3, TGM2, TIMP-1, TIMM211, Trail-R2, Trail-R4, TrkB (poly), Trop2, Tsg 101, TWEAK, UNC93A, VCAN, VEGF A, wnt-5a(C-16), XAGE, XAGE-1

Prostate Vesicle

Membrane

ADAM 9, ADAM10, AGR2, ALDOA, ALIX, ANX1A, ANX2A, ANX4A, ARF6, ATP1A3, B7H3, BCHE, BCL2L14 (Bel G), BCN1P, BDKB2R, BDNFCAV1-Caveolin, CCR2 (CC chemokine receptor 2, CD192), CCR5 (CC chemokine receptor 5), CCT2 (TCP 1-beta), CD10, CD151, CD166/ALCAM, CD24, CD283/TLR3, CD41, CD46, CD49d (Integrin alpha 4, ITG4A), CD63, CD68, CD81, CD90/THY1, CDH1, CDH2, CDKNA1 cyclin-dependent kinase inhibitor (p21), CGA gene (coding for the alpha subunit of glycoprotein hormones), CLDN3- Claudin3, COX2 (PTGS2), CSEIL (Cellular Apoptosis Susceptibility), CXCR3, Cytokeratin 18, Eagg (KCNH1), EDIL3 (del-1), EDNRB - Endothelial Receptor Type B, EGRF, EpoR, EZH2 (enhancer of Zeste Homolog2), EZR, FABP5, Farnesyltransferase/geranylgeranyl diphosphate synthase 1 (GGPS1), Fatty acid synthase (FASN), FTI (light and heavy), GAL3, GDF15-Growth Differentiation Factor 15, G16, GM-CSF, GSTP1, H3F3A, HGF (hepatecty growth factor), hK2 / Kif2a, HSP90AA1, HSPA1A / HSP70-1, HSPB1, IGFBP-2, IGFBP-3, ILalpah, IL-6, IQGAP1, ITGAL (Integrin alpha L chain), Ki67, KLK1, KLK10, KLK11, KLK12, KLK13, KLK14, KLK15, KLK4, KLK5, KLK6, KLK7, KLK8, KLK9, Lamp-2, LDH-A, LGALS3BP, LGALS8, MMP 1, MMP 2, MMP 25, MMP 3, MMP10, MMP-14/MT1-MMP, MMP7, MTAlnAnS, Nav1.7, NKX3-1, Notch1, NRPI / CD304, PAP (ACPP), PGP, PhIP, PIP3 / BPTNT1, PKM2, PKPI (plakophilin3), PLP (plakophilin3), Plasma chromogranin-A (Cga), PRDX2, Prostate secretory protein (PSNP4) / β-Microsominoprotein (MSP) / IGF, PSA, PSMA, PSMA1, PTEN/PPTN13/PTPL1, RPL19 seprase/FAP SET, SLC3A2 / CD98, SRVN, STEAP1, Syndecan/ CD138, TGF, TGM2, TIMP-1 TLR4 (CD284), TLR9 (CD289), TMRSS1 / hepsin, TMRSS2, TNF R1, TNFα, Transferrin receptor/CD71/TRF R, Trop2 (TACSTD2), TWEAK uPA (urokinase plasminoactivate) degrades extracellular matrix, uPAR (uPA receptor) / CD87, VEGFR1, VEGFR2

Prostate Vesicle

Markers

ADAM 34, ADAM 9, AGR2, ALDOA, ANX1A, ANX1A, ANX4A, ANX4A, ANX2A, ARF6, ATP1A1, ATP1A2, ATP1A3, B7H3, BCHE, BCL2L14 (Bel G), BDKB2R, CA125, CAV1-Caveolin, CCR2 (CC chemokine receptor 2, CD192), CCR5 (CC chemokine receptor 5), CCT2 (TCP 1-beta), CD166/ALCAM, CD49b (Integrin alpha 2, ITG4A), CD90/THY1, CDH1, CDH2, CDKNA1 cyclin-dependent kinase inhibitor (p21), CGA gene (coding for the alpha subunit of glycoprotein hormones), CHMP4B, CLDN3- Claudin3, CLSTN1 (Calsyntenin-1), COX2 (PTGS2), CSEIL (Cellular Apoptosis Susceptibility), Cytokeratin 18, Eagg (KCNH1) (plasma membrane-K+-voltage gated channel), EDIL3 (del-1), EDNRB- Endothelial Receptor Type B, Endoglin/CD105, ENOX2 - Ecto-NOX disulphide Thiol exchanger 2, EPAC-2 Early prostate cancer antigen2, EpO2, EZH2 (enhancer of Zeste Homolog2), EZR, FABP5, Farnesyltransferase/geranylgeranyl diphosphate synthase 1 (GGPS1), Fatty acid synthase (FASN, plasma membrane protein), FTI (light and heavy), GDF15-Growth Differentiation Factor 15, G16, GSTP1, H3F3A, HGF (hepatecty growth factor), hK2 (KLK2), HSP90AA1, HSPA1A / HSP70-1, IGFBP-2, IGFBP-3, ILalpah, IL-6, IQGAP1, ITGAL (Integrin alpha L chain), Ki67, KLK1, KLK10, KLK11, KLK12, KLK13, KLK14, KLK15, KLK4, KLK5, KLK6, KLK7, KLK8, KLK9, Lamp-2, LDH-A, LGALS3BP, LGALS8, MMP 1, MMP 2, MMP 25, MMP 3, MMP10, MMP-14/MT1-MMP, MMP7, MTAlnAnS, Nav1.7, NKX3-1, Notch1, NRPI / CD304, PAP (ACPP), PGP, PhIP, PIP3 / BPTNT1, PKM2, PKPI (plakophilin3), PKP3 (plakophilin3), Plasma chromogranin-A (Cga), PRDX2, Prostate secretory protein (PSNP4) / β-Microsominoprotein (MSP) / IGF, PSA, PSMA, PSMA1, PTEN, PTGFRN, PTPL1/PPL1, PKM2, RPL19, SCA-1 / ATXN1, SERINC5/TP01, SET,
Prostate Cancer Treatment
hsa-miR-1974, hsa-miR-27b, hsa-miR-103, hsa-miR-146a, hsa-miR-22, hsa-miR-382, hsa-miR-23a, hsa-miR-376c, hsa-miR-335, hsa-miR-142-5p, hsa-miR-221, hsa-miR-142-3p, hsa-miR-151-3p, hsa-miR-21, hsa-miR-16

Prostate Cancer
let-7d, miR-148a, miR-195, miR-25, miR-26b, miR-329, miR-376c, miR-574-3p, miR-888, miR-9, miR-1204, miR-16-2*, miR-497, miR-588, miR-614, miR-765, miR-92b*, miR-938, let-7f-2*, miR-300, miR-523, miR-525-5p, miR-1182, miR-1244, miR-520d-3p, miR-379, let-7b, miR-125a-3p, miR-1296, miR-134, miR-149, miR-150, miR-187, miR-32-324-3p, miR-324-5p, miR-342-3p, miR-378, miR-378*, miR-384, miR-451, miR-455-3p, miR-485-3p, miR-487a, miR-502-5p, miR-548a-5p, miR-550, miR-562, miR-593, miR-593*, miR-595, miR-602, miR-603, miR-654-5p, miR-877*, miR-886-5p, miR-125a-5p, miR-140-3p, miR-192, miR-196a, miR-2110, miR-212, miR-222, miR-224*, miR-30b*, miR-499-3p, miR-505*

Prostate Cancer (PCSA+ cMVs)
miR-182, miR-663, miR-155, miR-125a-5p, miR-548a-5p, miR-628-5p, miR-517*, miR-450a, miR-920, hsa-miR-619, miR-1913, miR-224*, miR-502-5p, miR-888, miR-376a, miR-542-5p, miR-30b*, miR-1179

Prostate Cancer

Prostate Cancer RAD23B, FBP1, TNFRSF9, CCNG2, NOTCH3, ETVI, B1D, SIM2, LETMD1, ANAX1, miR-519d, miR-647

Prostate Cancer RAD23B, FBP1, TNFRSF9, NOTCH3, ETVI, B1D, SIM2, ANAX1, BCL2

Prostate Cancer ANPEP, ABL1, PSCA, EFNAB1, HSPB1, INMT, TRIP13

Prostate Cancer E2F3, c-met, pRB, EZH2, e-cad, CAXII, CAIX, HIF-la, Jagged, PIM-1, hepsin, RECK, Clusterin, MMP9, MTSP-1, MMP24, MMP15, IGFBP-2, IGFBP-3, E2F4, caveolin, EF-1A, Kallikrein 2, Kallikrein 3, PSGR

Prostate Cancer A2MEL1, BAX, C10orf47, Clilorf62*, CSDA, EIF3C, ETFB, GABARAPL2, GUK1, GZMH, HIST1H3B, HLA-A, HSP90AA1, NRGN, PRDX5, PTMA, RABAC1, RABAGAP1L, RPL22, SAP18, SEPW1, SOX1

Prostate Cancer NY-ESO-1, SSX-2, SSX-4, XAGE-1b, AMACR, p90 autoantigen, LEDGF

Prostate Cancer A33, ABL2, ADAM10, AFP, ALA, ALIX, ALPL, ApoJ/CLU, ASCA, ASPH(A-IO), ASPH(DOIP), AURKB, B7H3, B7H3, B7H4, BCNP, BDNF, CA125(MUC16), CA-19-9, C-Bir, CD10, CD151, CD24, CD41, CD44, CD46, CD59(MEM-43), CD63, CD63, CD66eCEA, CD81, CD9, CD9, CD9, CDA, CDAD1C, CRMP-2, CRP, CXCL12, CXCR3, CYFRA21-1, DDX-1, DDL4, DDL4, EGFR, Epcam, EphA2, ErbB2, ERG, EZH2, FASL, FLNA, FRT, GAL3, GATA2, GM-CSF, Gro-alpha, HAP, HER3, ErbB3, HPS70, HSPB1, hVEGFR2, i3c3b, IL-1B, IL6R, IL6Unc, IL7Ralpha/CD127, IL8, INSIG-2, Integrin, KL2K, LAMN, Mammoglobin, M-CSF, MFG-E8, MIF, MISRII, MMP7, MMP9, MUC1, Muc1, MUC17, MUC2, Ncam, NDUF7, NGAL, NK-2R(C-21), NTSE (CD73), p53, BPB, PCSA, PCSA, PDGFRB, PIM1, PRL, PSA, PSA, PSMA, PSMA, RAGE, RANK, RegIV, RUNX2, S100A-4, seprase/FAP, SERPINB3, SIM2(C-15), SPARC, SPC, SPDEF, SPP1, STEAP, STEAP4, TFF3, TGM2, TIMP-1, TMEM211, Trail-R2, Trail-R4, TrKB(poly), Trop2, Tsgf01, TWEAK, UNC93A, VEGFA, Wnt-5a(C-16)

Prostate Vesicles CD9, CD63, CD81, PCSA, MUC2, MFG-E8


Prostate Cancer hsa-miR-877*, hsa-miR-593, hsa-miR-595, hsa-miR-300, hsa-miR-324-5p, hsa-miR-548a-5p, hsa-miR-329, hsa-miR-550, hsa-miR-886-5p, hsa-miR-603, hsa-miR-490-3p, hsa-miR-938, hsa-miR-149, hsa-miR-1296, hsa-miR-384, hsa-miR-487a, hsa-miRPlus-C1089, hsa-miR-485-3p, hsa-miR-525-5p

Prostate Cancer hsa-miR-451, hsa-miR-223, hsa-miR-593*, hsa-miR-1974, hsa-miR-486-5p, hsa-miR-19b, hsa-miR-320b, hsa-miR-92a, hsa-miR-21, hsa-miR-675*, hsa-miR-16, hsa-miR-876-5p, hsa-miR-144, hsa-miR-126, hsa-miR-137, hsa-miR-1913, hsa-miR-29b-1*, hsa-miR-15a, hsa-miR-93, hsa-miR-1266


Prostate Cancer A33, ADAM10, AMACR, ASPH(A-10), AURKB, B7H3, CA125, CA-19-9, C-Bir, CD24,
CD3, CD41, CD63, CD66e CEA, CD81, CD9, CDADCl, CSA, CXCL12, DCRN, EGRF, EphA2, ERG, FLNA, FRT, GAL3, GM-CSF, Gro-alpha, HER 3 (ErbB3), hVEGFR2, IL6 Unc, Integrin, Mammaglobin, MFG-E8, MMP9, MUC1, MUC17, MUC2, NGAL, NK-2R(C-21), NY-ESO-1, PBP, PCSA, PIM1, PRL, PSA, PSIP1/LEDGF, PSMA, RANK, S100-A4, seprase/FAP, SIM2 (C-15), SPDEF, SSX2, STEAP, TGM2, TIMP-1, Trail-R4, Tsg 101, TWEAK, UNC93A, VCAN, XAGE-1

Prostate Cancer
A33, ADAM10, ALIX, AMACR, ASCA, ASPH (A-10), AURKB, B7H3, BCNP, CA125, CA-19-9, C-Bir (Flagellin), CD24, CD3, CD41, CD63, CD66e CEA, CD81, CD9, CDADCl, CR, CPA, CXCL12, CYFRA21-1, DCRN, EGRF, EpCam, EphA2, ERG, FLNA, GAL3, GATA2, GM-CSF, Gro alpha, HER3 (ErbB3), HSP70, hVEGFR2, ic3b, IL-1B, IL6 Unc, IL8, Integrin, KLK2, Mammaglobin, MFG-E8, MMP7, MMP9, MS4A1, MUC1, MUC17, MUC2, NGAL, NK-2R(C-21), NY-ESO-1, p53, PBP, PCSA, PIM1, PRL, PSA, PSMA, RANK, RUNX2, S100-A4, seprase/FAP, SERPINB3, SIM2 (C-15), SPDEF, SSX2, SSX4, STEAP, TGM2, TIMP-1, TRAIL R2, Trail-R4, Tsg 101, TWEAK, VCAN, VEGF A, XAGE

Prostate Vesicles
EpCam, CD81, PCSA, MUC2, MFG-E8

Prostate Vesicles
CD9, CD63, CD81, EpCam

Prostate Cancer
let-7d, miR-148a, miR-195, miR-25, miR-26b, miR-329, miR-376c, miR-574-3p, miR-888, miR-9, miR1204, miR-16-2*, miR-497, miR-588, miR-614, miR-765, miR-92b*, miR-938, let-7f-2*, miR-300, miR-525, miR-525-5p, miR-1182, miR-1244, miR-200d-3p, miR-379, let-7b, miR-125a-3p, miR-1296, miR-134, miR-149, miR-150, miR-187, miR-32, miR-324-3p, miR-324-5p, miR-342-3p, miR-342-5p, miR-378, miR-378*, miR-384, miR-451, miR-455-3p, miR-485-3p, miR-487a, miR-490-3p, miR-502-5p, miR-502-5p, miR-548a-5p, miR-550, miR-562, miR-593, miR-593*, miR-595, miR-602, miR-654-5p, miR-877*, miR-886-5p, miR-125a-5p, miR-140-3p, miR-192, miR-196a, miR-2110, miR-212, miR-222, miR-224*, miR-30b*, miR-499-3p, miR-505*
<table>
<thead>
<tr>
<th>Prostate Markers (Ensembl ENSG identifiers)</th>
</tr>
</thead>
<tbody>
<tr>
<td>E.005893 (LAMP2), E.006756 (ARSD), E.010539 (ZNZF200), E.014257 (ACPP), E.015133 (CCDC83C), E.018699 (TTC27), E.044115 (CTNNA1), E.048828 (FAM120A), E.051620 (HEBP2), E.056097 (ZFR), E.060339 (CCAR1), E.063241 (SOCC2), E.064489 (MEF2BN-MEF2F), E.064886 (CHIL3L2), E.066455 (GOLGA5), E.069535 (MAOB), E.072042 (RDH11), E.072071 (LPHN1), E.074047 (GL2), E.076248 (UNG), E.076554 (TDP52), E.077147 (TM9SF3), E.077312 (SNRPA), E.081842 (PCDHA6), E.086717 (PPEF1), E.088888 (MAVS), E.089930 (XRN2), E.089902 (RCOR1), E.090612 (ZNZF26B), E.092199 (HRNPC), E.095380 (SNRPA), E.099783 (HRNPNP), E.100191 (SLCSA4), E.100216 (TOM22), E.100242 (SUN2), E.100284 (TOM1), E.100401 (RANGAPI), E.100412 (ACO2), E.100836 (PABPN1), E.102054 (RBBP7), E.102103 (PQBP1), E.103599 (IQCH), E.103978 (TMEM87A), E.104177 (MYEF2), E.104228 (TRIM35), E.105428 (ZNFR4), E.105518 (TMEM205), E.106603 (C7orf44; COA1), E.108405 (P2RX1), E.111057 (KRT18), E.111218 (PRMTS), E.112081 (SRSF3), E.112144 (ICK), E.113013 (HSPA9), E.113368 (LMNB1), E.115221 (ITGB6), E.116096 (SPR), E.116754 (SRFS1), E.118197 (DDX59), E.118898 (PPL), E.119121 (TRPM6), E.119711 (ALDH6A1), E.120656 (TAF12), E.121671 (CRY2), E.121774 (KHDDB3), E.122126 (OCRL), E.122556 (HRNPCA2B1), E.123901 (GPR38), E.124562 (SNRPC), E.124788 (ATXN1), E.124795 (DEK), E.126246 (CLYBL), E.126883 (NUP214), E.127616 (SMARCA4), E.127884 (ECHS1), E.128050 (PAICS), E.129351 (ILF3), E.129757 (CDKN1C), E.130338 (TULP4), E.130612 (CYP2G1P), E.131351 (HAUSB), E.131467 (PSE3), E.133315 (MACROD1), E.134452 (FBXO18), E.134851 (TMEM165), E.135940 (COX5B), E.136169 (SESTD2), E.136807 (CDK9), E.137727 (ARHGAP20), E.138031 (ADCY3), E.138050 (THUMP2), E.138069 (RAB1A), E.138594 (TMOD3), E.138760 (SCARB2), E.138796 (HADH), E.139613 (SMARCC2), E.139684 (ESD), E.140263 (SORD), E.140350 (ANP32A), E.140632 (GLYR1), E.142765 (SYTL1), E.143621 (ILF2), E.143933 (CALM2), E.144410 (CPO), E.147127 (RAB41), E.151304 (SRFBP1), E.151806 (GUF1), E.152207 (CYSLTR2), E.152234 (ATP5A1), E.152291 (TGN2L2), E.154272 (ATP5J), E.156467 (UCRBP), E.159387 (IRX6), E.159761 (C1orf96), E.161813 (LARP4), E.162613 (FUBP1), E.162694 (EXTL2), E.162562 (NDUFQ6), E.16713 (COQ4), E.16753 (CDT1), E.167772 (ANGPLTL4), E.167978 (SRRM2), E.168196 (ZNZF68), E.169763 (PRYP3), E.169789 (PRY), E.169807 (PRY2), E.170017 (ALCAM), E.17044 (HRNPCA3), E.170313 (STX8), E.170954 (ZNZF415), E.170959 (DCDC5), E.171302 (CANT1), E.171643 (S100Z), E.172283 (PRYP4), E.173961 (MRPL52), E.172867 (KRT2), E.173366 (TL9J), E.173599 (PC), E.177595 (PIDD), E.178473 (UCN3), E.179981 (TSHZ1), E.181163 (NPM1), E.182319 (Tyrosine-protein kinase Sgk223), E.182795 (Clorfil6), E.182944 (EWRS1), E.183092 (BEGAIN), E.183098 (GPC6), E.184254 (ALDH1A3), E.185619 (PCGF3), E.186889 (TMEM17), E.187837 (HIST11H1C), E.188771 (Clorfil4), E.188906 (HIF1), E.196491 (XRCC6), E.196436 (NPIPL2), E.196504 (PRPF40A), E.196796, E.196993, E.197451 (HRNPCA), E.197746 (PSAP), E.198131 (ZNFS44), E.198156, E.198732 (SMOC1), E.198793 (MOTOR), E.039068 (CDH1), E.173230 (GOLGB1), E.124193 (SRSP6), E.140497 (SCAMP2), E.168393 (DTYMK), E.184078 (EIF4EN1F1), E.124164 (VAPB), E.125753 (VAP), E.118260 (CREBI), E.135052 (GOLM1), E.010244 (ZNZF207), E.010278 (CD9), E.047597 (XK), E.049246 (PER3), E.069849 (ATP1B3), E.072506 (HSD17B10), E.081138 (CDH7), E.099785 (MARCH2), E.104331 (IMPAD1), E.104812 (GYSI), E.120868 (APAF1), E.122908 (EIF2C2), E.125492 (BARHL1), E.127328 (RAB3IP), E.127329 (PTPRB), E.129514 (FOXA1), E.129657 (SEC4L1), E.129990 (SYT5), E.132881 (RSGI), E.136521 (NDUFBS), E.138347 (MYPT1), E.141429 (GALNT1), E.144566 (RABSA), E.151715 (TMEM45B), E.152582 (SPEF2), E.154957 (ZNFB1), E.162385 (MAGOH)</td>
</tr>
</tbody>
</table>
Prostate vesicle

LAMP2, ACPP, CTNNAI, HEBP2, ISOC2, HNRNPC, HNRNPM, TOMM22, TOM1, AC02, KRT18, HSPA1A, LMBN1, SPR, PRL, ALDH6A1, HNRNPA2B1, ATXN1, SMARCA4, ECHS1, PAICS, ILF3, PSEME3, COX5B, RABIA, SCARB2, HADH, ESD, SORD, ILF2, CALM2, ATP5A1, TNOL2, ANGPTL4, ALCAM, KRT2, PC, NPM1, Clorf16, GPC6, ALDH1A3, HISTI1H1C, XRC6, HNRNPA8, POSA, C11H1, SCAMP2, VASP, CD9, ATP1B3, HSD17B10, APAFI, EIF2C2, RAB5A, CFL2, FARS2, XPNPEP3, ENTPD4, APL2, NUCBI, RAB3D, VEGFA, HPS3, TNSAXIP1, HNRPNL, PSMB7, GNA12, NONO, FOLHI, PRKAR2A, PHB, HIST3H3, MAP7, VCP, U2AF2, FUS, FKBP5, NDRG1, ATP2A3, NCL, RPL36, KRT8, C1GALT1C1, FASN, PTBP1, TNXDC6, DNAJC5, SLC37A2, HNRNPK, VDAC2, PRDX2, TALDO1, USP14, PSMD7, HSP1, DNAI1, BAI1, WYHAZ, RAB3B, COROIB, MH2, HISTH13A, LAMP1, STC2, DSTN, SLC20A2, ENPP4, WIZ, HSP90AB1, ID3B3, ECH1, C1QBP, SET, TNSF15, ITGB7, SPOCK1, EIF4A2, C3T, CLDN3, EEF2, LLRRC7, RUVBL2, CLDN5, APPL2, TM9SF2, EIF4A3, DBI, DBF4B, SVIP, CD151, ALOY5, SLC9A3R2, RAB27B, DGL1, ARCN1, CHICD3, RAB5B, RPS25, RPL10, DDAH1, HSP90B1, CTNNBI, PSMD2, PKP3, FLNB, EFTUD2, GL0I, PRKCSH, TMBIM1, SEC31A, TEMED10, RPL14, MATR3, APEX1, B4GALT1, HNRNPA1, CFD, HSPA1A, CAPNI, CHRD1, SPEN, SDF4, NAPA, SYNGR2, CHMP3, CND2, CDC64B, SERINC5, VPS37C, DNPEP, CLDN7, KTN1, SERPINB6, ATP5B, CANX, AKT1, TTBK2, DD1X, DLD, LNPEP, LTBP2, LRRPCC, EPS8, AZGP1, VPS28, DHC7, CIB1, DIX39B, HIST1H4B, UGDH, HSPD1, B2M, Tollip, CD276, CYCS, CUL3, GI2D, LLGL2, XRC5, CTNN, PHGDH, CST3, RBL2, SCLAI5, CD46, VAMP8, CLTA, ACSL3, MRPS26, SNX9, GLUD1, TMED4, PTTPN3, AP1G1, SYT9, DCTN2, ID2, GLUD2, TMED9, CLDN4, GM2A, CD2AP, MBM5.
Prostate Cancer vesicles
EGFR, GLUD2, ANXA3, APLP2, BclG, Cofilin 2/cFL2, DCTN-50 / DCTN2, DDAH1, ESD, FARSLA, GITRL, PRKCSH, SLCA20A2, Synaptogyrin 2/SYNGR2, TM9SF2, Calnexin, TOMM22, NDRG1, RPL10, RPL14, USP14, VDAC2, LLGL2, CD63, CD81, uPAR / CD87, ADAM 9, BDKRB2, CCR5, CRT (TCPI-beta), PSMA, PSMA1, HSPB1, VAMP8, Raba1, B4GALT1, Aspartyl Aminopeptidase /Dnpep, ATPase Na+K+ beta 3/ATP1B3, BDNF, ATBP, beta 2 Microglobulin, Calmodulin 2 /CALM2, CD9, XRC5C / Ku80, SMARC4, TOM1, Cytochrome C, Hsp0 /HSPC1, COX2 / PTGS2, Claudin4 /CLDN4, Cystokeratin 8, Cytokeratin 18, Cytokeratin 19, CTNND1 / delta 1-catenin /pl20-catenin, CTNND2, CXADR, CXCR3, CYCS, Cystatin C, Cytochrome C, FHL1, FTH1, HSF1, KRT19, TUBA3C, TUBA1C, TUBA3D, AARS, COMT, MUM1L1, NBL1, PRKACB, GGCT, PRDX6, DHX9, TUBA3E, TUBA1C, TUBA3C, ERP29, SOD2, KRT7, TUBA3D, AARS, COMT, MUM1L1, CDH5, ECE1, ACAT1, ENDOD1, TUBA8, ETFB, NME2, CS, VBP1, RAB9A, TXNRD1, LIF, B2AP12, HISTI1H3, GRN, HIBADH, H3F3B, CUL4B, HNRNPR, WHYAQ, PKHD1, TUBA1A, PARK7, ERLIN2, PDI4A, TUBA4A, PRKCD, ANXA3, H3F3A, PTP4A2, PDLA3, ETFA, CYB5R1, CRTAP, OXR1, YES1, EPCAM, ARHGDI1, DIABLO, SLC9A3R1, BLVRB, P4HA1, HISTI1H3B, ACTN4, UBC, FH, HIST4H4, TUBA1B, HSD17B4, PIK3CA, FLOT2, LMNA, TMEM192, HIST2H4B, YBX1, EIF3A, FLOT1, UTRN, HKI, ACLY, ATIC, WHYAHG, GNG5, GD1, HNRNPH2, NEDD4, BTN3A3, SLC16A1, HGS, ACTN2, SRM, PCNA, ACSL1, RAD21, ARHGAP11, IGF2R, WHYAE, ACTN1, EIF4G1, EPHX2, EIF4E, FTHI, CXADR, MTHFD1, AKR1A1, STXB4P, AHNK, MUC1, RPS27A, UBB, PDLIM5, FAM129A, SND1, FUC2, CRAYAB, EZR, TJP1, ANXA4, GPI, APAK9, CD44, GCN11D, ACTB, FLB, NLUDD3

Prostate Cancer vesicles
14-3-3 zeta/beta, Aconitase 2, ADAM 9, ADAM10, A2DE, AFM2, Ago2, ARG2, AKT1, ALDH1A3, ALDH6A1, ALDOA, ALIX, ANGPTL4, ANXA1, ANXA2, ANXA3, ANXA4, AP1G1, APAF1, APE1, APLP2, APLF2, ARF6, Aspartyl Aminopeptidase /Dnpep, Ataxin 1, ATP5A1, ATPase Na+/K+ alpha 3/ATP1A3, ATPaseta Na+/K+ beta 3/ATP1B3, ATPase Na+/K+ beta 3/ATPB3, ATPase Na+/K+ beta 3/ATP1B, ATPase Na+/K+ beta 3/ATP1B3, ATPase Na+/K+ beta 3/ATP1B, ATPase Na+/K+ beta 3/ATP1B3, ATPase Na+/K+ beta 3/ATP1B3, ATPase Na+/K+ beta 3/ATP1B3, B4GALT1, B7H3, BCHE, BclG, BDKRB2, BDNF, BDNF, beta 2 Microglobulin, beta catenin, BRG1, CALM2, Calmodulin 2 /CALM2, Calnexin, Calpain 1, Catenin Alpha 1, Catenin B, Catenin C, CD101, CD151, CD166/ALCAM, CD24, CD276, CD41, CD46, CD49d, CD63, CD71 / TRFR, CD81, CD9, CD90/THY1, CDH1, CDH2, CDKN1A, CGA, CgA, CHRD1L2, CIB1, CIB1, Claudin 4 /CLDN4, Claudin 5, CLDN3, CLDN3- Claudin3, CLDN4, CLDN7, CNDP2, Coatomer Subunit Delta, Coflin 2/cFL2, CORO1B, Contactin/CTTN, COX2 / PTGS2, COX5b, CSE1L, CTH, CTNND1 / delta 1-catenin /pl20-catenin, CTNND2, CXCR3, CYCS, Cystatin C, Cytochrome C, Cytokeratin 2,
| Protein/Family | Cytokeratin basic, DBF4B /DBF1, DB1*, DCTN-50 / DCTN2, DDAH1, DDAH1, DD1X1, Destrin, DIP13B /appl2, DIP13B /appl2, DLG1, Dnpep, E-Cadherin, ECH1, ECH1S, ECH1S1, EDIL3 (del-1), EDN-3, EDNRB/EDN3, EGFR, EIF4A3, ENTPD4, ErpR, ErpQ, ERAB, ESD, ESD, ETS1, ETS1, ETS-2, EZH2, EZH2 /KMT6, EYR, FABP5, FARSLA, FASN, FKBP5/FKBP51, FLNB, FTL (light and heavy), FUS, GAL3, gamma-catenin, GDF15, GI2D, GGSP1, GGSP1, GITRL, Gol, GLUD2, GM2A, GM-CSF, GOLM1/GOLPH2, Mab; clone 3B10, GOLPH2, GOLPH2, GPC6, GRP94, GSTP1, GSTP1, H3F3A, HADH/HADHSc, HGF, HIST3H3, Histone H4, hK2 /Kif2a, hrNRP A1, hrNRP A2B1, hrNRP C1 + C2, hrNRP K (F45)*, hrNRP L, hrNRP M1-M4, HOXB13, HspO / HSPE1, HspO4/DNAJB1, HspO60, HSP90AA1, Hsp90B, HSP90A1, HSPB1, IDH2, IDH3B, IDH3B, IGFBP-2, IGFBP-3, IgG, IgG2A, IgG2B, ILalpha, ILalpha, Integrinbeta 7, IQGAP1, ITGAL, KLHL12/C3IP1, KLK1, KLK10, KLK11, KLK12, KLK13, KLK14, KLK15, KLK4, KLK5, KLK6, KLK7, KLK8, Ku70 (XRCC6), LaminiB1, LAMPI, Lamp-2, LDH-A, LGALS3BP, LGALS8, Lipoamide Dehydrogenase, LGG2, LSP1, LSP1, LTBP2, MATR3, MBD5, MDH2, MDM4, MEI, MKI67/Kit67, MMP 1, MMP 2, MMP 25, MMP10, MMP-14/M1-MMP, MMP3, MMP7*, Mortalin, MTA1, nAnS, nAnS, Nav1.7/SNC9A, NCL, NDRG1, NXX-3-1, NONO, Notch1, NOTCH4, NRPI / CD304, Nucleobindin-1, Nucleophosmin, p30 / RBL2, p97 / VCP, Pap - same as ACPP, PHGDH, PhIP, PIP3 / BPNT1, PKA R2, PKM2, PKM2, PKP1, PKP3, PLEKH1 / Kindlin-2, PRDX2, PRKCSH, Prohibitin, Proteasome 195 10b, Proteasome 205 2007, PSA, PSA, PSA, PSA, PSA1, PSA1, PSMD7, PSMD7, PSME3, PSPI4 / MSP / IGBF, PTBP1, PTEN, PTPN13/PTPL1, Rabla, RAB3B, Rab5a, Rad51b, RPL10, RPL10, RPL14, RPL14, RPL19, RUVBL2, SCARB2, seprase/FAP, SerpinB6, SET, SH3PX1, SLC20A2, SLC3A2 / CD98, SLC9A3R2, SMARCA4, Sorbitol Dehydrogenase, SPEN/RBM15, SPOCK1, SPR, SRV, Striational, S100A15 / STC2, STEAP1, Synaptogyrin 2 /SYNGR2, Syndecan, SYNGR2, SYT9, TAFIB / GRHL1, TBX5, TGBF, TGM2, TGN46 /TGOLN2, TMP-1, TLR3, TLR4 (CD284), TLR9 / CD289, TM9SF2, TM9SF2B, TM9SF3, TM9SF3, TM9SF3, TNFR1, TNFR1, TNFSF18 / GITLF, TNFa, Tolllip, Tom1, TOMM22, Trop2 / TACSTD2, TNSAXIP1, TWEAK, U2AF2, uPA, uPA / CD87, USP14, USP14, VAMP8, VAS, VDAC2, VEGFA, VEGFR1/FLT1, VEGFR2, VP528, XRC5 / Ku80, XRC5 / Ku80 |
| Ribonucleoprotein complexes & vesicles | GW182, Ago2, miR-let-7a, miR-16, miR-22, miR-148a, miR-451, miR-92a, CD9, CD63, CD81 |
| Prostate Cancer vesicles | PCSA, Muc2, AdamO |
| Prostate Cancer vesicles | Alkaline Phosphatase (AP), CD63, MyoD1, Neuron Specific Enolase, MAP1B, CNPase, Prohibitin, CD45RO, Heat Shock Protein 27, Collagen II, LaminiB1/b1, Gail, CDw75, bcl-XL, Lamini-s, Ferritin, CD21, ADP-ribosylation Factor (ARF-6) |
| Prostate Cancer vesicles | CD56/NCAM-1, Heat Shock Protein 27 /hsp27, CD45RO, MAP1B, MyoD1, CD45/T200/LCA, CD3zeta, Lamini-s, bcl-XL, Rad8, Gail, Thymidylate Synthase, Alkaline Phosphatase (AP), CD63, MMP-16 / MT3-MMP, Cyclin C, Neuron Specific Enolase, SIRP al, Lamini B1/b1, Amyloid Beta (APP), SODD (Silencer of Death Domain), CDC37, Gab-1, E2F-2, Cd6, Mast Cell Chymase, Gamma Glutamylcysteine Synthetase (GCS) |
| Prostate Cancer vesicles | EpCAM, MMAP7, PCSA, BCNP, ADAM10, KLK2, SPDEF, CD81, MFGE8, IL-8 |
| Prostate Cancer vesicles | EpCAM, KLK2, PBP, SPDEF, SSX2, SSX4 |
| Prostate Cancer vesicles | ADAM-10, BCNP, CD9, EGFR, EpCam, IL1B, KLK2, MMP7, p53, PBP, PCSA, SERPINB3, SPDEF, SSX2, SSX4 |
| Androgen Receptor (AR) pathway members in cMVs | GTF2F1, CTNNB1, PTEN, APPL1, GAPDH, CDC37, PNRC1, AES, UXT, RAN, PA2G4, JUN, BAG1, UBE21, HDAC1, COX5B, NCO2, STUB1, HIPK3, PXN, NCOA4 |
| EGFR1 pathway members in cMVs | RALBP1, SH3BGR1, RBP7, REPS1, SNRPD2, CEBPB, APPL1, MAP3K3, EEF1A1, GRB2, RAC1, SNCA, MAP2K3, CEBPA, CDC42, SH3KB1, CBL, PTNPD6, WYHAB, FOXO1, JAK1, KRT8, RALGDS, SMAD2, VAV1, NDUFA13, PRKCB1, MYC, JUN, RFXANK, HDAC1, HIST3H3, PEBP1, PXN, TPIN1, PKN2 |
| TNF-alpha | BCL3, SMACRE1, RPS11, CDC37, RPL6, RPL8, PAPOLA, PSMC1, CASP3, AKT2, -109- |
pathway members in cMVs

MAP3K7IP2, POLR2L, TRADD, SMARCA4, HIST3H3, GNBL1, PSMD1, PEBP1, HSBP1, TNIP1, RPS13, ZFAND5, YWHAQ, COMM1, COP3, POLR1D, SMARC2C, MAP3K3, BIRC3, UBE2D2, HDAC2, CASP8, MCM7, PSMD7, YWHAG, NFkBIA, CAST, YWHA8, G3BP2, PSMD13, RELB, YWHAZ, SKP1, UBE2D3, PDCD2, HSP90AA1, HDAC1, KPN2A, RPL30, GTF2F, PFN2

Colorectal cancer

MAP3K7IP2, POLR2L, TRADD, SMARCA4, HIST3H3, GNBL1, PSMD1, PEBP1, HSBP1, TNIP1, RPS13, ZFAND5, YWHAQ, COMM1, COP3, POLR1D, SMARC2C, MAP3K3, BIRC3, UBE2D2, HDAC2, CASP8, MCM7, PSMD7, YWHAG, NFkBIA, CAST, YWHA8, G3BP2, PSMD13, RELB, YWHAZ, SKP1, UBE2D3, PDCD2, HSP90AA1, HDAC1, KPN2A, RPL30, GTF2F, PFN2

Colorectal cancer

MAP3K7IP2, POLR2L, TRADD, SMARCA4, HIST3H3, GNBL1, PSMD1, PEBP1, HSBP1, TNIP1, RPS13, ZFAND5, YWHAQ, COMM1, COP3, POLR1D, SMARC2C, MAP3K3, BIRC3, UBE2D2, HDAC2, CASP8, MCM7, PSMD7, YWHAG, NFkBIA, CAST, YWHA8, G3BP2, PSMD13, RELB, YWHAZ, SKP1, UBE2D3, PDCD2, HSP90AA1, HDAC1, KPN2A, RPL30, GTF2F, PFN2
Colorectal Cancer
hsa-miR-16, hsa-miR-25, hsa-miR-125b, hsa-miR-451, hsa-miR-200c, hsa-miR-143-3p, hsa-miR-658, hsa-miR-370, hsa-miR-1296, hsa-miR-502-5p

Breast cancer
mir-21, mir-155, mir-206, mir-122a, mir-210, miR-21, let-7, miR-10b, miR-125a, miR-125b, miR-143, miR-145, miR-145-3p, miR-145-5p

Breast cancer
GAS5

Breast cancer
ER, PR, HER2, MUC1, EGFR, KRAS, B-Raf, CYP2D6, hsp70, MART-1, TRP, HER2, hsp70, MART-1, TRP, HER2, ER, PR, Class III b-tubulin, VEGFA, ETVD-NTRK3, BCA-225, hsp70, MART1, ER, VEGFA, Class III b-tubulin, HER2/neu (e.g., for Her2+ breast cancer), GPR30, ErbB4 (JM isoform), MPR8, MISIIR, CD9, EphA2, EGFR, B7H3, PSM, PCSA, CD63, STEAP, CD81, ICAM1, A33, DR3, CD66e, MEG-F8, TROP-2, Mammaglobin, Hepsin, NPGP/NPFF2, PSCA, ST4, NGAL, EpCam, neurokinin receptor-1 (NK-1 or NK-1R), NK-2, P1, CD45, CD10, HER2/ERBB2, AGTR1, NPY1R, MUC1, ESA, CD133, GPR30, BCA225, CD24, CA15.3 (MUC1 secreted), CA27.29 (MUC1 secreted), NMDAR1, NMDAR2, MAGEA, CTAG1B, NY-ESO-1, SPB, SPC, NSE, PGP9.5, progesterone receptor (PR) or its isoform (PR(A) or PR(B)), P2RX7, NDUFB7, NSE, GAL3, osteopontin, CHI3L1, IC3b, mesothelin, SPA, AQ5P, GPCR, hCEA-CAM, PTP IA-2, CAVB4, TMEM211, ADAM28, UNC93A, MUC17, MUC2, IL10-beta, BCMA, HVEM/TNFRSF14, Trappin-2, Elafin, ST2/IL1R4, TNFRF14, CEACAM1, TPA1, LAMP, WF, WH1000, PECA, BSA, TNFR

Breast cancer
CD9, MIS Rii, ER, CD63, MUC1, HER3, STAT3, VEGFA, BCA, CA125, CD24, EPCAM, ERBB4

Breast cancer
CD10, NPGP/NPFF2, HER2/ERBB2, AGTR1, NPY1R, neurokinin receptor-1 (NK-1 or NK-1R), NK-2, MUC1, ESA, CD133, GPR30, BCA225, CD24, CA15.3 (MUC1 secreted), CA27.29 (MUC1 secreted), NMDAR1, NMDAR2, MAGEA, CTAG1B, NY-ESO-1

Breast cancer
SPB, SPC, NSE, PGP9.5, CD9, P2RX7, NDUFB7, NSE, GAL3, osteopontin, CHI3L1, EGFR, B7H3, IC3b, MUC1, mesothelin, SPA, PCSA, CD63, STEAP, AQ5P, CD81, DR3, PSM, GPCR, EphA2, hCEA-CAM, PTP IA-2, CAVB4, TMEM211, ADAM28, UNC93A, A33, CD24, CD10, NGAL, EpCam, MUC17, TROP-2, MUC2, IL10-beta, BCMA, HVEM/TNFRSF14, Trappin-2, Elafin, ST2/IL1R4, TNFRF14, CEACAM1, TPA1, LAMP, WF, WH1000, PECA, BSA, TNFR

Breast cancer
BRC4, MUC-1, MUC 16, CD24, ErbB4, ErbB2 (HER2), ErbB3, HSP70, Mammaglobin, PR, PR(B), VEGFA

Breast cancer
CD9, HSP70, Gal3, MIS, EGFR, ER, ICB3, CD63, B7H4, MUC1, DLL4, CD81, ERB3, VEGF, BCA225, BRC4, CA125, CD174, CD24, ErbB2, NGAL, GPR30, CYFRA21, CD31, cMET, MUC2, ERBB4

Breast cancer
CD9, EphA2, EGFR, B7H3, PSMA, PCSA, CD63, STEAP, CD81, STEAP1, ICAM1 (CD54), PSMA, A33, DR3, CD66e, MEG-F8, TMEM211, TROP-2, EGFR, Mammaglobin, Hepsin, NPGP/NPFF2, PSCA, ST4, NGAL, NK-2, EpCam, NK-IR, PSMA, ST4, PAI-1, CD45
| Breast cancer | PGP9.5, CD9, HSP70, gal3-b2cd0, EGFR, iC3b, PSMA, PCSA, CD63, MUC1, DLLU, CD81, B7-H3, HER 3 (ErbB3), MART-1, PSA, VEGF A, TIMP-1, GPCR GPR110, EphA2, MMP9, mmp7, TMEM211, UNC93a, BRCA, CA125 (MUC16), Mammoglobin, CD174 (Levy s), CD66e CEA, CD24 csn3, C-erbB2, CD 10, NGAL, epcam, CEA (carcinoembryonic Antigen), PGP30, CYFRA21-1, OPN, MUC17, hVEGF, MUC2, NCAM, AsPH, ErbB4, SPB, SP, CD9, MS4A1, EphA2, MIS RII, HER2 (ErbB2), ER, PR (B), MRPA8, CD63, B7H4, TGM2, CD81, DR3, STAT 3, MACC-1, TrkKB, IL 6 Unc, OPG - 13, IL6R, EZH2, SCRNN, TWEAK, SERPINC, CDAC1, BCA-225, DR3, A33, NPGP/NPFF2, TIMPI, BDNF, FPT, Ferritin heavy chain, seprase, p53, LDH, HSP, ost, p53, CXCL12, HAP, CRP, Gro-alpha, Tsg 101, GDF15 |
| Breast Cancer | ST4 (trophoblast), ADAM10, AGER/RAGE, APC, APP (β-amyloid), AsPH (A-10), B7H3 (CD276), BACE1, BA13, BRC1, BDNF, BIRC2, C1GALT1, CA125 (MUC16), Calmodulin 1, CCL2 (MCP-1), CD9, CD10, CD127 (IL7R), CD174, CD24, CD44, CD63, CD81, CEA, CRMP-2, CXCR3, CXCR4, CXCR6, CYFRA 21, derlin 1, DLLL4, DP6, E-CAD, EpCaM, EphA2 (H-77), ER(1) ESR1 a, ER(2) ESR2 b, Erb B4, Erb b2, erb3 (Erb-B3), PA2G4, FRT (FLT1), Ga3, PGP30 (G-coupled ER1), HAP1, HER3, HSP-27, HSP70, IC3b, ILS, insig, junction plakoglobin, Keratin 15, KRAS, Mammoglobin, MARTI, MCT2, MFGE8, MMP9, MRPA8, Mucl, MUC17, MUC2, NCAM, NG2 (CSPG4), Ngal, NHE-3, NTSE (CD73), ODC1, OPG, OPN, p53, PARK7, PCSA, PGP9.5 (PARK5), Pr(B), PSA, PSMA, RAGE, STXB4P, Survivin, TFF3 (secreted), TIMPI, TIM2, TMEM211, TRAF4 (scaffolding), TRAIL-R2 (death Receptor 5), Trk KB, Tsg 101, UNC93a, VEGFA, VEGF, YB-1, VEGFR1, GCPDF-15 (PIP), BigH3 (TGFβ1-inhibited protein), 5HT2B (serotonin receptor 2B), BRCA2, BACE 1, CDH1-cadherin |
| Breast Cancer | AK5.2, ATP6V1B1, CRABP1 |
| Breast Cancer | AK5.2, ATP6V1B1, CRABP1, DST.3, GATA3, KRT81 |
| Breast Cancer | TRAP; Renal Cell Carcinoma; Filamin; 14.3.3.Pan; Prohibitin; c-fos; Ang-2; GSTmu; Ang-1; FHT1; Rad51; Inhibin alpha; Cadherin-P; 14.3.3 gamma; pBINK4c; P504S; XRCC2; Caspase 5; CREB-Binding Protein; Estrogen Receptor; IL7; Claudin2; Keratin 8; GAPDH; CD1; Keratin, LMW; Gamma Glutamylcysteine Synthetase(GCS)/Glutamate-cysteine Ligase; a-B-Crystallin; Pax-5; MMP-19; APC; IL-3; Keratin 8 (phospho-specific Ser73); TGFBeta 2; ITK; Oct-2; DJ-1; B7-H2; Plasma Cell Marker; Rad8; Estroil; Chkl; Prolactin Receptor; Laminin Receptor; Histone HI; G0/G1; G2/M; Actin; Nucleoside Specific; S100; Dystrophin; Tubulin-a; CD3Zeta; CDC37; GABA a Receptor 1; MMP-7 (Matriillin); Heregulin; Caspase 3; CD56/NCAM-1; Gastrin 1; SREBP-1 (Sterol Regulatory Element Binding Protein-1); MLH1; PGP9.5; Factor VII Related Antigen; ADP-ribosylation Factor (ARF-6); MHC II (HLA-DR) Ia; Survivin; CD23; G-CSF; CD2; Calretinin; Neuron Specific Enolase; CD165; Calponin; CD95 / Fas; Urocortin; Heat Shock Protein 27/hsp27; Topo II beta; Insulin Receptor; Keratin 5/8; sm; Actin, skeletal muscle; CA19-9; GruRl; GRIP1; CD79a mb-1; Tdt; HRP; CD94; CCK-8; Thymidine Phosphorylase; CD57; Alkaline Phosphatase (AP); CD59 / MACTF / MIRL / Protectin; GLUT-1; alpha-1-antitrypsin; Presenilin; Mucin 3 (MUC3); pS2; 14-3-3 beta; MMP-13 (Collagenase-3); Fli-1; mGlur5; Mast Cell Chymase; Laminin B1/b; Neurofilament (160kDa); CNPase; Amylin Peptide; Gail; CD6; alpha-1-antichymotrypsin; E2F-2; MyoD |
| Ductal carcinoma in situ (DCIS) | Laminin B1/b; E2F-2; Tdt; ApolipoproteinD; Granulocyte; Alkaline Phosphatase (AP); Heat Shock Protein 27/hsp27; CD95 / Fas; ps2; Estroil; GLUT-1; Fibronecin; CD6; CCK-8; sm; Factor VIII Related Antigen; CD57; Plasminogen; CD71 / Transferrin Receptor; Keratin 5/8; Thymidine Phosphorylase; CD45/T200/LCA; Epithelial Specific Antigen; Macrophage; CD10; MyoD1; Gail; bcl-XL; Ipl; Caspase 3; Actin, skeletal muscle; IP10/CRG2; GnRH Receptor; p53nck5a; ADP-ribosylation Factor (ARF-6); Cdk4; alpha-1-antitrypsin; IL17; Neuron Specific Enolase; CD56/NCAM-1; Prolactin Receptor; Cdk7; CD79a mb-1; Collagen IV; CD94; Myeloid Specific Marker; Keratin 10; Pax-5; IgM (m-Heavy Chain); CD45RO; CA19-9; Mucin 2; Glucagon; Mast Cell Chymase; MLH1; CD1; CNPase; Parkin; MHC II (HLA-DR) Ia; B7-H2; Chkl; Lambda Light Chain; MHC II (HLA-DR and D); Myogenin; MMP-7 (Matriillin); Topo II beta; CD53; Keratin 19; Rad8; Ret Oncoprotein; MHC II (HLA-DP); E3-binding protein (ARM1); Progesterone Receptor; Keratin 8; IgG; IgA; Tubulin; Insulin Receptor Substrate-1; Keratin 15; DR3; IL-3; Keratin 10/13; CyclinD3; MHC I (HLA25 and HLA-Aw32); Calmodulin; Neurofilament (160kDa) |
| Ductal carcinoma | Macrophage; Fibronecin; Granulocyte; Keratin 19; CyclinD3; CD45/T200/LCA; EGFR |
### Breast cancer

<table>
<thead>
<tr>
<th>Gene</th>
<th>Function</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD81/TAPA-1</td>
<td>Ruv C</td>
<td>Plasminogen</td>
</tr>
<tr>
<td>Laminin IV</td>
<td>Laminin B/B</td>
<td>CDIO</td>
</tr>
<tr>
<td>TdT</td>
<td>Filamin b/c</td>
<td>CDIO</td>
</tr>
<tr>
<td>in situ (DCIS) v. Thrombospondin; CD81/TAPA-1; Ruv C; Plasminogen; Collagen IV; Laminin B/B; TdT; Filamin b/c; CDIO;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>other Breast cancer</td>
<td>CDIO;</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ruv C</td>
<td>CD81/TAPA-1</td>
<td>Ruv C</td>
</tr>
<tr>
<td>Collagen IV</td>
<td>Laminin B/B</td>
<td>CDIO;</td>
</tr>
<tr>
<td>CDIO;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>other Breast cancer</td>
<td>CDIO;</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Lung cancer

<table>
<thead>
<tr>
<th>Gene</th>
<th>Function</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prognostic, CD23, Amylin Peptide, HRP, Rad51, Pax-5, Oct-3, GLUT-1, PSCA, Prognostic, CD23, Amylin Peptide, HRP, Rad51, Pax-5, Oct-3, GLUT-1, PSCA,</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Thrombospondin, a-B-Crystallin, Lewis A, Vascular Endothelial Growth Factor (VEGF), Hepatocyte Factor Homologue-
| Factor (VEGF), Hepatocyte Factor Homologue-4, Fli-4, Glur67, Prostate Apoptosis |
| Response Protein-4, GluR1, Urocin, S100A4, 14-3-3 beta, P504S, HDAC1, PGP9.5, D1-1, COX2, MMP-19, Actin, skeletal muscle, Claudin 3, Cadherin-P, Collagen IX, |
| p27Kipl, CathepsinD, CD30 (Reed-Sternberg Cell Marker), Ubiquitin, FSH-b, TrxR2, |
| CCK-8, Cyclin C, CD138, TGF-beta 2, Adrenocorticotropic Hormone, PPAR-gamma, Bel-6, GLUT-1, IGF-I, mRANKL, Fas-ligand, Filamin, Calretinin, O-ct-1, Parathyroid Hormone, |
| Claudin 5, Claudin 4, Raf-1 (Phospho-specific), CDC14A, Phosphatase, Mitochondria, APC, Gastrin 1, Ku (p80), Gail, XPA, Maltose Binding Protein, Melanoma (p600), Phosphotyrosine, Amyloid A, CXC4/Fusin, Hepatic Nuclear Factor-3B, Caspase 1, HPV 16-17, Axonal Growth Cones, Lck, Ornthine Decarboxylase, Gamma Glutamylcysteine Synthetase (GCS)/Glutamate-cysteine Lyase, ERCC1, Calmodulin, Caspase 7 (Mch 3), CD137 (4-1BB), Nitric Oxide Synthase, brain (bNOS), E2F-2, IL-10, L-Plastin, CD18, Vimentin, CD50/ICAM-3, Superoxide Dismutase, Adenovirus Type 5 E1A, PHAS-I, Progesterone Receptor (phospho-specific) - Serine 294, MHC II (HLA-DQ), XPG, ER Ca2+ ATPase2, Laminin-s, E3-binding protein (ARM1), CD45RO, CD1, Cd2k, MMP-10 (Stromelysin-2), sm, Surfactant Protein B (Pro), Apolipoprotein D, CD46, Keratin 8 (phospho-specific Ser73), PCNA, PLAP, CD20, Syk, LH, Keratin 19, ADP-ribosylation Factor (ARF-6), Int-2 Oncoprotein, Luciferase, AIF (Apoptosis Inducing Factor), Grb2, bcl-X, CD16, Paxillin, MHC II (HLA-DF and DR), B-Cell, p21WAF1, MHC II (HLA-DR), Tyrosinase, E2F-1, Pds, Calponin, Notch, CD26/DPP IV, SV40 Large T Antigen, Ku (p70/p80), Perforin, XPF, SIM A (SIMA-4D3), Cdklp/34cdc2, Neuron Specific Enolase, b-2-Microglobulin, DNA Polymerase Beta, Thyroid Hormone Receptor, Human, Alkaline Phosphatase (AP), Plasminogen, |
| Prostate, Heat Shock Protein 70/hs70, TRP75 / gp75, SRF |

---

-113-
<table>
<thead>
<tr>
<th>Disease Markers</th>
<th>Lung Cancer</th>
<th>Ovarian Cancer</th>
<th>Prostate Cancer</th>
<th>Breast Cancer</th>
<th>Colorectal Cancer</th>
<th>Lung Cancer</th>
<th>Prostate Cancer</th>
<th>Integins</th>
<th>Glycoprotein</th>
<th>Transcription factors</th>
<th>Kinases</th>
<th>Disease Markers</th>
</tr>
</thead>
</table>
| Serum Response Factor, LamininBl, Mast Cell Chymase, Caldesmon, CEA / CD66e, CD24, Retinoid X Receptor (hRXR), CD45/T200/LCA, Rabies Virus, Cytochrome c, DR3, bcl-XL, Fascin, CD71 / Transferrin Receptor | miR-497 | CA-125, CA 19-9, c-reactive protein, CD95(also called Fas, Fas antigen, Fas receptor, FasR, TNFRSF6, APT1 or APO-1), FAP-1, miR-200 microRNAs, EGFR, EGFRIII, apolipoprotein AI, apolipoprotein CHI, myoglobin, tenasin C, MS6H, claudin-3, claudin-4, caveolin-1, coagulation factor III, CD9, CD36, CD37, CD53, CD63, CD81, CD136, CD147, Hsp70, Hsp90, Rab13, Desmocollin-1, EMP-2, CK7, CK20, GCDF15, CD82, Rab-5b, Annexin V, MFG-E8, HLA-DR. MiR-200 microRNAs (miR-200a, miR-200b, miR-200c), miR-141, miR-429, JNK, Jun | AQP2, BMP5, C16orf56, CXCL13, DST, ERCC1, GNAO1, KLHL5, MAP4K1, NELL2, PENK, PGF, POU3F1, PRSS21, SCLM1, SEMG1, SMARCD3, SNA2, TAC1, TNNT3 | ADRB2, ARG2, C22orf32, CYorf4, EIF1AY, FEV, KLK2, KLK4, LRRRC26, MAOA, NLGN4Y, PNPLA7, PVRL3, SIM2, SLC30A4, SLC45A3, STX19, TRIM36, TRPM8 | ADRB2, BAIAP2L2, C19orf33, CDX1, CEACAM6, EEF1A2, ERN2, FAM110B, FOXA2, KLK2, KLK4, LOC389816, LRRRC26, MUC1, PNPLA7, SIM2, SLC45A3, SPDEF, TRIM31, TRIM36, ZNF613 | ASTN2, CAB39L1, CRIP1, FAM110B, FEV, GSTP1, KLK2, KLK4, LOC389816, LRRRC26, MUC1, PNPLA7, SIM2, SLC45A3, SPDEF, TRIM36, TRPV6, ZNF613 | miR-26a+b, miR-15, miR-16, miR-195, miR-497, miR-424, miR-206, miR-342-5p, miR-186, miR-1271, miR-600, miR-216b, miR-519 family, miR-203 | ITGAI (CD49a, VLA1), ITGAI2 (CD49h, VLA2), ITGAI3 (CD49c, VLA3), ITGA4 (CD49d, VLA4), ITGA5(CD49e, VLA5), ITGA6 (CD49f, VLA6), ITGA7 (FLJ25220), ITGA8, ITGA9 (RLC), ITGA10, ITGA11 (HsT19864), ITGAD (CD11D, FLJ39841), ITGAE (CD103, HUMNAEC), ITGAL (CDila, LFA1A), ITGAM (CDlb, MAC-1), ITGA8 (CD51, VNRA, MSK8), ITGAW, ITGAX (CDlic), ITGB1 (CD29, FNBR, MSK12, MDF20), ITGB2 (CD18, LFA-1, MAC-1, MF17), ITGB3 (CD61, GP3A, GPlIIIa), ITGB4 (CD104), ITGB5 (FLJ26658), ITGB6, ITGB7, ITGB8 | Gpfa-Ila, GpIB-IIla, GpIib, GpIb, GpIX | STAT3, EZH2, p53, MACC1, SPDEF, RUNX2, YB-1 | AURKA, AURKB | 6Chkine, Adiponectin, Adrenocorticotropic Hormone, Agouti-Related Protein, Aldose Reductase, Alpha-1-Antichymotrypsin, Alpha-1-Antitrypsin, Alpha-1-Microglobulin, Alpha-2-Macroglobulin, Alpha-Fetoprotein, Amphiregulin, Angiogenin, Angiopoietin-2, Angiotensin-Converting Enzyme, Angiotensinogen, Annexin AI, Apolipoprotein A-I, Apolipoprotein A-II, Apolipoprotein A-IV, Apolipoprotein B, Apolipoprotein C-I, Apolipoprotein C-III, Apolipoprotein D, Apolipoprotein E, Apolipoprotein H, Apolipoprotein(a), AXL Receptor Tyrosine Kinase, B cell-activating Factor, B Lymphocyte Chemoattractant, B-2-like protein 2, Beta-2-Microglobulin, Betacellulin, Bone Morphogenetic Protein 6, Brain-Derived Neurotrophic Factor, Calbindin, Calcitonin, Cancer Antigen 125, Cancer Antigen 15-3, Cancer Antigen 19-9, Cancer Antigen 72-4, Carcinoembryonic Antigen, Cathepsin D, CD 40 antigen, CD40 Ligand, CD5 Antigen-like, Cellular Fibronectin, Chemokine CC-4, Chromogranin-A, Ciliary Neurotrophic Factor, Clusterin, Collagen IV, Complement C3, Complement Factor H, Connective Tissue Growth Factor, Cortisol, C-Terpeptide, C-Reactive Protein, Creatine Kinase-MB, Cystatin-C, Endoglin, Endostatin, Endothelin-1, EN-RAGE, Eotaxin-1, Eotaxin-2, Eotaxin-3, Epidermal Growth Factor, Epiuregulin, Epithelial cell adhesion molecule, Epithelial-Derived Neutrophil-Activating Protein 78, Erythropoietin, E-Selectin, Evasion Factor, Factor VII, Fas Ligand, FASLG Receptor, Fatty Acid-Binding Protein (adipocyte), Fatty Acid-Binding Protein (heart), Fatty Acid-Binding Protein (liver), Ferritin, Feto-A, Fibrinogen, Fibroblast Growth Factor 4, Fibroblast Growth Factor basic, Fibulin-IC, Follicle-Stimulating Hormone, Galectin-3, Gelsofin, Glucagon, Glucagon-like Peptide 1. Glucose-6-phosphate Isomerase, Glutamate-Cysteine Ligase Regulatory subunit, Glutathione S-Transferase alpha, Glutathione S-Transferase Mu 1, Granulocyte Colony-Stimulating Factor, Granulocyte-Macrophage Colony-Stimulating Factor, Growth Hormone, Growth-Regulated alpha protein, Haptoglobin, HE4, Heat Shock Protein 60, Heparin-Binding EGF-Like Growth Factor, Hepatocyte Growth Factor, Hepatocyte Growth Factor Receptor, Hepsin, Human Chorionic Gonadotropin beta, Human Epidermal Growth Factor Receptor 2, Immunoglobulin A, Immunoglobulin E, Immunoglobulin M, Insulin, Insulin-like Growth Factor 1, Insulin-like Growth Factor-Binding Protein 1, Insulin-like Growth Factor-Binding Protein 2, Insulin-like Growth Factor-


**Oncology**

| Cardiovascular | Adiponectin, Apolipoprotein A-I, Apolipoprotein B, Apolipoprotein C-III, Apolipoprotein D, Apolipoprotein E, Apolipoprotein H, Apolipoprotein(a), Clusterin, C-Reactive Protein, Cystatin-C, EN-RAGE, E-Selectin, Fatty Acid-Binding Protein (heart), Ferritin, Fibrinogen, Haptoglobin, Immunoglobulin M, Intercellular Adhesion Molecule 1, Interleukin-6, Interleukin-8, Lectin-Like Oxidized LDL Receptor 1, Leptin, Macrophage Inflammatory Protein-1 alpha, Macrophage Inflammatory Protein-1 beta, Malondialdehyde-Modified Low-Density Lipoprotein, Matrix Metalloproteinase-1, Matrix Metalloproteinase-10, Matrix Metalloproteinase-2, Matrix Metalloproteinase-3, Matrix Metalloproteinase-7, Matrix Metalloproteinase-9, Monocyte Chemoattractant Protein 1, Myeloperoxidase, Myoglobin, NT-proBNP, Osteopontin, Plasminogen Activator Inhibitor 1, P-Selectin, Receptor for advanced glycosylation end products, Serum Amyloid P-Component, Sex Hormone-Binding Globulin, T-Cell-Specific Protein RANTES, Thrombomodulin, Thyroxine-Binding Globulin, Tissue Inhibitor of Metalloproteinases 1, Tumor Necrosis Factor alpha, Tumor necrosis Factor Receptor 2, Vascular Cell Adhesion Molecule-1, von Willebrand Factor |
| Inflammatory | Alpha-1-Antitrypsin, Alpha-2-Macroglobulin, Brain-Derived Neurotrophic Factor, Complement C3, C-Reactive Protein, Eotaxin-1, Factor VII, Ferritin, Fibrinogen, Granulocyte-Macrophage Colony-Stimulating Factor, Haptoglobin, Intercellular Adhesion Molecule 1, Interferon gamma, Interleukin-1 alpha, Interleukin-1 beta, Interleukin-
<table>
<thead>
<tr>
<th>Metabolic</th>
<th>Kidney</th>
<th>Cytokines</th>
<th>Protein</th>
</tr>
</thead>
</table>
| Adiponectin, Adrenocorticotropic Hormone, Angiotensin-Converting Enzyme, Angiotensinogen, Complement C3 alpha des arg, Cortisol, Follicle-Stimulating Hormone, Galanin, Glucagon, Glucagon-like Peptide 1, Insulin, Insulin-like Growth Factor 1, Leptin, Luteinizing Hormone, Pancreatic Polypeptide, Peptide YY, Progesterone, Prolactin, Resistin, Secretin, Testosterone | Alpha-1-Microglobulin, Beta-2-Microglobulin, Calbindin, Clusterin, Connective Tissue Growth Factor, Creatinine, Cystatin-C, Glutathione S-Transferase alpha, Kidney Injury Molecule-1, Microalbumin, Neutrophil Gelatinase-Associated Lipocalin, Osteopontin, Tamm-Horsfall Urinary Glycoprotein, Tissue Inhibitor of Metalloproteinases 1, Trefoil Factor 3, Vascular Endothelial Growth Factor | Granulocyte-Macrophage Colony-Stimulating Factor, Interferon gamma, Interleukin-2, Interleukin-3, Interleukin-4, Interleukin-5, Interleukin-6, Interleukin-7, Interleukin-8, Interleukin-10, Interleukin-12, Interleukin-23, Macrophage Inflammatory Protein-1 alpha, Macrophage Inflammatory Protein-1 beta, Matrix Metalloproteinase-2, Matrix Metalloproteinase-3, Matrix Metalloproteinase-9, Monocyte Chemotactic Protein 1, Stem Cell Factor, T-Cell-Specific Protein RANTES, Tissue Inhibitor of Metalloproteinases 1, Tumor Necrosis Factor alpha, Tumor Necrosis Factor beta, Tumor necrosis factor receptor 2, Vascular Cell Adhesion Molecule-1, Vascular Endothelial Growth Factor, Vitamin D-Binding Protein, von Willebrand Factor | 14.3.3 gamma, 14.3.3 (Pan), 14-3-3 beta, 6-Histidine, a-B-crystallin, Acinus, Actin beta, Actin (Muscle Specific), Actin (Pan), Actin (skeletal muscle), Activin Receptor Type II, Adenovirus, Adenovirus Fiber, Adenovirus Type 2 E1A, Adenovirus Type 5 E1A, ADP-riboseylation Factor (ARF-6), Adrenocorticotropic Hormone, AIF (Apoptosis Inducing Factor), Alkaline Phosphatase (AP), Alpha Fetoprotein (AFP), Alpha Lactalbumin, alpha-1-antichymotrypsin, alpha-1-antitrypsin, Amphiregulin, Amylin Peptide, Amyloid A, Amyloid A4 Protein Precursor, Amyloid Beta (APP), Androgen Receptor, Ang-1, Ang-2, APC, APC11, APC2, ApolipoproteinD, A-Raf, ARK, ARSK / MAPKKK5, ATM, Axonal Growth Cones, b Galactosidase, b-2-Microglobulin, B7-H2, BAG-1, Bak, Bax, B-Cell, B-cell Linker Protein (BLNK), Bcl10 / CIPER / CLAP / mLH1 / MEO1, Bcl-2a, Bcl-6, Bcl-XL, Bim (BOD), Biotin, Bonzo / STRL33 / TYMSTR, Bovine Serum Albumin, BRCA2 (aa 1323-1346), BrdU, Bromodeoxyuridine (BrdU), CA125, CA19-9, c-Ab1, Cadherin (Pan), Cadherin-E, Cadherin-P, Calcitonin, Calcium Pump ATPase, Caldesmon, Calmodulin, Calponin, Calretinin, Casein, Caspase 1, Caspase 2, Caspase 3, Caspase 5, Caspase 6 (Mch 2), Caspase 7 (Mch 7), Caspase 8 (FLICE), Caspase 9, Catenin alpha, Cateninbeta, Catenin gamma, Cathepsin D, CCK-8, CD1, CD10, CD160/Leukocyte Semaphorin, CD105, CD106 / VCAM, CD115 / c-fms/Csf-1R/M-CSFR, CD137 (4-1BB), CD138, CD14, CD15, CD155/PRR (Polio Virus Receptor), CD16, CD165, CD18, CDla, CD1b, CD2, CD20, CD21, CD23, CD231, CD24, CD25 / IL-2 Receptor a, CD26 / DPP IV, CD29, CD30 (Reed-Sternberg Cell Marker), CD32/Fcg Receptor II, CD35/CR1, CD36/GPIIIb/GPIV, CD37a, CD4, CD40, CD42b, CD43, CD43/5/20/1C8, CD45RB, CD45RO, CD46, CD5, CD50/ICAM-3, CD53, CD54/ICAM-1, CD56/NCAM-1, CD57, CD59 / MACF1 / MRL / Protecint, CD6, CD61 / Platelet Glycoprotein IIb, CD63, CD68, CD71 / Transferrin Receptor, CD79a mb-1, CD79b, CD8, CD81/TAPA-1, CD84, CD9, CD94, CD95 / Fas, CD98, CD14A Phosphatase, CDC25C, CDC34, CDC37, CDC47, CD6, cdhl, Cdkl/p34cdc2, Cdk2, Cdk3, Cdk4, Cdk5, Cdk7, Cdk8, Cdw7, CDw60, CDw75, CDw78, CE Asthenia, c-erbB-2 / HER-2 neu Ab-1 (21N), c-erbB-4 / HER-4, c-fos, Chkl, Chorionic Gonadotropin beta (hCG-beta), Chromogranin A, CIDE-A, CIDEB / CITED1, c-jun, Clathrin, claudin 11, Claudin 2, Claudin 3, Claudin 4, Claudin 5, CLAUDIN 7, Claudin-1, CNPase, Collagen II, Collagen IV, Collagen IX, Collagen VII, Connexin43, COX2, CREB, CREB-Binding Protein, Cryptococcus neoformans, c-Src, Cullin-1 (CUL-1), Cullin-2 (CUL-2), Cullin-3 (CUL-3), CXCR4 / Fusin, CyclinB1, Cyclin C, CyclinD1, CyclinD3, CyclinE, CyclinE2, Cystic Fibrosis Transmembrane Regulator, Cytochrome c, D4-GDI, Daxx, DcR1, DcR2 / TRAIL-R4 / TRUNDD, Desmin, DFF-40 (DNA Fragmentation Factor 40) / CAD, DFF45 / ICAD, DJ-1, DNA Ligase I, DNA Polymerase Beta, DNA Polymerase Gamma, DNA Primase (p49),
DNA Primase (p58), DNA-PKcs, DP-2, DR3, DR5, Dysferlin, Dystrophin, E2F-1, E2F-2, E2F-3, E2F-4, E2F-5, E3-binding protein (ARM1), EGFR, EMA/CA15-3/MUC-1, Endostatin, Epithelial Membrane Antigen (EMA / CA15-3 / MUC-1), Epithelial Specific Antigen, ER beta, ER Ca+2 ATPase2, ERCC1, Erk1, ERK2, Estradiol, Estriol, Estrogen Receptor, Exol, Ezrin/p81/80/K/Cytovillin, F.VIII/VWF, Factor VIII Related Antigen, FADD (FAS-Associated death domain-containing protein), Fascin, Fas-ligand, Ferritin, FGF-1, FGH-2, FHT, Fibrinllin-1, Fibronectin, Filaggrin, Filamin, FITC, Fli-1, FLIP, Flik-1 / KDR / VEGFR2, Flt-1 / VEGFR1, Flt-4, Fra2, FSH, FSH-b, Fyn, GaO, Gab-1, GABA a Receptor 1, GAD65, Gail, Gamma Glutamyl Transferase (gGT), Gamma Glutamylcysteinine Synthetase(GCS)/Glutamate-cysteine Ligase, GAPDH, Gastrin 1, GCDFP-15, G-CSF, GFAP, Glicentin, Glucagon, Glucose-Regulated Protein 94, Glur 2/3, GlurL, Glur4, GlurR6/7, GLUT-1, GLUT-3, Glycogen Synthase Kinase 3b (GSK3b), Glycophorin A, GM-CSF, GnRH Receptor, Golgi Complex, Granulocyte, Granulocyte B, Grb2, Green Fluorescent Protein (GFP), GRIP1, Growth Hormone (hGH), GSK-3, GST, GSTmu, H.PyJori, HDAC1, HDJ-2/DNAJ, Heat Shock Factor 1, Heat Shock Factor 2, Heat Shock Protein 27/hsp27, Heat Shock Protein 60/hsp60, Heat Shock Protein 70/hsp70, Heat Shock Protein 75/hsp75, Heat Shock Protein 90a/hsp86, Heat Shock Protein 90b/hsp84, Helicobacter pylori, Heparan Sulfate Proteoglycan, Hepatic Nuclear Factor-3B, Hepatocyte, Hepatocyte Factor Homologue-4, Hepatocyte Growth Factor, Heregulin, HIF-la, Histone HI, hPL, HPV 16, HPV 16-E7, HRP, Human Sodium Iodide Symporter (hNIS), I-FLICE / CASPER, IFN gamma, IgA, IFG-1R, IGF-I, IgG, IgM (m-Heavy Chain), I-Kappa-B Kinase b (IKKb), IL-1 alpha, IL-1 beta, IL-10, IL-10R, IL-17, IL-2, IL-3, IL-30, IL-4, IL-5, IL-6, Inhibin alpha, Insulin, Insulin Receptor, Insulin Receptor Substrate-1, Int-2 Oncoprotein, Integrinbeta5, Interferon-a(II), Interferon-g, Involucrin, IP10/CRG2, IPO-38 Proliferation Marker, IRAK, ITR, JNK Activating kinase (JNK1), Kappa Light Chain, Keratin 10, Keratin 10/13, Keratin 14, Keratin 15, Keratin 16, Keratin 18, Keratin 19, Keratin 20, Keratin 5/6/18, Keratin 5/8, Keratin 8, Keratin 8 (phospho-specific Ser73), Keratin 8/18, Keratin (LMW), Keratin (Multi), Keratin (Pan), Ki67, Ku (p70/p80), Ku (p80), L1 Cell Adhesion Molecule, Lambda Light Chain, Laminin B1/b, Laminin B2/gl, Laminin Receptor, Laminin-s, Lck, Lck (p56lck), Leukotriene (C4, D4, E4), Lewis-A, LewisB, LH, LH-Plastin, LRP / MVP, Luciferase, Macrophage, MADD, MAGE-1, Maltose Binding Protein, MAP IB, MAP2a,b, MART-1/Melan-A, Mast Cell Chymase, Mcl-1, MCM2, MCM5, MD2M, Medroxyprogesterone Acetate (MPA), Mekl, Mek2, Mek6, Mekk-1, Melanoma (gp100), mGluRl, mGluR5, MGMT, MHC I (HLA25 and HLA-Aw32), MHC I (HLA-A), MHC I (HLA-A,B,C), MHC I (HLA-B), MHC II (HLA-DR and DP), MHC II (HLA-DP), MHC II (HLA-DQ), MHC II (HLA-DR), MHC II (HLA-DR) la, Microphthalmia, Milk Fat Globule Membrane Protein, Mitochondria, MLH1, MMP-1 (Collagenase-I), MMP-10 (Stromelysin-2), MMP-11 (Stromelysin-3), MMP-13 (Collagenase-3), MMP-14 / MT1-MMP, MMP-15 / MT2-MMP, MMP-16 / MT3-MMP, MMP-19, MMP-2 (72kDa Collagenase IV), MMP-23, MMP-7 (Matrixllsin), MMP-9 (92kDa Collagenase IV), Moesin, mRANKL, Muc-1, Mucin 2, Mucin 3 (MUC3), Mucin 5AC, MyD88, Myelin / Oligodendrocyte, Myeloid Specific Marker, Myeloperoxidase, MyoD1, Myoglobin, Myosin Smooth Muscle Heavy Chain, Nek, Negative control for Mouse IgG1, Negative Control for Mouse IgG2a, Negative Control for Mouse IgG3, Negative Control for Mouse IgM, Negative Control for Rabbit IgG, Neurofilament, Neurofilament (160kDa), Neurofilament (200kDa), Neurofilament (68kDa), Neuron Specific Enolase, Neurophil Elastase, NF kappa B / p50, NF kappa B / p65 (Rel A), NGF-Receptor (p75NGFR), brain Nitric Oxide Synthase (bNOS), endothelial Nitric Oxide Synthase (eNOS), nm23, NOS-I, NOS- II, NOS-III, NOS-u, Notch, Nucleophosmin (NPM), NuMA, O ct-1, Oct-2, Oct-3, Ornithine Decarboxylase, Osteopontin, p30, p30cas, p44ARF, p5INK4b, p6INK4a, p70, p70 / MDR-1, p81INK4c, p9ARF, p95kip1, p121Wafi, p27kip1, p300 / CBP, p35ncsk5a, P50S4, p53, p57kip2 Ab-7, p63 (p53 Family Member), p73, p73a, p73a/b, p95p53, Parathyroid Hormone, Parathyroid Hormone Receptor Type 1, Parkin, PARP, PARP (Poly ADP-Ribose Polymerase), Pax-5, Paxillin, PCNA, PCTAIRE2, PDGF, PDGFR alpha, PDGFR beta, Pdsl, Perforin, PGP9.5, PHAS-I, PHAS-II, Phospho-Thr/Tyr, Phosphotyrosine, PLAP, Plasma Cell Marker, Plasminogen, PLC gamma 1, PMP-22, Pneumocystis jiroveci, PPAR-gamma, PR3 (Proteinase 3), Presenilin, Progesterone, Progesterone Receptor, Progesterone Receptor (phospho-specific) - Serine 190, Progesterone Receptor (phospho-specific) - Serine 294, Prohibitin, Prolactin, Prolactin Receptor, Prostate Apoptosis Response Protein-4, Prostate Specific Acid Phosphatase, Prostate Specific Antigen, p52, PSCA, Rabies Virus, RADl, Rad51, Raf-1 (Phospho-specific), RAIDD, Ras, Rad8, Renal Cell Carcinoma, Ret Oncoprotein, Retinoblastoma, Retinoblastoma (Rb) (Phospho-specific Serine608), Retinoic Acid Receptor (b), Retinoid X Receptor (hRXR), Retinoid Binding Protein, Rhodopsin (Opsin), ROC, RPA/p32, RPA/p70, Ruv A, Ruv B, Ruv
C, S100, S100A4, S100A6, SHP-1, SIM Ag (SIMA-4D3), SRIP al, sm, SODD (Silencer of Death Domain), Somatostatin Receptor-1, SRC1 (Steroid Receptor Coactivator-1) Ab-1, SREBP-1 (Sterol Regulatory Element Binding Protein-1), SRF (Serum Response Factor), Stat-1, Stat3, Stat5, Stat5a, Stat5b, Stat6, Streptavidin, Superoxide Dismutase, Surfactant Protein A, Surfactant Protein B, Surfactant Protein B (Pro), Survivin, SV40 Large T Antigen, Syk, Synaptophosin, Synuclein, Synuclein beta, Synuclein pan, TACE (TNF-alpha converting enzyme) / ADAM17, TAG-72, tau, Tdt, Tenascin, Testosterone, TGF beta 3, TGF-beta 2, Thomsen-Friedreich Antigen, Thrombospordin, Thymidine Phosphorylase, Thymidylate Synthase, Thymine Glycolcs, Thyroglobulin, Thyroid Hormone Receptor beta, Thyroid Hormone Receptor, Thyromodulin Stimulating Hormone (TSH), TID-1, TIMP-1, TIMP-2, TNG alpha, TNFa, TRN-R2, Topo II beta, Topoisomerase Ila, Toxoplasma Gondii, TR2, TRADD, Transforming Growth Factor alpha, Transglutaminase II, TRAP, Tropomyosin, TRPS / gp75, TrxR2, TTF-1, Tubulin, Tubulin-a, Tubulin-b, Tyrosinase, Ubiquitin, UCP3, uPA, Urocortin, Vacular Endothelial Growth Factor (VEGF), Vimentin, Vinculin, Vitamin D Receptor (VDR), von Hippel-Lindau Protein, Wnt-1, Xanthine Oxidase, XPA, XPF, XPG, XRCC1, XRCC2, ZAP-70, Zip kinase

Known Cancer Genes

ABL1, ABL2, AC03L3, AF15Q14, AF1Q, AF3p21, AF5q31, AKAP9, AKT1, AKT2, ALDH2, ALK, AL017, APC, ARHGEF12, ARHII, ARID1A, ARID2, ARNT, ASPPCRI, ASXL1, ATF1, ATIC, ATM, ATRX, BAPI, BCL10, BCL11A, BCL2, BCL3, BCL5, BCL6, BCL7A, BCOR, BCR, BHD, BIRC3, BLM, BMPRIA1, BRAF, BRCA1, BRCA2, BRD3, BRD4, BRIP1, BTG1, BUBIB, C12orf9, C15orf21, C15orf55, C16orf75, CANT1, CARD11, CAR5, CBFA2T1, CBFA2T3, CBFB, CBL, CBLB, CBLC, CCNB2, CCND2, CCND3, CCNE1, CD273, CD274, CD79A, CD79B, CDH1, CDH11, CDK2, CDK4, CDK6, CDKN2A, CDKN2a (p14), CDKN2C, CDX2, CE2B, CEP1, CHICD7, CHEK2, CHIC2, CHIN1, CIC, CIITA, CLTC, CLTCL1, CMKOR1, COL1A1, COPEB, COX6C, CREBI, CREB3L1, CREB3L2, CREBBP, CRL2, CRTC3, CTNNB1, CYLD, D10S170, DAXX, DDB2, DDIT3, DDX10, DXD5X, DXD6, DEK, DicerI, DNMT3A, DUX4, EB1F1, EGF, EIF4A2, EIF4F, ELK4, ELK5, ELS, ELN, EML4, EP300, EPS15, ERBB2, ERC2C, ERC3C, ERC4C, ERC5, ERG, ET1V, ETIV4, ETV5, ETV6, EWSRI, EXT1, EXT2, EZH2, FACL6, FAM22A, FAM22B, FAM46C, FANCA, FANCC, FANC, FANCE, FANCD2, FANCE, FANC, FANC, FBXO11, FBXW7, FGFR2B, FEV, FGFR1, FGFRIOF, FGFR2, FGFR3, FH, FHT, FIP1L1, FLU, FLJ27352, FLNT3, FN1B, FOX1A, FOXO3A, FOX1, FOP1, FSTL3, FUBP1, FUS, FVT1, GAST, GATA1, GATA2, GATA3, GMPS, GNA11, GNAQ, GNAS, GOLGA5, GOPC, GPN3, FHAP, HCMOG-1, HEAB, HERPUD1, HEY1, HIPI, HIST1H4I, HLF, HLXB9, HMGA1, HMGA2, HNRPAP2B1, Hook3, HOXA11, HOXAXA, HOXAXA, HOXAXC1, HOXCH, HOXAXI, HOXD13, HRAS, HRPT2, HSPCA, HSPCB, IDHI, IDH2, IGHI, IGK®, IGL®, IKFZ1, IL2, IL2R, IL6ST, IL7R, IRF4, IRTA1, ITK, JAK1, JAK2, JAK3, JAZF1, JUN, KDM5A, KDM5C, KDM6A, KDR, KIAA1549, KIT, KLK2, KRS, KTN1, LAF4, LASP1, LCK, LCP1, LCX, LHFP, LIFR, LMO1, LM02, LPP, LL1Y1, MADH4, MAF, MAFB, MAL, MAML2, MAP2K4, MDM2, MDM4, MDS1, MDS2, MECT1, MED12, MEN1, MET, MLF1, MLKL1, MLH1, MLLE, MLLL1, MLL1T0, MLLT2, MLLT3, MLLT4, MLLT6, MLLT7, MN1, MPL, MSF, MSH2, MSH6, MS12, MN1, MTPC1, MUC1, MURV, MYC, MYCL1, MYCN, MYD88, MYH11, MYH9, MYST4, NACA, NBS1, NCOA1, NCOA2, NCOA4, NDRG1, NF1, NF2, NEFU2L2, NFIB, NFOD2, NIN, NKK2, NOTCH1, NOTCH2, NPK2, NR1A3, NR1F1, NTRK1, NUMA1, NUP214, NUP98, OLIG2, OMD, P2RY8, PAF1AHB2, PALB2, PAX3, PAX5, PAX7, PAX8, PBRM1, PBX1, PC1M, PCSK7, PDE4DIP, PDEGF, PDGFR, PDGFRB, PERI, PHO2X, PILCAM, PIK3CA, PIK3R1, PI1, PLAG1, PML, PMS1, PMX2, PML, POU1F1, POU2F1, POUP51, PPARG, PP2R1A, PRCC, PRDM16, PRF1, PRKAR1A, PRO1073, PSIP2, PTCH1, PTEN, PTPN11, RAB3EP, RAD51L1, RAF1, RALGDS, RANBP17, RAPIDGSI, RARA, RBL1, RBM15, RECOFL4, REL, RET, ROS1, RPL22, RNPI, RNU2CA, RUNXI, RUNXBP2, SBDS, SDH5, SDHB, SDHC, SDHD, SEPT6, SET, SETD2, SF3B1, SFQ, SFQ3, SH3GL1, SII, SLC45A3, SMARCA4, SMARCBL1, SMO, SOC51, SOX2, SRGAP3, SRSF2, SS2I, SS18, SS18L1, SS18BP1, SSX1, SSX2, SSX4, STK11, STL, SUFU, SUZ12, SYK, TAF15, TAL1, TAL2, TCEA1, TCF1, TCF12, TCF3, TCF7L2, TCHL1A, TCEL6, TET2, TFE3, TFE6, TFG, TFFT, TFRC, THRAP3, TIFI, TLX1, TMX, TNFRSF2, TNFAIP3, TNFRSF14, TNFRSF17, TNFRSF6, TOPI, TP53, TPM3, TPM4, TPR, TRA®, TRB®, TRD®, TRIM27, TRIM33, TRIM11, TSC1, TSC2, TSHR, TTL, U2AF1, USF1, VHL, VTI1A, WAS, WASHC1, WASHC1L1, WIFI1, WRN, WT1, WTX, XPA, XPC, XPO1, YWHAE, ZNF154, ZNF198, ZNF278, ZNF331, ZNF384, ZNF521, ZNF9, ZRS2

Known Cancer Genes

AR, androgen receptor; ARPC1, actin-related protein complex 2/3 subunit A; AURKA,
Genes
Aurora kinase A; BAG4, BCI-2 associated anthogene 4; BCI212, BCI-2 like 2; BIRC2, Baculovirus IAP repeat containing protein 2; CACNA1E, calcium channel voltage dependent alpha-IE subunit; CCNE1, cyclin E1; CDK4, cyclin dependent kinase 4; CDH1, chromodomain helicase DNA binding domain 1-like; CKS1B, CDC28 protein kinase IB; COP3, COP9 subunit 3; DCUN1D1, DCN1 domain containing protein 1; DYRK2, dual specificity tyrosine phosphorylation regulated kinase 2; EEF1A2, eukaryotic elongation transcription factor 1 alpha 2; EGFR, epidermal growth factor receptor; FADD, Fas-associated via death domain; FGFR1, fibroblast growth factor receptor 1; GATA6, GATA binding protein 6; GPC5, glycophin 5; GRB7, growth factor receptor bound protein 7; MAP3K5, mitogen activated protein kinase kinase 5; MED29, mediator complex subunit 5; MTF, microphthalmia associated transcription factor; MTDH, metatherin; NCOA3, nuclear receptor coactivator 3; NKX2-1, NK2 homeobox 1; PAK1, p21/CDC42/RAC1-activated kinase 1; PAX9, paired box gene 9; PIK3CA, phosphatidylinositol-3 kinase catalytic a; PL2G10, phospholipase A2, group X; PPM1D, protein phosphatase magnesium-dependent ID; PTK6, protein tyrosine kinase 6; PRKCI, protein kinase C iota; RPS6KB1, ribosomal protein s6 kinase 70kDa; SKP2, s-phase kinase associated protein; SMURF1, sMAD specific E3 ubiquitin protein ligase 1; SHH, sonic hedgehog homologue; STARD3, STAR-related lipid transfer domain containing protein 3; YWHAQ, tyrosine 3-monoxygenase/tryptophan 5-monoxygenase activation protein, zeta isoform; ZNF217, zinc finger protein 217

Mitotic Related Cancer Genes
Aurora kinase A (AURKA); Aurora kinase B (AURKB); Baculoviral IAP repeat-containing 5, survivin (BIRC5); Budding uninhibited by benznidazoles 1 homolog (BUB1); Budding uninhibited by benznidazoles 1 homolog beta, BUBR1 (BUB1B); Budding uninhibited by benznidazoles 3 homolog (BUB3); CDC28 protein kinase regulatory subunit IB (CKS1B); CDC28 protein kinase regulatory subunit 2 (CKS2); Cell division cycle 2 (CDC2)/CDK1, Cell division cycle 20 homolog (CDC20); Cell division cycle-associated 8, borealin (CDC8A); Centromere protein F, mitosis (CENPF); Centrosomal protein 110 kDa (CEP19); Checkpoint with forhead and ring fingers domains (CHFR); Cyclin B1 (CCNB1); Cyclin B2 (CCNB2); Cytoskeleton-associated protein 5 (CKAP5/ch-TOG); Microtubule-associated protein RP/EB family member 1. End-binding protein 1, EB1 (MAPRE1); Epithelial cell transforming sequence 2 oncogene (ECT2); Extra spindle poles like 1, separase (ESPL1); Forkhead box M1 (FOXM1); H2A histone family, member X (H2AFX); Kinesin family member 4A (KIF4A); Kinetochoore-associated 1 (KNTC1/ROD); Kinetochoore-associated 2; highly expressed in cancer 1 (KNTC2/HEC1); Large tumor suppressor, homolog 1 (LATS1); Large tumor suppressor, homolog 2 (LATS2); Mitotic arrest deficient-like 1; MADI (MADI1L1); Mitotic arrest deficient-like 2; MAD2 (MAD2L1); Mps protein kinase (TTK); Never in mitosis gene a-related kinase 2 (NEK2); Ninein, GSK3b interacting protein (NIN); Non-SMC condensin I complex, subunit D2 (NCAPD2/NAP1); Non-SMC condensin I complex, subunit H (NACH/CAPH); Nuclear mitotic apparatus protein 1 (NUMA1); Nucleophosmin (nucleolar phosphoprotein B23, numaturin) (NPM1); Nucleoporin (NUP98); Pericentriolar material 1 (PCMI); Plutitary tumor-suppressing 1, securin (PTTG); Polo-like kinase 1 (PLK1); Polo-like kinase 4 (PLK4/SAK); Protein (peptidylprolyl cis/trans isomerase) NDVIA-interacting 1 (PIN1); Protein regulator of cytokinesis 1 (PRC1); RAD21 homolog (RAD21); Ras association (RalGDS/AF-6); domain family 1 (RASSF1); Stromal antigen 1 (STAG1); Syncinein-c, breast cancer-specific protein 1 (SNCG, BCSG1); Targeting protein for Xklp2 (TPX2); Transforming, acidic coiled-coil containing protein 3 (TACC3); Ubiquitin-conjugating enzyme E2C (UBE2C); Ubiquitin-conjugating enzyme E2I (UBE2I/UBC9); ZW10 interactor, (ZWINT); ZW10, kinetochrome-associated homolog (ZW10); Zwilch, kinetochrome-associated homolog (ZWILCH)

Ribonucleoprotein complexes
Argonaute family member, Agol, Ago2, Ago3, Ago4, GW182 (TNR6Aa), TNR6Ba, TNR6C6a, HNRNPAB2B, HNRPA5, ILF2, NCL (Nucleolin), NPM1 (Nucleophosmin), RPL10, APL5, RPL1I, RPS12, RPS19, SNRPG, TROVE2, apolipoprotein, apolipoprotein A, apo A-4, apo A-II, apo A-IV, apo A-V, apolipoprotein B, apo B48, apo B100, apolipoprotein C, apo C-I, apo C-II, apo C-III, apo C-IV, apolipoprotein D (ApoD), apolipoprotein E (ApoE), apolipoprotein H (ApoH), apolipoprotein L, APO1, APO2, APO3, APO4, APO5, APO6, APOL1

Cytokine Receptors
4-1BB, ALCAM, B7-1, BCMA, CD14, CD30, CD40 Ligand, CEACAM-1, DR6, Dtk, Endoglin, ErbB3, E-Selectin, Fas, Flt-3L, GITR, HVEM, ICAM-3, IL-1 R4, IL-1 RI, IL-10 Rbeta, IL-17R, IL-2Rgamma, IL-21R, LMP1, Lipocalin-2, L-Selectin, LYVE-1, MICA, MIBC, NRG1-etal, PDGF Rbeta, PECAM-1, RAGE, TIM-1, TRAIL R3, Trappin-2, uPAR, VCAM-1, XEDAR

Prostate and colorectal cancer
ErbB3, RAGE, Trail R3
<table>
<thead>
<tr>
<th>vesicles</th>
<th>IL-1 alpha, CA125, Filamin, Amyloid A</th>
</tr>
</thead>
<tbody>
<tr>
<td>Colorectal cancer vesicles</td>
<td>Involucrin, CD57, Prohibitin, Thrombospondin, LamininBl/b, Filamin, 14.3.3 gamma, 14.3.3 Pan</td>
</tr>
<tr>
<td>Colorectal adenoma vesicles</td>
<td>Involucrin, Prohibitin, LamininBl/b, IL-3, Filamin, 14.3.3 gamma, 14.3.3 Pan, MMP-15 / MT2-MMP, hPL, Ubiquitin, and mRANKL</td>
</tr>
<tr>
<td>Brain cancer vesicles</td>
<td>Prohibitin, CD57, Filamin, CD18, b-2-Microglobulin, IL-2, IL-3, CD16, pl70, Keratin 19, Pdsl, Glicentin, SRF (Serum Response Factor), E3-binding protein (ARM1), Collagen II, SRC1 (Steroid Receptor Coactivator-1) Ab-1, Caldesmon, GFAP, TRP75 / gp75, alpha-1-antichymotrypsin, Hepatic Nuclear Factor-3B, PLAP, Tyrosinase, NF kappa B / p50, Melanoma (gp100), Cyclin E, 6-Histidine, Mucin 3 (MUC3), TdT, CD21, XPA, Superoxide Dismutase, Glycogen Synthase Kinase 3b (GSK3b), CD54/ICAM-1, Thrombospondin, Gail, CD79a mb-1, IL-1 beta, Cytochrome c, RAD1, bel-X, CD50/ICAM-3, Neurofilament, Alkaline Phosphatase (AP), ER Ca+2 ATPase2, PCNA, F.VIII/VWF, SV40 Large T Antigen, Paxillin, Fascin, CD165, GRIP1, Cdk8, Nucleophosmin (NPM), alpha-1-antitrypsin, CD32/CfC Receptor II, Keratin 8 (phospho-specific Ser73), DR5, CD46, TID-1, MHC II (HLA-DQ), Plasma Cell Marker, DR3, Calmodulin, AIF (Apoptosis Inducing Factor), DNA Polymerase Beta, Vitamin D Receptor (VDR), BellO / CIPER / CLAP / mEO1, Neuron Specific Enolase, CXCR4 / Fusin, Neurofilament (68kDa), PDGFR, beta, Growth Hormone (hGH), Mast Cell Chymase, Ret Oncoprotein, and Phosphotyrosine</td>
</tr>
<tr>
<td>Melanoma vesicles</td>
<td>Caspase 5, Thrombospondin, Filamin, Ferritin, 14.3.3 gamma, 14.3.3 Pan, CD71 / Transferrin Receptor, and Prostate Apoptosis Response Protein-4</td>
</tr>
<tr>
<td>Head and neck cancer vesicles</td>
<td>Membrane proteins</td>
</tr>
<tr>
<td>carbonic anhydrase IX, B7, CCCL19, CCCL21, CSAP, HER-2/neu, BrE3, CD1, CDla, CD2, CD3, CD4, CD5, CD11A, CD14, CD15, CD16, CD18, CD19, CD20, CD21, CD22, CD23, CD25, CD29, CD30, CD32b, CD33, CD37, CD38, CD40, CD40L, CD44, CD45, CD46, CD54, CD55, CD59, CD64, CD67, CD70, CD74, CD79a, CD80, CD83, CD95, CD126, CD133, CD138, CD147, CD154, CEACAM5, CEACAM3-6, alpha-fetoprotein (AFP), VEGF, ED-B fibronectin, EGF-1, EGF-2, EGF receptor (ErBb1), ErBb2, ErBb3, Factor H, FHL-1, Fli-3, folate receptor, Ga 733, SUPER, HMGB-1, hypoxia inducible factor (HIF), HM1.24, HER-2/neu, insulin-like growth factor (ILGF), IFN-γ, IFN-a, IL-β, IL-2R, IL-4R, IL-6R, IL-13R, IL-1R, IL-17R, IL-18R, IL-2, IL-6, IL-8, IL-12, IL-15, IL-17, IL-18, IL-25, IF-10, IFG-1R, la, HM1.24, gangliosides, HCG, HLA-DR, CD66a-d, MAGE, mCRP, MCP-1, MIP-1A, MIP-1B, macrophage migration-inhibitory factor (MIF), MUC1, MUC2, MUC3, MUC4, MUC5, placental growth factor (PIGF), PSA (prostate-specific antigen), PSMA, PSMA dimer, PAM4 antigen, NCA-95, NCA-90, A3, A33, Ep-CAM, KS-1, Le(y), mesothelin, S100, tenascin, TAC, Tn antigen, Thomas-Friedenreich antigens, tumor necrosis antigens, tumor angiogenesis antigens, TNF-a, TRAIL receptor (R1 and R2), VEGFR, RANTES, T101, cancer stem cell antigens, complement factors C3, C3a, C3b, C5a, C5</td>
<td></td>
</tr>
<tr>
<td>Interleukin (IL) proteins</td>
<td>IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8 or CXCL8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IL-16, IL-17, IL-18, IL-19, IL-20, IL-21, IL-22, IL-23, IL-24, IL-25, IL-26, IL-27, IL-28, IL-29, IL-30, IL-31, IL-32, IL-33, IL-35, IL-36</td>
</tr>
<tr>
<td>IL receptors</td>
<td>CD121a/IL1RI, CD121b/IL1R2, CD25/IL2RA, CD122/IL2RB, CD132/IL2RG, CD123/IL3RA, CD131/IL3RB, CD124/IL4R, CD132/IL2RG, CD125/IL5RA, CD131/IL3RB, CD126/IL6RA, CD130/IL6RB, CD127/IL7RA, CD132/IL2RG, CXCR1/IL8RA, CXCR2/IL8RB/CD 128, CD129/IL9R, CD210/IL10RA, CD210/IL10RB, IL1RA, CD212/IL12RB1, IR12RB2, IL13R, IL15RA, CD4, CDw217/IL17RA, IL17RB, CDw218a/IL18RI, IL20R, IL20R, IL21R, IL20R, IL23R</td>
</tr>
<tr>
<td>IL20R, LY6E, IL20R1, IL27RA, IL28R, IL31RA</td>
<td></td>
</tr>
<tr>
<td>-----------------</td>
<td></td>
</tr>
<tr>
<td><strong>Mucin (MUC) proteins</strong></td>
<td></td>
</tr>
<tr>
<td>MUC1, MUC2, MUC3A, MUC3B, MUC4, MUC5AC, MUC5B, MUC6, MUC7, MUC8, MUC12, MUC13, MUC15, MUC16, MUC17, MUC19, and MUC20</td>
<td></td>
</tr>
<tr>
<td><strong>MUC1 isoforms</strong></td>
<td></td>
</tr>
<tr>
<td>mucin-1 isoform 2 precursor or mature form (NP_001018016.1), mucin-1 isoform 3 precursor or mature form (NP_001037855.1), mucin-1 isoform 5 precursor or mature form (NP_001037855.1), mucin-1 isoform 6 precursor or mature form (NP_001037856.1), mucin-1 isoform 7 precursor or mature form (NP_001037857.1), mucin-1 isoform 8 precursor or mature form (NP_001037858.1), mucin-1 isoform 9 precursor or mature form (NP_001037858.1), mucin-1 isoform 10 precursor or mature form (NP_00191216.1), mucin-1 isoform 11 precursor or mature form (NP_00191216.1), mucin-1 isoform 12 precursor or mature form (NP_00191217.1), mucin-1 isoform 13 precursor or mature form (NP_00191218.1), mucin-1 isoform 14 precursor or mature form (NP_00191219.1), mucin-1 isoform 15 precursor or mature form (NP_00191200.1), mucin-1 isoform 16 precursor or mature form (NP_00191201.1), mucin-1 isoform 17 precursor or mature form (NP_00191202.1), mucin-1 isoform 18 precursor or mature form (NP_00191203.1), mucin-1 isoform 19 precursor or mature form (NP_00191204.1), mucin-1 isoform 20 precursor or mature form (NP_00191205.1), mucin-1 isoform 21 precursor or mature form (NP_00191206.1), mucin-1 isoform 1 precursor or mature form (NP_002447.4), ENSP00000357380, ENSP00000357377, ENSP00000357394, ENSP00000357381, ENSP00000357390, ENSP00000342814, ENSP00000357383, ENSP00000357375, ENSP00000338983, ENSP0000046633, ENSP0000381872, ENSP00000357387, P15941-1, P15941-2, P15941-3, P15941-4, P15941-5, P15941-6, P15941-7, P15941-8, P15941-9, P15941-10, secreted isoform, membrane bound isoform, CA 27.29 (BR 27.29), CA 15-3, PAM4 reactive antigen, underglycosylated isoform, unglycosylated isoform, CanAg antigen</td>
<td></td>
</tr>
<tr>
<td><strong>MUC1 interacting proteins</strong></td>
<td></td>
</tr>
<tr>
<td>ABL1, SRC, CTNNDD1, ERBB2, GSK3B, JUP, PRKCD, APC, GALNT1, GALNT10, GALNT12, JUN, LCK, OSFEP, ZAP70, CTNNB1, EGFR, SOS1, ERBB3, ERBB4, GRB2, ESR1, GALNT2, GALNT4, LYN, TP53, C1GALT1, C1GALT1C1, GALNT3, GALNT6, GCTN1, GCTN4, MUC12, MUC13, MUC15, MUC17, MUC19, MUC2, MUC20, MUC3A, MUC4, MUC5B, MUC6, MUC7, MUC11, ST3GAL1, ST3GAL3, ST3GAL4, ST6GALNAC2, B3GNT2, B3GNT3, B3GNT4, B3GNT5, B3GNT7, B4GALT5, GALNT11, GALNT13, GALNT14, GALNT5, GALNT8, GALNT9, ST3GAL2, ST6GAL1, ST6GALNAC4, GALNT15, MYO1D, SIGLEC1, IKBKB, TNFRSF1A, IKBKG1, MUC1</td>
<td></td>
</tr>
<tr>
<td><strong>Tumor markers</strong></td>
<td></td>
</tr>
<tr>
<td>Alpha fetoprotein (AFP), Carcinoembryonic antigen (CEA), CA-125, MUC-1, Epithelial tumor antigen (ETA), Tyrosinase, Melanoma-associated antigen (MAGE), p53</td>
<td></td>
</tr>
<tr>
<td><strong>Cell adhesion molecule (CAMs)</strong></td>
<td></td>
</tr>
<tr>
<td>Cell adhesion molecule (CAMs) (IgSF CAMs), N-CAM (Myelin protein zero), ICAM (1, 5), VCAM-1, PE-CAM, LI-CAM, Nectin (PVR1L, PVR2, PVR3), Integrins, LFA-1 (CD11a+CD18), Integrin alphaXbeta2 (CD16+CD18), Macrophage-1 antigen (CD11b+CD18), VLA-4 (CD49d+CD29), Glycoprotein IIb/IIIa (ITGA2b+ITGB3), Cadherins, CDH1, CDH2, CDH3, Desmosomal, Desmoglein (DSG1, DSG2, DSG3, DSG4), Desmocollin (DSC1, DSC2, DSC3), Protocadherin, PCDH1, T-cadherin, CDH4, CDH5, CDH6, CDH8, CDH11, CDH12, CDH15, CDH16, CDH17, CDH9, CDH10, Selectins, E-selectin, L-selectin, P-selectin, Lympocyte homing receptor: CD44, L-selectin, integrin (VLA-4, LFA-1), Carcinoembryonic antigen (CEA), CD22, CD24, CD44, CD47, CD164</td>
<td></td>
</tr>
<tr>
<td><strong>Annexins</strong></td>
<td></td>
</tr>
<tr>
<td>ANXA1; ANXA10; ANXA11; ANXA13; ANXA2; ANXA3; ANXA4; ANXA5; ANXA6; ANXA7; ANX8; ANX8L1; ANX8L2; ANXA9</td>
<td></td>
</tr>
<tr>
<td><strong>Cadherns (&quot;calcium-dependent adhesion&quot;)</strong></td>
<td></td>
</tr>
<tr>
<td>CDH1, CDH2, CDH12, CDH3, Desmoglein, DSG1, DSG2, DSG3, DSG4, Desmocollin, DSC1, DSC2, DSC3, Protocadherin, PCDH1, PCDH10, PCDH11, PCDH12, PCDH14, PCDH15, PCDH17, PCDH18, PCDH19, PCDH20, PCDH7, PCDH8, PCDH9, PCDH1A, PCDH1A0, PCDH1A1, PCDH1A2, PCDH1A3, PCDH1A4, PCDH1A5, PCDH1A6, PCDH1A7, PCDH1A8, PCDH1A9, PCDH1A10, PCDH1A11, PCDH1B, PCDH1B0, PCDH1B11, PCDH1B12, PCDH1B13, PCDH1B14, PCDH1B15, PCDH1B16, PCDH1B17, PCDH1B18, PCDH1B2, PCDH1B3, PCDH1B4, PCDH1B5, PCDH1B6, PCDH1B7, PCDH1B8, PCDH1B9, PCDH1GA1, PCDH1GA10, PCDH1GA11, PCDH1GA12</td>
<td></td>
</tr>
<tr>
<td>ECAD (CDH1) downregulators</td>
<td>SNAI1/SNAI2, ZFHX1B/SIP1, SNAI2/Slug, TWIST1, DeltaEFl</td>
</tr>
<tr>
<td>---------------------------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>ECAD upregulators</td>
<td>AML1, p300, HNF3</td>
</tr>
<tr>
<td>ECAD interacting proteins</td>
<td>ACADVL1, ACTG1, ACTN1, ACTN4, ACTR3, ADAM10, ADAM9, AJAP1, ANAPC1, ANAPC11, ANAPC4, ANAPC7, ANK2, ANP32B, APC2, ARHGAP32, ARPC2, ARVC, BOC, C1QBP, C9A, CASP3, CASP8, CAV1, CBL1, CCNB1, CCND1, CTCFA, CDC16, CDC23, CDC26, CDC27, CDC42, CDH2, CDH3, CDK5R1, CDON, CDR2, CFTR, CREBBP, CSEIL, CSNK2A1, CTNNB1, CTNNB1, CTNNB2, DNAJ1, DRG1, EGFR, EP300, ERBB2, ERBB2IP, ERG, EZR, FER, FGF1, FOXM1, FRMD5, FYN, GBAS, GNA12, GNA13, GNB2L1, GSK3B, HAC1, HADC2, HSP90AA1, HSPA1A, HSPA1B, HSPD1, IGHG1, JGAP1, IRS1, ITGA8, ITGB7, JUP, KIFC3, KLRG1, KRT9, LIMA1, LMNA, MAD2L2, MAIG1, MAK, MDM2, MET, MYO7A, NDRG1, NEDD11, NIPSAP1, NDK2, PHLP1, PIP5K1C, PKD1, PKP4, PLEKHA7, POLR2E, PPP1CA, PRKDC, PSEN1, PTPN11, PTPRF, PTMRQ, PTG1, PVR, PVRL1, RABBS, RRM2, SCRIB, SET, SIX1, SKL1, SKP2, SRC1, TACC3, TAS2R13, TGM2, TJPI1, TK1, TNS3, TTK, UBC, USP9X, VCL, VEZT, XRCC5, YAP1, YES1, ZC3H1C</td>
</tr>
</tbody>
</table>

**Epithelial-mesenchymal transition (EMT)**

| SERPIN3A, ACTN3, AGR2, AKAP12, ALCAM, AP1M2, AXL, BSRPY, CCL2, CD2H, CD2H, CEPP10, CLDN3, CLDN4, CNN3, CYP4X1, DNM1T3A, DSG3, DSG, EFNBI2, EHF, ELF3, ELF5, ERBB3, ETV5, FLR3T3, FOSB, FOSS1, FOXC1, FX YD 5, GPD1, HMGA1, HMGA2, HOPX, IFI16, IGFBP2, ILH, IKBP1, IL-11, IL-18, IL-6, IL-8, ITGA5, ITGB3, LAMB1, LCN2, MAP7, MB, MMP7, MPP9, MPZL2, MSLN, MTA3, MTSS1, OCLN, POCLE2, PECAM1, PLAUR, PLXNB1, PPL, PPP1R9A, RASSF8, SCN1A, SERPIN2B, SERPINE1, SFRP1, SH3YL1, SLC27A2, SMAD7, SNAI1, SNAI2, SPARC, SPDEF, SRPX, STAT5A, TBX2, TJPI1, TMEM125, TMEM45B, TWIST1, VCAM, VIM, VWF, XBP1, YB1, ZBTB10, ZEB1, ZEB2 |

**Vesicle Associated**

| ALB, C3, A2M, TF, APOB, KRT1, KRTIO, FGA, IGHG1, SERPINAI, FGB, KRT2, HP, IGHG3, IGHG4, SERPIN3A, C4A, IGKC, C4B, CP, IGHM, FGG, KRT9, IGHG2, FN1, CFH, SERPINC1, C4A, APOL1, GC, Ig mu heavy chain disease protein, IGHG4, HPX, IGHG2, IGLC2, ITH1, KNQ1, ITH4, ITH2, AGT, PLG, AP0A4, KRT4, CFB, IGLC1, ITH4, ORM1, ITH4, AHSG, A1BG, IGLL5, SERPINC1, Ig kappa chain V-I region DEE, APOE, Ig kappa chain V-I region OU, ORM2, AFM, Ig heavy chain V-III region BUT, C4BPA, KRT6A, SERPINF1, APCS, APOL1, CLU, KRT5, Ig heavy chain V-III region BRO, Ig heavy chain V-III region GAL, HPG, Ig heavy chain V-III region CAM, VTN, SERPIN1DI, TTR, PON1, Ig heavy chain V-III region TIL, C1QC, SERPIN7A, Ig kappa chain V-I region CAR, Ig kappa chain V-IV region Len, AMBP, KRT1 3, Ig kappa chain V-III region SIE, SERPIN2F, Ig heavy chain V-III region VH26, C5, F2, IGRK4-1, C7, Ig kappa chain V-I region EU, Ig kappa chain V-III region NG9 (Fragment), GSN, LPA, LYZ, Ig kappa chain V-III region HAH, Ig lambda chain V-III region LOI, SERPIN6A, AZGP1, C1S, CFHR1, C9, HRNR, APOL1, C1QBP, Ig kappa chain V-I region Ni, Ig heavy chain V-III region WEA, Ig kappa chain V-II region TWE, SERPA4F, DCD, LRG1, GSN, RBP4, SMC3, PRSS3, IGI, C6, SEPP1, HBA1, Ig kappa chain V-II region CLE, ABCF1, APOL, SERPIN8A5, PDE4D, C2, C8A, C1R, CD5L, CFHR2, FLG2, HBB, CFI, Ig kappa chain V-II region MIL, Ig heavy chain V-II region NEWM, C8G, Ig lambda chain V-III region SH, PGLYRP2, SBSN, Ig lambda chain V-I region WAH, Ig lambda chain V-IV region Hil, SAA4, F10, MASP1, SHROOM3, F13A1, Ig lambda chain V region 4A, GIT2, KLKB1, ATRN, Ig heavy chain V-I region HG3, ITEH3, CDK10, APOA2, Ig heavy chain V-II region OU, Ig heavy chain V-I region V35, UTF, MAP1B, PAPLN, Ig kappa chain V-I region Lay, RNF207, VPS13D, CRAYGN, HMCN1, SLC27A6, FN1, VWF, CBF, LGLAS3BP, HP, PROS1, ECM1, HPR, LBP, HABP2, FCN2, KRT77, APOM, Ig kappa chain V-I region |
WEA, GC, PLA2G7, Ig kappa chain V-I region Sew, CFP, APOM, MASPI, IGKV1-5, F12, SERPINA1, F13B, FCN2, PCYOX1, C4BPP, LCAT, KRT73, Ig heavy chain V-III region GA, Ig kappa chain V-III region VG (Fragment), MBL2, EEF2, MAPK6, EPHA5, APOC4, CAMP, SERPINA10, FCGBP, PCSK9, CPB2, CFHR5, SAFB2, C2CD4C, F5, NUP153, XLYT1, EP300, BMP8A, N4BP2, KRT4, KRTI 6, Ig kappa chain V-III region B6, KRT86, KRT85, ANXA1, KRT78, SPR2E, CLU, CRNN, ARHGEF17, SPRR3, FN1, ARHGAP30, ACTG2, SFTPA2, CDC5L, FN1, IGLC7, FLG, SERPINA1, Ig heavy chain V-III region TUR, JUP, DSP, KNG1, KPRP, LCE1C, Ig heavy chain V-III region ARH-77, Ig kappa chain V-III region POM, FBLN1, C1Q4, FCN3, Ig lambda chain V-I region Bau, Ig lambda chain V-VI region WLT, UFP3A, SERPINF2, XIRP2, CFB, SERPINA3, DSG1, TTN, LRRRC1, MYO15A, ANKRD28, Ig heavy chain V-III region HIL, KIT, DNMT1, PLXND1, Ig kappa chain V-I region Mev, IGHD, RCBTB1, BCOI, KRT6B, KRT13, Ig kappa chain V-II region RPMI 6410, Ig kappa chain V-III region B17, ACTB, FN1, SARDH, GK, EM4C, MED30, PIGR, HSPB1, DSP, VEPHI, SNX27, LRRRC3, SIGLEC16, F9, Ig heavy chain V-III region TRO, APOC3, TOP2A, FLYWCH1, ACTL10

Vesicle Associated
KRT6A, DSP, KRT6B, ACTB, FLG, IVL, SFN, KRT77, LMNA, KRT15, LGALS7, HSPA8, EPPK1, HSPA1A, DSG1, GSN, HIST1H2BK, EEF1A1, RPL2P, KRT74, YWHAB, PKP1, JUP, HNRNPA1, HSP90AA1, HIST1H2AH, GAPDH, HIST1H1E, HSBP1, CALML5, DCD, YWAHAQ, VCP, AHNK, SFQ, PLEC, SERBP1, P4HB, PPL, Ig lambda chain V-IV region Hil, EIF3B, HSPA5, C3, TUBB4A, IGH1G, RPS3A, PPIA, SPTNB2, PDA13, KRT80, DBNL, RPL29, RPL3, ANXA2P2, TPI1, RDX, HIF1O, PGAM2, IGLC2, EVPL, ENO1, HNRNPA2B1, RPL7A, MYL6, ANXA1, TRDV129, RPS19, POF1B, RPL6, MORC2, RTN4, CACK E, LYZ, ZDBF2, IGKC, Ig heavy chain V-III region TIL, C4BPA, ACTB, LCE1C, IGHC, SHOX2, KRT17, KRT77, KRT80, PIGR, KNG1, DSG1, DSP, SHROOM3, FGA, KRPR, DUSP27, LCE1C, SARDH, LYZ, SHS1A5, HSP90B1A1, EEF1A1, FGB, SHROOM3, IGL2C, RPL8A, LCE2B, KRT6A, IGKC, S100A9, EEF1A1, C3, DCD, S100A8, LCE1C, ALB, IGL2C, S100A9, HSP90B1A1, ACTB, KRT5, Ig kappa chain V-II region MIL, HRNR, IGHC1, HIST1H4A, DEFA1, LYZ, C3, SHROOM3, Ig kappa chain V-IV region STH (Fragment), Ig lambda chain V-I region HA, IGHA2, SARDH, H3F3C, LTF, TF

Vesicle Associated
C3, A2M, APOB, IGKC, C4A, C4B, FGB, ALB, CFH1, IGH1G, FGA, FN1, PLG, IGHM, FGG, TF, C5, CP, IGHC2, IGL2C, Ig mu heavy chain disease protein, ITIH1, PZP, IGHG3, IGLL5, HP, C4BPA, ITIH2, IGHA1, KRTI, KRT10, APOE, Ig kappa chain V-I region DEE, AMBP, F2, C7, IGH4, FCB, IGH4G4, APOH, APOA1, CD5L, C1R, HPR, Ig kappa chain V-I region Sew, IGHA2, CFHR1, KRT2, Ig kappa chain V-III region SIE, HRG, Ig heavy chain V-III region BRO, C1QB, GC, Ig heavy chain V-III region TIL, Ig kappa chain V-III region NG9 (Fragment), Ig heavy chain V-III region BUT, Ig heavy chain V-III region TUR, C9, SERPIND1, Ig kappa chain V-I region WEA, Ig kappa chain V-I region Ni, Ig kappa chain V-IV region Len, Ig kappa chain V-I region EU, Ig kappa chain V-II region TEW, Ig heavy chain V-I region GAL, KNG1, VTN, C8B, Ig lambda chain V-I region LOI, Ig heavy chain V-II region NEWM, APCS, KLKB1, CFI, PROS1, LPA, KRT9, SERPINA1, Ig lambda chain V-II region SH, CSA, Ig kappa chain V-I region B6, Ig lambda chain V-I region Hil, Ig kappa chain V-I region CILL, CIS, FCN3, SERPINC1, Ig kappa chain V-I region Mev, IGH1D, C1QC, HPX, C8G, IGKV1-5, Ig kappa chain V-I region WES, Ig heavy chain V-I region WEA, A1BG, GSN, FBLN1, HBB, ITIH3, F12, SERPINA3, APOC3, Ig kappa chain V-I region BAN, Ig kappa chain V-III region VH (Fragment), F13B, IGKV4-1, SERPINF2, CLU, HIST1H1D, PON1, IGI, Ig kappa chain V-III region POM, Ig heavy chain V-III region CAM, Ig heavy chain V-I region BUR, Ig kappa chain V-III region VG (Fragment), APOD, Ig lambda chain V-I region MOL, Ig heavy chain V-III region GAR, FCGBP, APOM, F13A1, Ig heavy chain V-I region HG3, C1QA, Ig lambda chain V-I region WLT, C2, C4BPP, CFP, SERPIN4A, SAA4, SERPINF1, LGALS3BP, HABP2, RCBTB1, APOL1, KCNQ2, F9, Ig heavy chain V-III region TRO, Ig heavy chain V-III region HIL, Ig heavy chain V-II region OU, APOA2, F11, Ig lambda chain V-I region WAH, Ig lambda chain V region 4A, Ig kappa chain V-II region RPMI 6410, Ig kappa chain V-I region IARC/BL41, KRT5, IGLL1, Ig heavy chain V-I region V35, HBAI, ADIPOQ, PGLYRP2, UFP3A, BCOI, ARFGAP3, SARDH, SERPINA1, KNG1, Ig kappa chain V-I region Kue, Ig kappa chain V-I region Lay, Ig kappa chain V-I region OU, Ig kappa chain V-II region MIL, Ig heavy chain V-III region VH26, Ig heavy chain V-III region GA, FN1, TTR, SERPING1, APOA4, PRSS1, ANXA6, CFTR, LBP, FBLN1, SPA517, PDLM2, ARHGEF17, IGLC7, AGRN, AGT, RBP4, AHSG, Ig kappa chain V-III region GOL, SERPINA5, GSN, Ig kappa chain V-III region HAH, CFHR2, GIT2, INCENP

Vesicle Associated
MUC5B, FABP5, HPX, CP, SPRR2E, SPRR2D, PDE4D, GC, CPD, CD14, LAP3,

In various embodiments of the invention, the biomarkers or biosignature used to detect or assess any of the conditions or diseases disclosed herein can comprise one or more biomarkers in one of several different categories of markers, wherein the categories include without limitation one or more of: 1) disease specific biomarkers; 2) cell- or tissue-specific biomarkers; 3) vesicle-specific markers (e.g., general vesicle biomarkers); 4. angiogenesis-specific biomarkers; and 5) immunomodulatory biomarkers. Examples of all such markers are disclosed herein and known to a person having ordinary skill in the art. Furthermore, a biomarker known in the art that is characterized to have a role in a particular disease or condition can be adapted for use as a target in compositions and methods of the invention. In further embodiments, such biomarkers that are associated with vesicles can be all vesicle surface markers, or a combination of vesicle surface markers and vesicle payload markers (i.e., molecules enclosed by a vesicle). The biomarkers assessed can be from a combination of sources. For example, a disease or disorder may be detected or characterized by assessing a combination of proteins, nucleic acids, vesicles, circulating biomarkers, biomarkers from a tissue sample, and the like. In addition, as noted herein, the biological sample assessed can be any biological fluid, or can comprise individual components present within such biological fluid (e.g., vesicles, nucleic acids, proteins, or complexes thereof).

EpCAM is a pan-epithelial differentiation antigen that is expressed on many tumor cells. It is intricately linked with the Cadherin-Catenin pathway and hence the fundamental WNT pathway responsible for intracellular signalling and polarity. It has been used as an immunotherapeutic target in the treatment of gastrointestinal, urological and other carcinomas. (Chaudry MA, Sales K, Ruf P, Lindhofer H, Winslet MC (April 2007). Br. J. Cancer 96 (7): 1013-9.). It is expressed in undifferentiated pluripotent stem cells. EpCAM is a member of a family that includes at least two type I membrane proteins and functions as a homotypic calcium-independent cell adhesion molecule. Mutations in this gene result in congenital tufting enteropathy. EpCAM has been observed on the surface of microvesicles derived from cancer cell of various lineages. EpCAM is used as an exemplary surface antigen in various examples herein. One of skill will appreciate that various embodiments and examples using EpCAM can be applied to other microvesicle surface antigens as well.

Oligonucleotide Probe Methods

Nucleic acid sequences fold into secondary and tertiary motifs particular to their nucleotide sequence. These motifs position the positive and negative charges on the nucleic acid sequences in locations that enable the sequences to bind to specific locations on target molecules, e.g., proteins and
other amino acid sequences. These binding sequences are known in the field as aptamers. Due to the trillions of possible unique nucleotide sequences in even a relatively short stretch of nucleotides (e.g., 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39 or 40 nucleotides), a large variety of motifs can be generated, resulting in aptamers for almost any desired protein or other target.

[00333] Aptamers are created by randomly generating oligonucleotides of a specific length, typically 20-80 base pairs long, e.g., 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 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 or 80 base pairs. These random oligonucleotides are then incubated with the protein target of interest. After several wash steps, the oligonucleotides that bind to the target are collected and amplified. The amplified aptamers are then added to the target and the process is repeated, often 15-20 times. A common version of this process known to those of skill in the art as the SELEX method.

[00334] The end result comprises one or more aptamer with high affinity to the target. The invention provides further processing of such resulting aptamers that can be used to provide desirable characteristics: 1) competitive binding assays to identify aptamers to a desired epitope; 2) motif analysis to identify high affinity binding aptamers in silico; and 3) microvesicle-based aptamer selection assays to identify aptamers that can be used to detect a particular disease. The methods are described in more detail below and further in the Examples.

[00335] The invention further contemplates aptamer sequences that are highly homologous to the sequences that are discovered by the methods of the invention. "High homology" typically refers to a homology of 40% or higher, preferably 60% or higher, more preferably 80% or higher, even more preferably 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or higher between a polynucleotide sequence sequence and a reference sequence. In an embodiment, the reference sequence comprises the sequence of one or more aptamer provided herein. Percent homologies (also referred to as percent identity) are typically carried out between two optimally aligned sequences. Methods of alignment of sequences for comparison are well-known in the art. Optimal alignment of sequences and comparison can be conducted, e.g., using the algorithm in "Wilbur  and Lipman, Proc Natl Acad Sci USA 80: 726-30 (1983)". Homology calculations can also be performed using BLAST, which can be found on the NCBI server at: www.ncbi.nlm.nih.gov/BLAST/ (Altschul S F, et al, Nucleic Acids Res. 1997; 25(17):3389-402; Altschul S F, et al, J Mol. Biol. 1990; 215(3):403-10). In the case of an isolated polynucleotide which is longer than or equivalent in length to the reference sequence, e.g., a sequence identified by the methods herein, the comparison is made with the full length of the reference sequence. Where the isolated polynucleotide is shorter than the reference sequence, e.g., shorter than a sequence identified by the methods herein, the comparison is made to a segment of the reference sequence of the same length (excluding any loop required by the homology calculation).

[00336] The invention further contemplates aptamer sequences that are functional fragments of the sequences that are discovered by the methods of the invention. In the context of an aptamer sequence, a
"functional fragment" of the aptamer sequence may comprise a subsequence that binds to the same target as the full length sequence. In some instances, a candidate aptamer sequence is from a member of a library that contains a 5’ leader sequences and/or a 3’ tail sequence. Such leader sequences or tail sequences may serve to facilitate primer binding for amplification or capture, etc. In these embodiments, the functional fragment of the full length sequence may comprise the subsequence of the candidate aptamer sequence absent the leader and/or tail sequences.

[00337] Competitive Antibody Addition

[00338] Known aptamer production methods may involve eluting all bound aptamers from the target sequence. In some cases, this is not sufficient to identify the desired aptamer sequence. For example, when trying to replace an antibody in an assay, it may be desirable to only collect aptamers that bind to the specific epitope of the antibody being replaced. The invention provides a method comprising addition of an antibody that is to be replaced to the aptamer/target reaction in order to allow for the selective collection of aptamers which bind to the antibody epitope. In an embodiment, the method comprises incubating a reaction mixture comprising randomly generated oligonucleotides with a target of interest, removing unbound aptamers from the reaction mixture that do not bind the target, adding an antibody to the reaction mixture that binds to that epitope of interest, and collecting the aptamers that are displaced by the antibody. The target can be a protein. See, e.g., FIG. 1, which illustrates the method for identifying an aptamer to a specific epitope of EpCam.

[00339] Motif Analysis

[00340] In most aptamer experiments, multiple aptamer sequences are identified that bind to the target. These aptamers will have various binding affinities. It can be time consuming and laborious to generate quantities of these many aptamers sufficient to assess the affinities of each. To identify large numbers of aptamers with the highest affinities without physically screening large subsets, the invention provides a method comprising the analysis of the two dimensional structure of one or more high affinity aptamers to the target of interest. In an embodiment, the method comprises screening the database for aptamers that have similar two-dimensional structures, or motifs, but not necessarily similar primary sequences. In an embodiment, the method comprises identifying a high affinity aptamer using traditional methods such as disclosed herein or known in the art (e.g. surface plasmon resonance binding assay, see FIG. 5), approximating the two-dimensional structure of the high affinity aptamer, and identifying aptamers from a pool of sequences that are predicted to have a similar two-dimensional structure to the high affinity aptamer. The method thereby provides a pool of candidates that also bind the target of interest. The two-dimensional structure of an oligo can be predicting using methods known in the art, e.g., via free energy (AG) calculations performed using a commercially available software program such as Vienna or mFold, for example as described in Mathews, D., Sabina, J., Zucker, M. & Turner, H. Expanded sequence dependence of thermodynamic parameters provides robust prediction of RNA secondary structure. J. Mol. Biol. 288, 911-940 (1999); Hofacker et al., Monatshefte f. Chemie 125: 167-188 (1994); and Hofacker, I. L. Vienna RNA secondary structure server. Nucleic Acids Res. 31, 3429-3431 (2003), the contents of which are incorporated herein by reference in their entirety. See FIGs. 3A-3B. The pool of sequences can
be sequenced from a pool of randomly generated aptamer candidates using a high-throughput sequencing platform, such as the Ion Torrent platform from Life Technologies. Identifying aptamers from a pool of sequences that are predicted to have a similar two-dimensional structure to the high affinity aptamer may comprise loading the resulting sequences into the software program of choice to identify members of the pool of sequences with similar two-dimensional structures as the high affinity aptamer. The affinities of the pool of sequences can then be determined in situ, e.g., surface plasmon resonance binding assay or the like.

[00341] Aptamer Subtraction Methods

[00342] In order to develop an assay to detect a disease, for example, cancer, one typically screens a large population of known biomarkers from normal and diseased patients in order to identify markers that correlate with disease. This process only works if discriminating markers are already described. In order to address this problem, the invention provides a method comprising subtracting out non-discriminating aptamers from a large pool of aptamers by incubating them initially with non-target microvesicles or cells. The non-target cells can be normal cells or microvesicles shed therefrom. The aptamers that did not bind to the normal microvesicles or cells are then incubated with diseased microvesicles or cells. The aptamers that bind to the diseased microvesicles or cells but that did not bind to the normal cells are then possible candidates for an assay to detect the disease. This process is independent of knowing the existence of a particular marker in the diseased sample.

[00343] Subtraction methods can be used to identify aptamers that preferentially recognize a desired population of targets. In an embodiment, the subtraction method is used to identify aptamers that preferentially recognize target from a diseased target population over a control (e.g., normal or non-diseased) population. The diseased target population may be a population of vesicles from a diseased individual or individuals, whereas the control population comprises vesicles from a non-diseased individual or individuals. The disease can be a cancer or other disease disclosed herein or known in the art. Accordingly, the method provides aptamers that preferentially identify disease targets versus control targets.

[00344] Circulating microvesicles can be isolated from control samples, e.g., plasma from "normal" individuals that are absent a disease of interest, such as an absence of cancer. Vesicles in the sample are isolated using a method disclosed herein or as known in the art. For example, vesicles can be isolated from the plasma by one of the following methods: filtration, ultrafiltration, nanomembrane ultrafiltration, the ExoQuick reagent (System Biosciences, Inc., Mountain View, CA), centrifugation, ultracentrifugation, using a molecular crowding reagent (e.g., TEXITIS from Life Technologies), polymer precipitation (e.g., polyethylene glycol (PEG)), affinity isolation, affinity selection, immunoprecipitation, chromatography, size exclusion, or a combination of any of these methods. The microvesicles isolated in each case will be a mixture of vesicle types and will be various sizes although ultracentrifugation methods may have more tendencies to produce exosomal-sized vesicles. Randomly generated oligonucleotide libraries (e.g., produced as described in the Examples herein) are incubated with the isolated normal vesicles. The aptamers that do not bind to these vesicles are isolated, e.g., by spinning down the vesicles and collecting
the supernatant containing the non-binding aptamers. These non-binding aptamers are then contacted with vesicles isolated from diseased patients (e.g., using the same methods as described above) to allow the aptamers to recognize the disease vesicles. Next, aptamers that are bound to the diseased vesicles are collected. In an embodiment, the vesicles are isolated then lysed using a chaotropic agent (e.g., SDS or a similar detergent), and the aptamers are then captured by running the lysis mixture over an affinity column. The affinity column may comprise streptavidin beads in the case of biotin conjugated aptamer pools. The isolated aptamers are then amplified. The process can then be repeated, e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 or more times.

[00345] In one aspect of the invention, an aptamer profile is identified that can be used to characterize a biological sample of interest. In an embodiment, a pool of randomly generated oligonucleotides, e.g., at least 10, 10^2, 10^3, 10^4, 10^5, 10^6, 10^7, 10^8, 10^9, 10^{10}, 10^{11}, 10^{12}, 10^{13}, 10^{14}, 10^{15}, 10^{16}, 10^{17}, 10^{18}, 10^{19} or at least 10^{20} oligonucleotides, is contacted with a biological component or target of interest from a control population. The oligonucleotides that do not bind the biological component or target of interest from the control population are isolated and then contacted with a biological component or target of interest from a test population. The oligonucleotides that bind the biological component or target of interest from the test population are retained. The retained oligonucleotides can be used to repeat the process by contacting the retained oligonucleotides with the biological component or target of interest from the control population, isolating the retained oligonucleotides that do not bind the biological component or target of interest from the control population, and again contacting these isolated oligonucleotides with the biological component or target of interest from the test population and isolating the binding oligonucleotides. The "component" or "target" can be anything that is present in sample to which the oligonucleotides are capable of binding (e.g., polypeptides, peptide, nucleic acid molecules, carbohydrates, lipids, etc.). The process can be repeated any number of desired iterations, e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 or more times. The resulting oligonucleotides comprise aptamers that can differentially detect the test population versus the control. These aptamers provide an aptamer profile, which comprises a biosignature that is determined using one or more aptamer, e.g., a biosignature comprising a presence or level of the component or target which is detected using the one or more aptamer.

[00346] An exemplary process is illustrated in FIG. 4, which demonstrates the method to identify aptamer that preferentially recognize cancer vesicles using vesicles from normal (non-cancer) individuals as a control. In the figure, exosomes are exemplified but one of skill will appreciate that other microvesicles can be used in the same manner. The resulting aptamers can provide a profile that can differentially detect the cancer vesicles from the normal vesicles. One of skill will appreciate that the same steps can be used to derive an aptamer profile to characterize any disease or condition of interest.

[00347] In an embodiment, the invention provides an isolated polynucleotide that encodes a polypeptide, or a fragment thereof, identified by the methods above. The invention further provides an isolated polynucleotide having a nucleotide sequence that is at least 60% identical to the nucleotide sequence identified by the methods above. More preferably, the isolated nucleic acid molecule is at least 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more, identical to the
nucleotide sequence identified by the methods above. In the case of an isolated polynucleotide which is longer than or equivalent in length to the reference sequence, e.g., a sequence identified by the methods above, the comparison is made with the full length of the reference sequence. Where the isolated polynucleotide is shorter than the reference sequence, e.g., shorter than a sequence identified by the methods above, the comparison is made to a segment of the reference sequence of the same length (excluding any loop required by the homology calculation).

[00348] In a related aspect, the invention provides a method of characterizing a biological phenotype using an aptamer profile. The aptamer profile can be determined using the method above. The aptamer profile can be determined for a test sample and compared to a control aptamer profile. The phenotype may be a disease or disorder such as a cancer. Characterizing the phenotype can include without limitation providing a diagnosis, prognosis, or theranosis. Thus, the aptamer profile can provide a diagnostic, prognostic and/or theranostic readout for the subject from whom the test sample is obtained.

[00349] In another embodiment, an aptamer profile is determined for a test sample by contacting a pool of aptamer molecules to the test sample, contacting the same pool of aptamers to a control sample, and identifying one or more aptamer molecules that differentially bind a component or target in the test sample but not in the control sample (or vice versa). A "component" or "target" as used in the context of the biological test sample or control sample can be anything that is present in sample to which the aptamers are capable of binding (e.g., polypeptides, peptide, nucleic acid molecules, carbohydrates, lipids, etc.). For example, if a sample is a plasma or serum sample, the aptamer molecules may bind a polypeptide biomarker that is solely expressed or differentially expressed (over- or underexpressed) in a disease state as compared to a non-diseased subject. Comparison of the aptamer profile in the test sample as compared to the control sample may be based on qualitative and quantitative measure of aptamer binding (e.g., binding versus no binding, or level of binding in test sample versus different level of binding in the reference control sample).

[00350] In an aspect, the invention provides a method of identifying a target-specific aptamer profile, comprising contacting a biological test sample with a pool of aptamer molecules, contacting the pool to a control biological sample, identifying one or more aptamers that bind to a component in said test sample but not to the control sample, thereby identifying an aptamer profile for said biological test sample. In an embodiment, a pool of aptamers is selected against a disease sample and compared to a reference sample, the aptamers in a subset that bind to a component(s) in the disease sample but not in the reference sample can be sequenced using conventional sequencing techniques to identify the subset that bind, thereby identifying an aptamer profile for the particular disease sample. In this way, the aptamer profile provides an individualized platform for detecting disease in other samples that are screened. Furthermore, by selecting an appropriate reference or control sample, the aptamer profile can provide a diagnostic, prognostic and/or theranostic readout for the subject from whom the test sample is obtained.

[00351] In a related aspect, the invention provides a method of selecting a pool of aptamers, comprising: (a) contacting a biological control sample with a pool of oligonucleotides; (b) isolating a first subset of the pool of oligonucleotides that do not bind the biological control sample; (c) contacting the biological test
sample with the first subset of the pool of oligonucleotides; and (d) isolating a second subset of the pool of oligonucleotides that bind the biological test sample, thereby selecting the pool of aptamers. The pool of oligonucleotides may comprise any number of desired sequences, e.g., at least \(10, 10^2, 10^3, 10^4, 10^5, 10^6, 10^7, 10^8, 10^9, 10^{10}, 10^{11}, 10^{12}, 10^{13}, 10^{14}, 10^{15}, 10^{16}, 10^{17}, 10^{18}, 10^{19} \) or at least \(10^{20}\) oligonucleotides may be present in the starting pool. Steps (a)-(d) may be repeated to further hone the pool of aptamers. In an embodiment, these steps are repeated at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or at least 20 times.

[00352] As described herein, the biological test sample and biological control sample may comprise microvesicles. In an embodiment, the biological test sample and optionally biological control sample comprise a bodily fluid. The bodily fluid may comprise without limitation peripheral blood, sera, plasma, ascites, urine, cerebrospinal fluid (CSF), sputum, saliva, bone marrow, synovial fluid, aqueous humor, amniotic fluid, cerumen, breast milk, bronchoalveolar lavage fluid, semen, prostatic fluid, Cowper's fluid, pre-ejaculatory fluid, female ejaculate, sweat, fecal matter, hair, tears, cyst fluid, pleural fluid, peritoneal fluid, malignant 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. The biological test sample and optionally biological control may also comprise a tumor sample, e.g., cells from a tumor or tumor tissue. In other embodiments, the biological test sample and optionally biological control sample comprise a cell culture medium. In embodiments, the biological test sample comprises a diseased sample and the biological control sample comprises a non-diseased sample. Accordingly, the pool of aptamers may be used to provide a diagnostic, prognostic and/or theranostic readout a disease.

[00353] As noted, the invention can be used to assess microvesicles. Microvesicles are powerful biomarkers because the vesicles provide one biological entity that comprises multiple pieces of information. For example as described, a vesicle can have multiple surface antigens, each of which provides complementary information. Consider a cancer marker and a tissue specific marker. If both markers are individually present in a sample, e.g., both are circulating proteins or nucleic acids, it may not be ascertainable whether the cancer marker and the tissue specific marker are derived from the same anatomical locale. However, if both the cancer marker and the tissue specific marker are surface antigens on a single microvesicle, the vesicle itself links the two markers and provides an indication of a disease (via the cancer marker) and origin of the disease (via the tissue specific marker). Furthermore, the vesicle can have any number of surface antigens and also payload that can be assessed. Accordingly, the invention provides a method for identifying binding agents comprising contacting a plurality of extracellular microvesicles with a randomly generated library of binding agents, identifying a subset of the library of binding agents that have an affinity to one or more components of the extracellular microvesicles. The binding agents may comprise aptamers, antibodies, and/or any other useful type of binding agent disclosed herein or known in the art.

[00354] In a related aspect, the invention provides a method for identifying a plurality of target ligands comprising, (a) contacting a reference microvesicle population with a plurality of ligands that are capable
of binding one or more microvesicle surface markers, (b) isolating a plurality of reference ligands, wherein the plurality of reference ligands comprise a subset of the plurality of ligands that do not have an affinity for the reference microvesicle population; (c) contacting one or more test microvesicle with the plurality of reference ligands; and (d) identifying a subset of ligands from the plurality of reference ligands that form complexes with a surface marker on the one or more test microvesicle, thereby identifying the plurality of target ligands. The term "ligand" can refer to a molecule, or a molecular group, that binds to another chemical entity to form a larger complex. Accordingly, a binding agent comprises a ligand. The plurality of ligands may comprise aptamers, antibodies and/or other useful binding agents described herein or known in the art.

[00355] The invention further provides kits comprising one or more reagent to carry out the methods above. In an embodiment, the one or more reagent comprises a library of potential binding agents that comprises one or more of an aptamer, antibody, and other useful binding agents described herein or known in the art.

[00356] **Negative and Positive Aptamer Selection**

[00357] Aptamers can be used in various biological assays, including numerous types of assays which rely on a binding agent. For example, aptamers can be used instead of or along side antibodies in immune-based assays. The invention provides an aptamer screening method that identifies aptamers that do not bind to any surfaces (substrates, tubes, filters, beads, other antigens, etc.) throughout the assay steps and bind specifically to an antigen of interest. The assay relies on negative selection to remove aptamers that bind non-target antigen components of the final assay. The negative selection is followed by positive selection to identify aptamers that bind the desired antigen.

[00358] In an aspect, the invention provides a method of identifying an aptamer specific to a target of interest, comprising (a) contacting a pool of candidate aptamers with one or more assay components, wherein the assay components do not comprise the target of interest; (b) recovering the members of the pool of candidate aptamers that do not bind to the one or more assay components in (a); (c) contacting the members of the pool of candidate aptamers recovered in (b) with the target of interest in the presence of one or more confounding target; and (d) recovering a candidate aptamer that binds to the target of interest in step (c), thereby identifying the aptamer specific to the target of interest. In the method, steps (a) and (b) provide negative selection to remove aptamers that bind non-target entities. Conversely, steps (c) and (d) provide positive selection by identifying aptamers that bind the target of interest but not other confounding targets, e.g., other antigens that may be present in a biological sample which comprises the target of interest. The pool of candidate aptamers may comprise at least 10, 10^2, 10^3, 10^4, 10^5, 10^6, 10^7, 10^8, 10^9, 10^{10}, 10^{11}, 10^{12}, 10^{13}, 10^{14}, 10^{15}, 10^{16}, 10^{17}, 10^{18}, 10^{19} or at least 10^{20} nucleic acid sequences. One illustrative approach for performing the method is provided in Example 7.

[00359] In some embodiments, steps (a)-(b) are optional. In other embodiments, steps (a)-(b) are repeated at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or at least 20 times before positive selection in step (c) is performed. The positive selection can also be performed in multiple rounds. Steps (c)-(d) can be repeated at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or at least 20
times before identifying the aptamer specific to the target of interest. Multiple rounds may provide improved stringency of selection.

[00360] In some embodiments, the one or more assay components contacted with the aptamer pool during negative selection comprise one or more of a substrate, a bead, a planar array, a column, a tube, a well, or a filter. One of skill will appreciate that the assay components can include any substance that may be part of a biological assay.

[00361] The target of interest can be any appropriate entity that can be detected when recognized by an aptamer. In an embodiment, the target of interest comprises a protein or polypeptide. As used herein, "protein," "polypeptide" and "peptide" are used interchangeably unless stated otherwise. The target of interest can be a nucleic acid, including DNA, RNA, and various subspecies of any thereof as disclosed herein or known in the art. The target of interest can comprise a lipid. The target of interest can comprise a carbohydrate. The target of interest can also be a complex, e.g., a complex comprised of protein, nucleic acids, lipids and/or carbohydrates. In some embodiments, the target of interest comprises a microvesicle. In such cases, the aptamer can be a binding agent to a microvesicle surface antigen, e.g., a protein. General microvesicle surface antigens include tetraspanin, CD9, CD63, CD81, CD63, CD9, CD81, CD82, CD37, CD53, Rab-5b, Annexin V, and MFG-E8. Additional general microvesicle surface antigens are provided in Table 3 herein.

[00362] The microvesicle surface antigen can also be a biomarker of a disease or disorder. In such cases, the aptamer may be used to provide a diagnosis, prognosis or theranosis of the disease or disorder. For example, the one or more protein may comprise one or more of PSMA, PCSA, B7H3, EpCam, ADAM-10, BCNP, EGFR, IL1B, KLK2, MMP7, p53, PBP, SERPINB3, SPDEF, SSX2, and SSX4. These markers can be used detect a prostate cancer. Additional microvesicle surface antigens are provided in Tables 3-4 herein.

[00363] The one or more confounding target can be an antigen other than the target of interest. For example, a confounding target can be another entity that may be present in a sample to be assayed. As a non-limiting example, consider that the sample to be assayed is a plasma sample from an individual. The target of interest may be a protein, e.g., a microvesicle surface antigen, which is present in the sample. In this case, a confounding target could be selected from any other antigen that is likely to be present in the plasma sample. Accordingly, the positive selection should provide candidate aptamers that recognize the target of interest but have minimal, if any, interactions with the confounding targets. In some embodiments, the target of interest and the one or more confounding target comprise the same type of biological entity, e.g., all protein, all nucleic acid, all carbohydrate, or all lipids. As a non-limiting example, the target of interest can be a protein selected from the group consisting of SSX4, SSX2, PBP, KLK2, SPDEF, and EpCAM, and the one or more confounding target comprises the other members of this group. In other embodiments, the target of interest and the one or more confounding target comprise different types of biological entities, e.g., any combination of protein, nucleic acid, carbohydrate, and lipids. The one or more confounding targets may also comprise different types of biological entities, e.g., any combination of protein, nucleic acid, carbohydrate, and lipids.
[00364] In an embodiment, the invention provides an isolated polynucleotide, or a fragment thereof, identified by the methods above. The invention further provides an isolated polynucleotide having a nucleotide sequence that is at least 60% identical to the nucleotide sequence identified by the methods above. More preferably, the isolated nucleic acid molecule is at least 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more, identical to the nucleotide sequence identified by the methods above. In the case of an isolated polynucleotide which is longer than or equivalent in length to the reference sequence, e.g., a sequence identified by the methods above, the comparison is made with the full length of the reference sequence. Where the isolated polynucleotide is shorter than the reference sequence, e.g., shorter than a sequence identified by the methods above, the comparison is made to a segment of the reference sequence of the same length (excluding any loop required by the homology calculation).

[00365] In a related aspect, the invention provides a method of selecting a group of aptamers, comprising: (a) contacting a pool of aptamers to a population of microvesicles from a first sample; (b) enriching a subpool of aptamers that show affinity to the population of microvesicles from the first sample; (c) contacting the subpool to a second population of microvesicles from a second sample; and (d) depleting a second subpool of aptamers that show affinity to the second population of microvesicles from the second sample, thereby selecting the group of aptamers that have preferential affinity for the population of microvesicles from the first sample.

[00366] The first sample and/or second sample may comprise a biological fluid such as disclosed herein. For example, the biological fluid may include without limitation blood, a blood derivative, plasma, serum or urine. The first sample and/or second sample may also be derived from a cell culture.

[00367] In an embodiment, the first sample comprises a cancer sample and the second sample comprises a control sample, such as a non-cancer sample. The first sample and/or the second sample may each comprise a pooled sample. For example, the first sample and/or second sample can comprise bodily fluid from 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 40, 50, 60, 70, 80, 90, 100 or more than 100 individuals. In such cases, the members of a pool may be chosen to represent a desired phenotype. In a non-limiting example, the members of the first sample pool may be from patients with a cancer and the members of the second sample pool may be from non-cancer controls.

[00368] Steps (a)-(d) can be repeated a desired number of times in order to further enrich the pool in aptamers that have preferential affinity for the population of microvesicles from the first sample. For example, steps (a)-(d) can be repeated 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more than 20 times. The output from step (d) can be used as the input to repeated step (a). In embodiment, the first sample and/or second sample are replaced with a different sample before repeating steps (a)-(d). In a non-limiting example, members of a first sample pool may be from patients with a cancer and members of a second sample pool may be from non-cancer controls. During subsequent repetitions of steps (a)-(d), the first sample pool may comprise samples from different cancer patients than in the prior round/s. Similarly, the second sample pool may comprise samples from different controls than in the prior round/s.
In still another related aspect, the invention provides a method of enriching a plurality of oligonucleotides, comprising: (a) contacting a first microvesicle population with the plurality of oligonucleotides; (b) fractionating the first microvesicle population contacted in step (a) and recovering members of the plurality of oligonucleotides that fractionated with the first microvesicle population; (c) contacting the recovering members of the plurality of oligonucleotides from step (b) with a second microvesicle population; (d) fractionating the second microvesicle population contacted in step (c) and recovering members of the plurality of oligonucleotides that did not fractionate with the second microvesicle population; (e) contacting the recovering members of the plurality of oligonucleotides from step (d) with a third microvesicle population; and (f) fractionating the third microvesicle population contacted in step (a) and recovering members of the plurality of oligonucleotides that fractionated with the third microvesicle population; thereby enriching the plurality of oligonucleotides. The first and third microvesicle populations may have a first phenotype while the second microvesicle population has a second phenotype. Thus, positive selection occurs for the microvesicle populations associated with the first phenotype and negative selection occurs for the microvesicle populations associated with the second phenotype. An example of such selection schemes is described in Example 18 herein, wherein the first phenotype comprises biopsy-positive breast cancer and the second phenotype comprises non-breast cancer (biopsy-negative or healthy).

In some embodiments, the first phenotype comprises a medical condition, disease or disorder and the second phenotype comprises a healthy state or a different state of the medical condition, disease or disorder. The first phenotype can be a healthy state and the second phenotype comprises a medical condition, disease or disorder. The medical condition, disease or disorder can be any detectable medical condition, disease or disorder, including without limitation a cancer, a premalignant condition, an inflammatory disease, an immune disease, an autoimmune disease or disorder, a cardiovascular disease or disorder, neurological disease or disorder, infectious disease or pain. Various types of such conditions are disclosed herein. See, e.g., Section "Phenotypes" herein.

Any useful method to isolate microvesicles in whole or in part can be used to fractionate the samples. See, e.g., Section "Microvesicle Isolation and Analysis" herein. In an embodiment, the fractionating comprises ultracentrifugation in step (b) and polymer precipitation in steps (d) and (f). The polymer can be polyethylene glycol (PEG). Any appropriate form of PEG may be used. For example, the PEG may be PEG 8000. The PEG may be used at any appropriate concentration. For example, the PEG can be used at a concentration of 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14% or 15% to isolate the microvesicles. In some embodiments, the PEG is used at a concentration of 6%.

The contacting can be performed in the presence of a competitor, which may reduce non-specific binding events. Any useful competitor can be used. In an embodiment, the competitor comprises at least one of salmon sperm DNA, tRNA, dextran sulfate and carboxymethyl dextran. As desired, different competitors or competitor concentrations can be used at different contacting steps.

The method can be repeated to achieve a desired enrichment. In an embodiment, steps (a)-(f) are repeated at least once. These steps can be repeated 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17.
18, 19, 20, or more than 20 times as desired. At the same time, each of the contacting steps can be repeated as desired. In some embodiments, the method further comprises: (i) repeating steps (a)-(b) at least once prior to step (c), wherein the recovered members of the plurality of oligonucleotides that fractionated with the first microvesicle population in step (b) are used as the input plurality of oligonucleotides for the repetition of step (a); (ii) repeating steps (c)-(d) at least once prior to step (e), wherein the recovered members of the plurality of oligonucleotides that did not fractionate with the second microvesicle population in step (d) are used as the input plurality of oligonucleotides for the repetition of step (c); and/or (iii) repeating steps (e)-(f) at least once, wherein the recovered members of the plurality of oligonucleotides that fractionated with the third microvesicle population in step (f) are used as the input plurality of oligonucleotides for the repetition of step (e). Repetitions (i)-(iii) can be repeated any desired number of times, e.g., (i)-(iii) can be repeated 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more than 20 times. In an embodiment, (i)-(iii) each comprise three repetitions.

[00374] The method may further comprise identifying the members of the selected group of aptamers or oligonucleotides, e.g., by DNA sequencing. The sequencing may be performed by Next Generation sequencing as desired and after or before any desired step in the method.

[00375] The method may also comprise identifying the targets of the selected group of aptamers/oligonucleotides. Methods to identify such targets are disclosed herein.

[00376] **Oligonucleotide Probe Target Identification**

[00377] The methods and kits above can be used to identify binding agents that differentiate between two biomarker populations. The invention further provides methods of identifying the targets of binding agent. For example, the methods may further comprise identifying a surface marker of a target microvesicle that is recognized by the binding agent.

[00378] In an embodiment, the invention provides a method of identifying a target of a binding agent comprising: (a) contacting the binding agent with the target to bind the target with the binding agent, wherein the target comprises a surface antigen of a microvesicle; (b) disrupting the microvesicle under conditions which do not disrupt the binding of the target with the binding agent; (c) isolating the complex between the target and the binding agent; and (d) identifying the target bound by the binding agent. The binding agent can be a binding agent identified by the methods above, e.g., an aptamer, ligand, antibody, or other useful binding agent that can differentiate between two populations of biomarkers.

[00379] An illustrative schematic for carrying on the method is shown in FIG. 9. The figure shows a binding agent 902, here an oligonucleotide probe or aptamer for purposes of illustration, tethered to a substrate 901. The binding agent 902 can be covalently attached to substrate 901. The binding agent 902 may also be non-covalently attached. For example, binding agent 902 can comprise a label which can be attracted to the substrate, such as a biotin group which can form a complex with an avidin/streptavidin molecule that is covalently attached to the substrate. This can allow a complex to be formed between the aptamer and the microvesicle while in solution, followed by capture of the aptamer using the biotin label. The binding agent 902 binds to a surface antigen 903 of microvesicle 904. In the step signified by arrow
(i), the microvesicle is disrupted while leaving the complex between the binding agent 902 and surface antigen 903 intact. Disrupted microvesicle 905 is removed, e.g., via washing or buffer exchange, in the step signified by arrow (ii). In the step signified by arrow (iii), the surface antigen 903 is released from the binding agent 902. The surface antigen 903 can be analyzed to determine its identity using methods disclosed herein and/or known in the art. The target of the method can be any useful biological entity associated with a microvesicle. For example, the target may comprise a protein, nucleic acid, lipid or carbohydrate, or other biological entity disclosed herein or known in the art.

[00380] In some embodiments of the method, the target is cross-linked to the binding agent prior disrupting the microvesicle. Without being bound by theory, this step may assist in maintaining the complex between the binding agent and the target while the vesicle is disrupted. Any useful method of crosslinking disclosed herein or known in the art can be used. In embodiments, the cross-linking comprises photocrosslinking, an imidoester crosslinker, dimethyl suberimidate, an N-Hydroxysuccinimide-ester crosslinker, bisulfosuccinimidyl suberate (BS3), an aldehyde, acrolein, crotonaldehyde, formaldehyde, a carbodiimide crosslinker, N,N'-dicyclohexylcarbodiimide (DDC), N,N'-diisopropylcarbodiimide (DIC), 1-Ethyl-3-[3-dimethylaminopropyl]carbodiimide hydrochloride (EDC or EDAC), Succinimidyl-4-(N-maleimidomethyl)cyclohexane-1-carboxylate (SMCC), a Sulfosuccinimidyl-4-(N-maleimidomethyl)cyclohexane-1-carboxylate (Sulfo-SMCC), a Sulfo-N-hydroxysuccinimidyl-2-(6-[biontanamido]-2-(p-azido benzamido)-hexanoamido) ethyl-L3'-dithiopropionate (Sulfo-SBED), 2-[N2-(4-Azido-2,3,5,6-tetrafluorobenzoyl)-N6-(6-biotin-amidocaproyl)-L-lysiny1]ethyl methanethiosulfonate (Mts-Atf-Biotin; available from Thermo Fisher Scientific Inc, Rockford IL.), 2-[N2-[N6-(4-Azido-2,3,5,64tetrafluorobenzoyl-6-amino-caproyl)-N6-(6-biotinamidocaproyl)-L4ysinylamido]ethyl methanethiosulfonate (Mts-Atf-LC-Biotin; available from Thermo Fisher Scientific Inc), a photoreactive amino acid (e.g., L-Photo-Leucine and L-Photo-Methionine, see, e.g., Suchanek, M., et al. (2005). Photo-leucine and photo-methionine allow identification of protein-protein interactions. Nat. Methods 2:261-267), an N-Hydroxysuccinimide (NHS) crosslinker, an NHS-Azide reagent (e.g., NHS-Azide, NHS-PEG4-Azide, NHS-PEG12-Azide; each available from Thermo Fisher Scientific, Inc.), an NHS-Phosphine reagent (e.g., NHS-Phosphine, Sulfo-NHS-Phosphine; each available from Thermo Fisher Scientific, Inc.), or any combination or modification thereof.

[00381] A variety of methods can be used to disrupt the microvesicle. For example, the vesicle membrane can be disrupted using mechanical forces, chemical agents, or a combination thereof. In embodiments, disrupting the microvesicle comprises use of one or more of a detergent, surfactant, a solvent, an enzyme, or any useful combination thereof. The enzyme may comprise one or more of lysozyme, lysostaphin, zymolase, cellulase, mutanolysin, aglycanase, a protease, and mannase. The detergent or surfactant may comprise one or more of octylthioglucoside (OTG), octyl beta-glucoside (OG), a nonionic detergent, Triton X, Tween 20, a fatty alcohol, a cetyl alcohol, a stearyl alcohol, cetostearyl alcohol, an oleyl alcohol, a polyoxyethylene glycol alkyl ether (Brij), octaethylene glycol monododecyl ether, pentaethylene glycol monododecyl ether, a polyoxypropylene glycol alkyl ether, a glucose alkyl ether, decyl glucoside, lauryl glucoside, octyl glucoside, a polyoxyethylene glycol octylphenol ethers, a
polyoxyethylene glycol alkylphenol ether, nonoxynol-9, a glycerol alkyl ester, glyceryl laurate, a polyoxyethylene glycol sorbitan alkyl esters, polysorbate, a sorbitan alkyl ester, cocamide MEA, cocamide DEA, dodecyldimethylamine oxide, a block copolymers of polyethylene glycol and polypropylene glycol, poloxamers, polyethoxylated tallow amine (POEA), a zwitterionic detergent, 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS), a linear alkylbenzene sulfonate (LAS), a alkyl phenol ethoxylate (APE), cocamidopropyl hydroxysultaine, a betaine, cocamidopropyl betaine, lecithin, an ionic detergent, sodium dodecyl sulfate (SDS), cetrimonium bromide (CTAB), cetyl trimethylammonium chloride (CTAC), octenidine dihydrochloride, cetylpyridinium chloride (CPC), benzalkonium chloride (BAC), benzethonium chloride (BZT), 5-Bromo-5-nitro-1,3-dioxane, dimethyldioctadecylammonium chloride, dioctadecyldimethylammonium bromide (DODAB), sodium deoxycholate, nonyl phenoxypolyethoxylethanol (Tergitol-type NP-40; NP-40), ammonium lauryl sulfate, sodium laureth sulfate (sodium lauryl ether sulfate (SLES)), sodium myreth sulfate, an alkyl carboxylate, sodium stearate, sodium lauroyl sarcosinate, a carboxylate-based fluorosurfactant, perfluorononanoate, perfluorooctanoate (PFOA or PFO), and a biosurfactant. Mechanical methods of disruption that can be used comprise without limitation mechanical shear, bead milling, homogenation, microfluidization, sonication, French Press, impingement, a colloid mill, decompression, osmotic shock, thermolysis, freeze-thaw, desiccation, or any combination thereof.

[00382] As shown in FIG. 9, the binding agent may be tethered to a substrate. The binding agent can be tethered before or after the complex between the binding agent and target is formed. The substrate can be any useful substrate such as disclosed herein or known in the art. In an embodiment, the substrate comprises a microsphere. In another embodiment, the substrate comprises a planar substrate. The binding agent can also be labeled. Isolating the complex between the target and the binding agent may comprise capturing the binding agent via the label. For example, the label can be a biotin label. In such cases, the binding agent can be attached to the substrate via a biotin-avidin binding event.

[00383] Methods of identifying the target after release from the binding agent will depend on the type of target of interest. For example, when the target comprises a protein, identifying the target may comprise use of mass spectrometry (MS), peptide mass fingerprinting (PMF; protein fingerprinting), sequencing, N-terminal amino acid analysis, C-terminal amino acid analysis, Edman degradation, chromatography, electrophoresis, two-dimensional gel electrophoresis (2D gel), antibody array, and immunoassay. Nucleic acids can be identified by sequencing.

[00384] One of skill will appreciate that the method can be used to identify any appropriate target, including those not associated with a vesicle. For example, with respect to the FIG. 9, all steps except for the step signified by arrow (i) (i.e., disrupting the microvesicle), could be performed for a circulating target such as a protein, nucleic acid, lipid, carbohydrate, or combination thereof. The target can be any useful target, including without limitation a cell, an organelle, a protein complex, a lipoprotein, a carbohydrate, a microvesicle, a virus, a membrane fragment, a small molecule, a heavy metal, a toxin, a drug, a nucleic acid, mRNA, microRNA, a protein-nucleic acid complex, and various combinations, fragments and/or complexes of any of these.
In an aspect, the invention provides a method of identifying at least one protein associated with at least one microvesicle in a biological sample, comprising: a) contacting the at least one microvesicle with an oligonucleotide probe library, b) isolating at least one protein bound by at least one member of the oligonucleotide probe library in step a); and c) identifying the at least one protein isolated in step b). The isolating can be performed using any useful method such as disclosed herein, e.g., by immunoprecipitation or capture to a substrate. Similarly, the identifying can be performed using any useful method such as disclosed herein, including without limitation use of mass spectrometry, 2-D gel electrophoresis or an antibody array. Examples of such methodology are presented herein in Examples 23-25.

The targets identified by the methods of the invention can be detected, e.g., using the oligonucleotide probes of the invention, for various purposes as desired. For example, an identified microvesicle surface antigen can then be used to detect a microvesicle. In an aspect, the invention provides a method of detecting at least one microvesicle in a biological sample comprising contacting the biological sample with at least one binding agent to at least one microvesicle surface antigen and detecting the at least one microvesicle recognized by the binding agent to the at least one protein. In an embodiment, the at least one microvesicle surface antigen is selected from Tables 3-4 herein. The at least one microvesicle surface antigen can be a protein in any of Tables 18-25. See Example 23. The at least one binding agent may comprise any useful binding agent, including without limitation a nucleic acid, DNA molecule, RNA molecule, antibody, antibody fragment, aptamer, peptoid, peptide, locked nucleic acid (LNA), lectin, peptide, dendrimer, membrane protein labeling agent, chemical compound, or a combination thereof. In some embodiments, the at least one binding agent comprises at least one oligonucleotide, such as an oligonucleotide probe as provided herein.

The at least one binding agent can be used to capture and/or detect the at least one microvesicle. Methods of detecting biomarkers and microvesicle using binding agents are provided herein. See, e.g., FIGs. 2A-B, which figures describe sandwich assay formats. In some embodiments, the at least one binding agent used to capture the at least one microvesicle is bound to a substrate. Any useful substrate can be used, including without limitation a planar array, a column matrix, or a microbead. See, e.g., FIGs. 2A-B. In some embodiments, the at least one binding agent used to detect the at least one microvesicle is labeled. Various useful labels are provided herein or known in the art, including 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, a semiconductor nanocrystal, a nanoparticle, a quantum dot, a gold particle, a fluorophore, or a radioactive label.

In an embodiment, the detecting is used to characterize a phenotype. The phenotype can be any appropriate phenotype of interest. In some embodiments, the phenotype is a disease or disorder. The characterizing may comprise providing diagnostic, prognostic and/or theranostic information for the disease or disorder. The characterizing may be performed by comparing a presence or level of the at least one microvesicle to a reference. The reference can be selected per the characterizing to be performed. For example, when the phenotype comprises a disease or disorder, the reference may comprise a presence or
level of the at least one microvesicle in a sample from an individual or group of individuals without the disease or disorder. The comparing can be determining whether the presence or level of the microvesicle differs from that of the reference. In some embodiments, the detected at least one microvesicle is found at higher levels in a healthy sample as compared to a diseased sample. In another embodiment, the detected at least one microvesicle is found at higher levels in a diseased sample as compared to a healthy sample. When multiplex assays are performed, e.g., using a plurality of binding agents to different biomarkers, some microvesicle antigens may be observed at a higher level in the biological samples as compared to the reference whereas other microvesicle antigens may be observed at a lower level in the biological samples as compared to the reference.

[00389] The method can be used to detect the at least one microvesicle in any appropriate biological sample. For example, the biological sample may comprise a bodily fluid, tissue sample or cell culture. The bodily fluid or tissue sample can be from a subject having or suspected of having a medical condition, a disease or a disorder. Thus, the method can be used to provide a diagnostic, prognostic, or theranostic read out for the subject. Any appropriate bodily fluid can be used, including without limitation peripheral blood, sera, plasma, ascites, urine, cerebrospinal fluid (CSF), sputum, saliva, bone marrow, synovial fluid, aqueous humor, amniotic fluid, serum, breast milk, bronchoalveolar lavage fluid, semen, prostatic fluid, cowper's fluid or pre-ejaculatory fluid, female ejaculate, sweat, fecal matter, hair oil, 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, blastocyl cavity fluid, or umbilical cord blood.

[00390] The method of the invention can be used to detect or characterize any appropriate disease or disorder of interest, including without limitation Breast Cancer, Alzheimer's disease, bronchial asthma, Transitional cell carcinoma of the bladder, Giant cellular osteoblastoclastoma, Brain Tumor, Colorectal adenocarcinoma, Chronic obstructive pulmonary disease (COPD), Squamous cell carcinoma of the cervix, acute myocardial infarction (AMI) / acute heart failure, Chron's Disease, diabetes mellitus type II, Esophageal carcinoma, Squamous cell carcinoma of the larynx, Acute and chronic leukemia of the bone marrow, Lung carcinoma, Malignant lymphoma, Multiple Sclerosis, Ovarian carcinoma, Parkinson disease, Prostate adenocarcinoma, psoriasis, Rheumatoid Arthritis, Renal cell carcinoma, Squamous cell carcinoma of skin, Adenocarcinoma of the stomach, carcinoma of the thyroid gland, Testicular cancer, ulcerative colitis, or Uterine adenocarcinoma.

[00391] In some embodiments, the disease or disorder comprises a cancer, a premalignant condition, an inflammatory disease, an immune disease, an autoimmune disease or disorder, a cardiovascular disease or disorder, neurological disease or disorder, infectious disease or pain. The cancer can include without limitation one of 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; 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 carcinoma 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; lung 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; Sezary 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; Waldenstrom macroglobulinemia; or Wilm's tumor. The premalignant condition can include without limitation Barrett's Esophagus. The autoimmune disease can include without limitation one of 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 erythematosus (SLE), Hashimoto's Thyroiditis, Grave's disease, Ankylosing Spondylitis Sjogrens Disease, CREST syndrome, Scleroderma, Rheumatic Disease, organ rejection, Primary Sclerosing Cholangitis, or sepsis. The
cardiovascular disease can include without limitation one of atherosclerosis, congestive heart failure, vulnerable plaque, stroke, ischemia, high blood pressure, stenosis, vessel occlusion or a thrombotic event. The neurological disease can include without limitation one of Multiple Sclerosis (MS), Parkinson's Disease (PD), Alzheimer's Disease (AD), schizophrenia, bipolar disorder, depression, autism, Prion Disease, Pick's disease, dementia, Huntington disease (HD), Down's syndrome, cerebrovascular disease, Rasmussen's encephalitis, viral meningitis, neuropsychiatric systemic lupus erythematosus (NPSLE), amyotrophic lateral sclerosis, Creutzfeldt-Jacob disease, Gerstmann-Straussler-Scheinker disease, transmissible spongiform encephalopathy, ischemic reperfusion damage (e.g. stroke), brain trauma, microbial infection, or chronic fatigue syndrome. The pain can include without limitation one of fibromyalgia, chronic neuropathic pain, or peripheral neuropathic pain. The infectious disease can include without limitation one of a bacterial infection, viral infection, yeast infection, Whipple's Disease, Prion Disease, cirrhosis, methicillin-resistant staphylococcus aureus, HIV, HCV, hepatitis, syphilis, meningitis, malaria, tuberculosis, or influenza. One of skill will appreciate that oligonucleotide probes or plurality of oligonucleotides or methods of the invention can be used to assess any number of these or other related diseases and disorders.

[00392] In a related aspect, the invention provides a kit comprising a reagent for carrying out the methods herein. In still another related aspect, the invention provides for use of a reagent for carrying out the methods. The reagent may comprise at least one binding agent to the at least one protein. The binding agent may be an oligonucleotide probe as provided herein.

[00393] Sample Characterization

[00394] The aptamers of the invention can be used to characterize a biological sample. For example, an aptamer can be used to bind a biomarker in the sample. The presence or level of the bound biomarker can indicate a characteristic of the example, such as a diagnosis, prognosis or theranosis of a disease or disorder associated with the sample.

[00395] In an aspect, the invention provides an aptamer comprising a nucleic acid sequence that is at least about 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 96, 97, 98, 99 or 100 percent homologous of any of: a) SEQ ID NOs. 11-24 or a sequence in Table 8; a) SEQ ID NOs. 25-44 or a sequence in Table 11; b) SEQ ID NOs. 47-130 or a sequence in Table 12; or c) a functional variation or fragment of any preceding sequence. A functional variation or fragment includes a sequence comprising modifications that is still capable of binding a target molecule, wherein the modifications comprise without limitation at least one of a deletion, insertion, point mutation, truncation or chemical modification. In a related aspect, the invention provides a method of characterizing a disease or disorder, comprising: (a) contacting a biological test sample with one or more aptamer of the invention, e.g., any of those in this paragraph or modifications thereof; (b) detecting a presence or level of a complex between the one or more aptamer and the target bound by the one or more aptamer in the biological test sample formed in step (a); (c) contacting a biological control sample with the one or more aptamer; (d) detecting a presence or level of a complex between the one or more aptamer and the target bound by the one or more aptamer in the
biological control sample formed in step (c); and (e) comparing the presence or level detected in steps (b) and (d), thereby characterizing the disease or disorder.

[00396] The biological test sample and biological control sample can each comprise a tissue sample, a cell culture, or a biological fluid. In some embodiments, the biological test sample and biological control sample may comprise the same sample type, e.g., both are tissue samples or both are fluid samples. In other embodiments, different sample types may be used for the test and control samples. For example, the control sample may comprise an engineered or otherwise artificial sample.

[00397] The biological fluid may comprise a bodily fluid. The bodily fluid may include without limitation one or more of peripheral blood, sera, plasma, ascites, urine, cerebrospinal fluid (CSF), sputum, saliva, bone marrow, synovial fluid, aqueous humor, amniotic fluid, cerumen, breast milk, bronchoalveolar 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, bile, interstitial fluid, menses, pus, sebum, vomit, vaginal secretions, mucusal secretion, stool water, pancreatic juice, lavage fluids from sinus cavities, bronchopulmonary aspirates, blastocly cavity fluid, or umbilical cord blood. In some embodiments, the bodily fluid comprises blood, serum or plasma.

[00398] The biological fluid may comprise microvesicles. For example, the biological fluid can be a tissue, cell culture, or bodily fluid which comprises microvesicles released from cells in the sample. The microvesicles can be circulating microvesicles.

[00399] The one or more aptamer can bind a target biomarker, e.g., a biomarker useful in characterizing the sample. The biomarker may comprise a polypeptide or fragment thereof, or other useful biomarker described herein or known in the art (lipid, carbohydrate, complex, nucleic acid, etc). In embodiments, the polypeptide or fragment thereof is soluble or membrane bound. Membrane bound polypeptides may comprise a cellular surface antigen or a microvesicle surface antigen. The biomarker can be a biomarker selected from Table 3 or Table 4.

[00400] The characterizing can comprises a diagnosis, prognosis or theranosis of the disease or disorder. Various diseases and disorders can be characterized using the compositions and methods of the invention, including without limitation a cancer, a premalignant condition, an inflammatory disease, an immune disease, an autoimmune disease or disorder, a cardiovascular disease or disorder, a neurological disease or disorder, an infectious disease, and/or pain. See, e.g., section herein "Phenotypes" for further details. In embodiments, the disease or disorder comprises a proliferative or neoplastic disease or disorder. For example, the disease or disorder can be a cancer. In some embodiments, the cancer comprises a breast cancer, ovarian cancer, prostate cancer, lung cancer, colorectal cancer, melanoma, or brain cancer.

[00401] FIG. 19A is a schematic 1900 showing an assay configuration that can be used to detect and/or quantify a target of interest using one or more aptamer of the invention. Capture aptamer 1902 is attached to substrate 1901. The substrate can be a planar substrate, well, microbead, or other useful substrate as disclosed herein or known in the art. Target of interest 1903 is bound by capture aptamer 1902. The target of interest can be any appropriate entity that can be detected when recognized by an aptamer or other binding agent. The target of interest may comprise a protein or polypeptide, a nucleic acid, including
DNA, RNA, and various subspecies thereof, a lipid, a carbohydrate, a complex, e.g., a complex comprising protein, nucleic acids, lipids and/or carbohydrates. In some embodiments, the target of interest comprises a microvesicle. The target of interest can be a microvesicle surface antigen. The target of interest may be a biomarker, including a vesicle associated biomarker, in Tables 3 or 4. The microvesicle input can be isolated from a sample using various techniques as described herein, e.g., chromatography, filtration, centrifugation, flow cytometry, affinity capture (e.g., to a planar surface, column or bead), and/or using microfluidics. Detection aptamer 1904 is also bound to target of interest 1903. Detection aptamer 1904 carries label 1905 which can be detected to identify target captured to substrate 1901 via capture aptamer 1902. The label can be a fluorescent, radiolabel, enzyme, or other detectable label as disclosed herein. Either capture aptamer 1902 or detection aptamer 1904 can be substituted with another binding agent, e.g., an antibody. For example, the target may be captured with an antibody and detected with an aptamer, or vice versa. When the target of interest comprises a complex, the capture and detection agents (aptamer, antibody, etc) can recognize the same or different targets. For example, when the target is a microvesicle, the capture agent may recognize one microvesicle surface antigen while the detection agent recognizes another microvesicle surface antigen. Alternately, the capture and detection agents can recognize the same surface antigen.

[00402] The aptamers of the invention may be identified and/or used for various purposes in the form of DNA or RNA. Unless otherwise specified, one of skill in the art will appreciate that an aptamer may generally be synthesized in various forms of nucleic acid. The aptamers may also carry various chemical modifications and remain within the scope of the invention.

[00403] In some embodiments, an aptamer of the invention is modified to comprise at least one chemical modification. The modification may include without limitation a chemical substitution at a sugar position; a chemical substitution at a phosphate position; and a chemical substitution at a base position of the nucleic acid. In some embodiments, the modification is selected from the group consisting of: biotinyla tion, incorporation of a fluorescent label, incorporation of a modified nucleotide, a 2′-modified pyrimidine, 3′ capping, conjugation to an amine linker, conjugation to a high molecular weight, non-immunogenic compound, conjugation to a lipophilic compound, conjugation to a drug, conjugation to a cytotoxic moiety, and labeling with a radioisotope, or other modification as disclosed herein. The position of the modification can be varied as desired. For example, the biotinylat ion, fluorescent label, or cytotoxic moiety can be conjugated to the 5′ end of the aptamer. The biotinylat ion, fluorescent label, or cytotoxic moiety can also be conjugated to the 3′ end of the aptamer.

[00404] In some embodiments, the cytotoxic moiety is encapsulated in a nanoparticle. The nanoparticle can be selected from the group consisting of: liposomes, dendrimers, and comb polymers. In other embodiments, the cytotoxic moiety comprises a small molecule cytotoxic moiety. The small molecule cytotoxic moiety can include without limitation vinblastine hydrazide, calicheamicin, vinca alkaloid, a cryptophycin, atubulysin, dolastatin-10, dolastatin-15, auristatin E, rhizoxin, epothilone B, epithilone D, taxoids, maytansinois and any variants and derivatives thereof. In still other embodiments, the cytotoxic moiety comprises a protein toxin. For example, the protein toxin can be selected from the group consisting
of diphtheria toxin, ricin, abrin, gelonin, and Pseudomonas exotoxin A. Non-immunogenic, high molecular weight compounds for use with the invention include polyalkylene glycols, e.g., polyethylene glycol. Appropriate radioisotopes include yttrium-90, indium-111, iodine-131, lutetium-177, copper-67, rhenium-186, rhenium-188, bismuth-212, bismuth-213, astatine-211 and actinium-225. The aptamer may be labeled with a gamma-emitting radioisotope.

[00405] In some embodiments of the invention, an active agent is conjugated to the aptamer. For example, the active agent may be a therapeutic agent or a diagnostic agent. The therapeutic agent may be selected from the group consisting of tyrosine kinase inhibitors, kinase inhibitors, biologically active agents, biological molecules, radionuclides, radionuclides, anthracyclines, ansamycin antibiotics, asparaginase, bleomycin, busulphan, cisplatin, carboplatin, carmustine, capecitabine, chlorambucil, cytarabine, cyclophosphamide, camptothecin, dacarbazine, dactinomycin, daunorubicin, dexamethasone, docetaxel, doxorubicin, etoposide, epothilones, floxuridine, fludarabine, fluorouracil, gemcitabine, hydroxyurea, idarubicin, ifosfamide, irinotecan, lomustine, mechloretamine, mercaptopurine, melphanal, methotrexate, rapamycin (sirolimus), mitomycin, mitotane, mitoxantrone, nitrosurea, paclitaxel, pamidronate, pentostatin, plicamycin, procarbazine, rituximab, streptozocin, teniposide, thioguanine, thiotapec, taxanes, vinblastine, vincristine, vinorelbine, taxol, combretastatins, discodermolides, transplatinum, anti-vascular endothelial growth factor compounds ("anti-VEGFs"), anti-epidermal growth factor receptor compounds ("anti-EGFRs"), 5-fluorouracil and derivatives, radionuclides, polypeptide toxins, apoptosis inducers, therapy sensitizers, enzyme or active protein, and combinations thereof.

[00406] Oligonucleotide Pools to Characterize a Sample

[00407] The complexity and heterogeneity present in biology challenges the understanding of biological systems and disease. Diversity exists at various levels, e.g., within and between cells, tissues, individuals and disease states. See, e.g., FIG. 20A. FIG. 20B overviews various biological entities that can be assessed to characterize such samples. As shown in the Figure, as one moves from assessing DNA, to RNA, to protein, and finally to protein complexes, the amount of diversity and complexity increases dramatically. The oligonucleotide probe library method of the invention can be used characterize complex biological sources, e.g., tissue samples, cells, circulating tumor cells, microvesicles, and complexes such as protein and proteolipid complexes.

[00408] Current methods to characterize biological samples may not adequately address such complexity and diversity. As shown in FIG. 20C, such current methods often have a trade off between measuring diversity and complexity. As an example, consider high throughput sequencing technology. Next generation approaches may query many 1000s of molecular targets in a single assay. However, such approaches only probe individual DNA and/or RNA molecules, and thus miss out on the great diversity of proteins and biological complexes. On the other hand, flow cytometry can probe biological complexes, but are limited to a small number of pre-defined ligands. For example, a single assay can probe a handful of differentially labeled antibodies to pre-defined targets.

[00409] The oligonucleotide probe library of the invention address the above challenges with current biological detection technologies. The size of the starting library can be adjusted to measure as many
different entities as there are library members. In this Example, the initial untrained oligonucleotide library has the potential to measure $10^{12}$ or more biological features. A larger and/or different library can be constructed as desired. The technology is adapted to find differences between samples without assumptions about what "should be different." For example, the probe library may distinguish based on individual proteins, protein modifications, protein complexes, lipids, nucleic acids, different folds or conformations, or whatever is there that distinguishes a sample of interest. Thus, the method provides an unbiased approach to identify differences in biological samples that can be used to identify different populations of interest.

[00410] In the context herein, the use of the oligonucleotide library probe to assess a sample may be referred to as Adaptive Dynamic Artificial Poly-ligand Targeting, or ADAPT™ (previously referred to as Topological Oligonucleotide Profiling: TOP™). Although as noted the terms aptamer and oligonucleotides are typically used interchangeable herein, some differences between "classic" individual aptamers and ADAPT probes are as follows. Individual aptamers may comprise individual oligonucleotides selected to bind to a known specific target in an antibody-like "key-in-lock" binding mode. They may be evaluated individually based on specificity and binding affinity to the intended target. However, ADAPT probes may comprise a library of oligonucleotides intended to produce multi-probe signatures. The ADAPT probes comprise numerous potential binding modalities (electrostatic, hydrophobic, Watson-Crick, multi-oligo complexes, etc.). The ADAPT probe signatures have the potential to identify heterogeneous patient subpopulations. For example, a single ADAPT probe library can be assembled to differentiate multiple disease states, as demonstrated herein. Unlike classic single aptamers, the binding targets may or may not be isolated or identified. It will be understood that screening methods that identify individual aptamers, e.g., SELEX, can also be used to enrich a naive library of oligonucleotides to identify a ADAPT probe library.

[00411] The general method of the invention is outlined in FIG. 20D. One input to the method comprises a randomized oligonucleotide library with the potential to measure $10^{12}$ or more biological features. As outlined in the figure, the method identifies a desired number (e.g., $\sim 10^2-10^9$) that are different between two input sample types. The randomized oligonucleotide library is contacted with a first and a second sample type, and oligonucleotides that bind to each sample are identified. The bound oligonucleotide populations are compared and oligonucleotides that specifically bind to one or the other biological input sample are retained for the oligonucleotide probe library, whereas oligonucleotides that bind both biological input samples are discarded. This trained oligonucleotide probe library can then be contacted with a new test sample and the identities of oligonucleotides that bind the test sample are determined. The test sample is characterized based on the profile of oligonucleotides that bound. See, e.g., FIG. 20H.

[00412] Extracellular vesicles provide an attractive vehicle to profile the biological complexity and diversity driven by many inter-related sources. There can be a great deal of heterogeneity between patient-to-patient microvesicle populations, or even in microvesicle populations from a single patient under different conditions (e.g., stress, diet, exercise, rest, disease, etc). Diversity of molecular phenotypes within microvesicle populations in various disease states, even after microvesicle isolation and sorting by
vesicle biomarkers, can present challenges identifying surface binding ligands. This situation is further complicated by vesicle surface-membrane protein complexes. The oligonucleotide probe library can be used to address such challenges and allow for characterization of biological phenotypes. The approach combines the power of diverse oligonucleotide libraries and high throughput (next-generation) sequencing technologies to probe the complexity of extracellular microvesicles. See FIG. 20E.

[00413] ADAPT™ profiling may provide quantitative measurements of dynamic events in addition to detection of presence/absence of various biomarkers in a sample. For example, the binding probes may detect protein complexes or other post-translation modifications, allowing for differentiation of samples with the same proteins but in different biological configurations. Such configurations are illustrated in FIGs. 20F-G. In FIG. 20F, microvesicles with various surface markers are shown from an example microvesicle sample population: Sample Population A. The indicated Bound Probing Oligonucleotides 2001 are contacted to two surface markers 2002 and 2003 in a given special relationship. Here, probes unique to these functional complexes and spatial relationships may be retained. In contrast, in microvesicle Sample Population B shown in FIG. 20F, the two surface markers 2002 and 2003 are found in disparate spatial relationship. Here, probes 2001 are not bound due to absence of the spatial relationship of the interacting components 2002 and 2003.

[00414] An illustrative approach 2010 for using ADAPT profiling to assess a sample is shown in FIG. 20H. The probing library 2011 is mixed with sample 2012. The sample can be as described herein, e.g., a bodily fluid from a subject having or suspected of having a disease. The probes are allowed to bind the sample 2020 and the microvesicles are pelleted 2015. The supernatant 2014 comprising unbound oligonucleotides is discarded. Oligonucleotide probes bound to the pellet 2015 are eluted 2016 and sequenced 2017. The profile 2018 generated by the bound oligonucleotide probes as determined by the sequencing 2017 is used to characterize the sample 2012. For example, the profile 2018 can be compared to a reference, e.g., to determine if the profile is similar or different from a reference profile indicative of a disease or healthy state, or other phenotypic characterization of interest. The comparison may indicate the presence of a disease, provide a diagnosis, prognosis or theranosis, or otherwise characterize a phenotype associated with the sample 2012. FIG. 20I illustrates another schematic for using TOP™ profiling to characterize a phenotype. A patient sample such as a bodily fluid disclosed herein is collected 2021. The sample is contacted with the ADAPT™ library pool 2022. Microvesicles (MVs) are isolated from the contacted sample 2023, e.g., using ultracentrifugation, filtration, polymer precipitation or other appropriate technique or combination of techniques disclosed herein. Oligonucleotides that bound the isolated microvesicles are collected and identity is determined 2024. The identity of the bound oligonucleotides can be determined by any useful technique such as sequencing, high throughput sequencing (e.g., NGS), amplification including without limitation qPCR, or hybridization such as to a planar or particle based array. The identity of the bound oligonucleotides is used to characterize the sample, e.g., as containing disease related microvesicles.

[00415] In an aspect, the invention provides a method of characterizing a sample by contacting the sample with a pool of different oligonucleotides (e.g., an aptamer pool), and determining the frequency at which
various oligonucleotides in the pool bind the sample. For example, a pool of oligonucleotides is identified that preferentially bind to microvesicles from cancer patients as compared to non-cancer patients. A test sample, e.g., from a patient suspected of having the cancer, is collected and contacted with the pool of oligonucleotides. Oligonucleotides that bind the test sample are eluted from the test sample, collected and identified, and the composition of the bound oligonucleotides is compared to those known to bind cancer samples. Various sequencing, amplification and hybridization techniques can be used to identify the eluted oligonucleotides. For example, when a large pool of oligonucleotides is used, oligonucleotide identification can be performed by high throughput methods such as next generation sequencing or via hybridization. If the test sample is bound by the oligonucleotide pool in a similar manner (e.g., as determined by bioinformatics classification methods) to the microvesicles from cancer patients, then the test sample is indicative of cancer as well. Using this method, a pool of oligonucleotides that bind one or more microvesicle antigen can be used to characterize the sample without necessarily knowing the precise target of each member of the pool of oligonucleotides. Examples 26-30 and others herein illustrate embodiments of the invention.

[00416] In an aspect, the invention provides a method for characterizing a condition for a test sample comprising: contacting a microvesicle sample with a plurality of oligonucleotide capable of binding one or more target(s) present in said microvesicle sample, identifying a set of oligonucleotides that form a complex with the sample wherein the set is predetermined to characterize a condition for the sample, thereby characterizing a condition for a sample.

[00417] In an related aspect, the invention provides a method for identifying a set of oligonucleotides associated with a test sample, comprising: (a) contacting a microvesicle sample with a plurality of oligonucleotides, isolating a set of oligonucleotides that form a complex with the microvesicle sample, (b) determining sequence and/or copy number for each of the oligonucleotides, thereby identifying a set of oligonucleotides associated with the test sample.

[00418] In still another related aspect, the invention provides a method of diagnosing a sample as cancerous or predisposed to be cancerous, comprising contacting a microvesicle sample with a plurality of oligonucleotides that are predetermined to preferentially form a complex with microvesicles from a cancer sample as compared to microvesicles from a non-cancer sample.

[00419] The oligonucleotides can be identified by sequencing, e.g., by dye termination (Sanger) sequencing or high throughput methods. High throughput methods can comprise techniques to rapidly sequence a large number of nucleic acids, including next generation techniques such as Massively parallel signature sequencing (MPSS; Polony sequencing; 454 pyrosequencing; Illumina (Solexa) sequencing; SOLiD sequencing; Ion Torrent semiconductor sequencing; DNA nanoball sequencing; Heliscope single molecule sequencing; Single molecule real time (SMRT) sequencing, or other methods such as Nanopore DNA sequencing; Tunnelling currents DNA sequencing; Sequencing by hybridization; Sequencing with mass spectrometry; Microfluidic Sanger sequencing; Microscopy-based techniques; RNAP sequencing; In vitro virus high-throughput sequencing. The oligonucleotides may also be identified by hybridization techniques. For example, a microarray having addressable locals to hybridize and thereby detect the
various members of the pool can be used. Alternately, detection can be based on one or more 
differentially labelled oligonucleotides that hybridize with various members of the oligonucleotide pool.
The detectable signal of the label can be associated with a nucleic acid molecule that hybridizes with a 
stretch of nucleic acids present in various oligonucleotides. The stretch can be the same or different as to 
one or more oligonucleotides in a library. The detectable signal can comprise fluorescence agents, 
including color-coded barcodes which are known, such as in U.S. Patent Application Pub. No. 
20140371088, 2013017837, and 20120258870. Other detectable labels (metals, radioisotopes, etc) can be 
used as desired.

The plurality or pool of oligonucleotides can comprise any desired number of oligonucleotides to 
allow characterization of the sample. In various embodiments, the pool comprises 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, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, 
400, 450, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, or at least 
10000 different oligonucleotide members.

The plurality of oligonucleotides can be pre-selected through one or more steps of positive or 
negative selection, wherein positive selection comprises selection of oligonucleotides against a sample 
having substantially similar characteristics compared to the test sample, and wherein negative selection 
comprises selection of oligonucleotides against a sample having substantially different characteristics 
compared to the test sample. Substantially similar characteristics mean that the samples used for positive 
selection are representative of the test sample in one or more characteristic of interest. For example, the 
samples used for positive selection can be from cancer patients or cell lines and the test sample can be a 
sample from a patient having or suspected to have a cancer. Substantially different characteristics mean 
that the samples used for negative selection differ from the test sample in one or more characteristic of 
interest. For example, the samples used for negative selection can be from individuals or cell lines that do 
not have cancer (e.g., "normal" or otherwise "control" samples) and the test sample can be a sample from 
a patient having or suspected to have a cancer. The cancer can be a breast cancer, ovarian cancer, prostate 
cancer, lung cancer, colorectal cancer, melanoma, brain cancer, or other cancer.

By selecting samples representative of the desired phenotypes to detect and/or distinguish, the 
characterizing can comprise a diagnosis, prognosis or theranosis for any number of diseases or disorders. 
Various diseases and disorders can be characterized using the compositions and methods of the invention, 
including without limitation a cancer, a premalignant condition, an inflammatory disease, an immune 
disease, an autoimmune disease or disorder, a cardiovascular disease or disorder, a neurological disease or 
disorder, an infectious disease, and/or pain. See, e.g., section herein "Phenotypes" for further details. In 
embodiments, the disease or disorder comprises a proliferative or neoplastic disease or disorder. For 
example, the disease or disorder can be a cancer. In some embodiments, the cancer comprises a breast 
cancer, ovarian cancer, prostate cancer, lung cancer, colorectal cancer, melanoma, or brain cancer.

FIG. 19B is a schematic showing use of an oligonucleotide pool to characterize a phenotype 
of a sample, such as those listed above. A pool of oligonucleotides to a target of interest is provided. 
For example, the pool of oligonucleotides can be enriched to target one or more microvesicle. The
members of the pool may bind different targets (e.g., a microvesicle surface antigen) or different epitopes of the same target present on the one or more microvesicle. The pool is contacted with a test sample to be characterized 1912. For example, the test sample may be a biological sample from an individual having or suspected of having a given disease or disorder. The mixture is washed to remove unbound oligonucleotides. The remaining oligonucleotides are eluted or otherwise disassociated from the sample and collected 1913. The collected oligonucleotides are identified, e.g., by sequencing or hybridization 1914. The presence and/or copy number of the identified is used to characterize the phenotype 1915. For example, the pool of oligonucleotides may be chosen as oligonucleotides that preferentially recognize microvesicles shed from cancer cells. The method can be employed to detect whether the sample retains oligonucleotides that bind the cancer-related microvesicles, thereby allowing the sample to be characterized as cancerous or not.

[00424] FIG. 19C is a schematic 1920 showing an implementation of the method in FIG. 19B. A pool of oligonucleotides identified as binding a microvesicle population is provided 1919. The input sample comprises a test sample comprising microvesicles 1922. For example, the test sample may be a biological sample from an individual having or suspected of having a given disease or disorder. The pool is contacted with the isolated microvesicles to be characterized 1923. The microvesicle population can be isolated before or after the contacting 1923 from the sample using various techniques as described herein, e.g., chromatography, filtration, ultrafiltration, centrifugation, ultracentrifugation, flow cytometry, affinity capture (e.g., to a planar surface, column or bead), polymer precipitation, and/or using microfluidics. The mixture is washed to remove unbound oligonucleotides and the remaining oligonucleotides are eluted or otherwise disassociated from the sample and collected 1924. The collected oligonucleotides are identified 1925 and the presence and/or copy number of the retained oligonucleotides is used to characterize the phenotype 1926 as above.

[00425] As noted, in embodiment of FIG. 19C, the pool of oligonucleotides 1919 is directly contacted with a biological sample that comprises or is expected to comprise microvesicles. Microvesicles are thereafter isolated from the sample and the mixture is washed to remove unbound oligonucleotides and the remaining oligonucleotides are disassociated and collected 1924. The following steps are performed as above. As an example of this alternate configuration, a biological sample, e.g., a blood, serum or plasma sample, is directly contacted with the pool of oligonucleotides. Microvesicles are then isolated by various techniques disclosed herein, including without limitation ultracentrifugation, ultrafiltration, flow cytometry, affinity isolation, polymer precipitation, chromatography, various combinations thereof, or the like. Remaining oligonucleotides are then identified, e.g., by sequencing, hybridization or amplification.

[00426] In a related aspect, the invention provides a composition of matter comprising a plurality of oligonucleotides that can be used to carry out the methods comprising use of an oligonucleotide pool to characterize a phenotype. The plurality of oligonucleotides can comprise any of those described herein.

[00427] In a related aspect, the invention provides a method of performing high-throughput sequencing comprising: performing at least one (i) negative selection or (ii) one positive selection of a plurality of oligonucleotides with a microvesicle sample; obtaining a set of oligonucleotides to provide a negative
binder subset or positive binder subset of the plurality of oligonucleotides, wherein the negative binder subset of the plurality of oligonucleotides does not bind the microvesicle sample and wherein the positive binder subset of the plurality of oligonucleotides does bind the microvesicle sample; contacting the negative binder subset or positive binder subset with a test sample; eluting oligonucleotides that bound to the test sample to provide a plurality of eluate oligonucleotides; and performing high-throughput sequencing of the plurality of eluate oligonucleotides to identify sequence and/or copy number of the members of the plurality of eluate oligonucleotides. Negative and positive selection of the plurality of oligonucleotides using microvesicle sample can be performed as disclosed herein. The oligonucleotide profile revealed by the sequence and/or copy number of the members of the plurality of eluate oligonucleotides can be used to characterize a phenotype of the test sample as described herein.

[00428] In a similar aspect, the invention provides a method for identifying oligonucleotides specific for a test sample. The method comprises: (a) enriching a plurality of oligonucleotides for a sample to provide a set of oligonucleotides predetermined to form a complex with a target sample; (b) contacting the plurality in (a) with a test sample to allow formation of complexes of oligonucleotides with test sample; (c) recovering oligonucleotides that formed complexes in (b) to provide a recovered subset of oligonucleotides; and (d) profiling the recovered subset of oligonucleotides by high-throughput sequencing or hybridization, thereby identifying oligonucleotides specific for a test sample. The test sample may comprise a plurality of microvesicles. The oligonucleotides may comprise RNA, DNA or both. In some embodiment, the method further comprises performing informatics analysis to identify a subset of oligonucleotides comprising sequence identity of at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or at least 99% to the oligonucleotides predetermined to form a complex with the target sample.

[00429] One of skill will appreciate that the method can be used to identify any appropriate target, including those not associated with a microvesicle. The target can be any useful target, including without limitation a cell, an organelle, a protein complex, a lipoprotein, a carbohydrate, a microvesicle, a virus, a membrane fragment, a small molecule, a heavy metal, a toxin, a drug, a nucleic acid (including without limitation microRNA (miR) and messenger RNA (mRNA)), a protein-nucleic acid complex, and various combinations, fragments and/or complexes of any of these. The target can, e.g., comprise a mixture of microvesicles and non-microvesicle entities.

[00430] In an aspect, the invention also provides a method comprising contacting an oligonucleotide or plurality of oligonucleotides with a sample and detecting the presence or level of binding of the oligonucleotide or plurality of oligonucleotides to a target in the sample, wherein the oligonucleotide or plurality of oligonucleotides can be those provided by the invention above. The sample may comprise a biological sample, an organic sample, an inorganic sample, a tissue, a cell culture, a bodily fluid, blood, serum, a cell, a microvesicle, a protein complex, a lipid complex, a carbohydrate, or any combination, fraction or variation thereof. The target may comprise a cell, an organelle, a protein complex, a lipoprotein, a carbohydrate, a microvesicle, a membrane fragment, a small molecule, a heavy metal, a toxin, or a drug.
[00431] In a related aspect, the invention provides a method comprising: a) contacting a biological sample comprising microvesicles with an oligonucleotide probe library, wherein optionally the oligonucleotide probe library comprises an oligonucleotide or plurality of oligonucleotides those provided by the invention above; b) identifying oligonucleotides bound to at least a portion of the microvesicles; and c) characterizing the sample based on a profile of the identified oligonucleotides.

[00432] In another aspect, the invention provides a method comprising: a) contacting a sample with an oligonucleotide probe library comprising at least $10^6$, $10^7$, $10^8$, $10^9$, $10^{10}$, $10^{11}$, $10^{12}$, $10^{13}$, $10^{14}$, $10^{15}$, $10^{16}$, $10^{17}$, or at least $10^{18}$ different oligonucleotide sequences oligonucleotides to form a mixture in solution, wherein the oligonucleotides are capable of binding a plurality of entities in the sample to form complexes, wherein optionally the oligonucleotide probe library comprises an oligonucleotide or plurality of oligonucleotides as provided by the invention above; b) partitioning the complexes formed in step (a) from the mixture; and c) detecting oligonucleotides present in the complexes partitioned in step (b) to identify an oligonucleotide profile for the sample. In an embodiment, the detecting step comprises performing sequencing of all or some of the oligonucleotides in the complexes, amplification of all or some of the oligonucleotides in the complexes, and/or hybridization of all or some of the oligonucleotides in the complexes to an array. The array can be any useful array, such as a planar or particle-based array.

[00433] In still another aspect, the invention provides a method for generating an enriched oligonucleotide probe library comprising: a) contacting a first oligonucleotide library with a biological test sample and a biological control sample, wherein complexes are formed between biological entities present in the biological samples and a plurality of oligonucleotides present in the first oligonucleotide library; b) partitioning the complexes formed in step (a) and isolating the oligonucleotides in the complexes to produce a subset of oligonucleotides for each of the biological test sample and biological control sample; c) contacting the subsets of oligonucleotides in (b) with the biological test sample and biological control sample wherein complexes are formed between biological entities present in the biological samples and a second plurality of oligonucleotides present in the subsets of oligonucleotides to generate a second subset group of oligonucleotides; and d) optionally repeating steps b)-c), one, two, three or more times to produce a respective third, fourth, fifth or more subset group of oligonucleotides, thereby producing the enriched oligonucleotide probe library. In a related aspect, the invention provides a plurality of oligonucleotides comprising at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10000, 20000, 30000, 40000, 50000, 60000, 70000, 80000, 90000, 100000, 200000, 300000, 400000, or 500000 different oligonucleotide sequences, wherein the plurality results from the method in this paragraph, wherein the library is capable of distinguishing a first phenotype from a second phenotype. In some embodiments, the first phenotype comprises a disease or disorder and the second phenotype comprises a healthy state; or wherein the first phenotype comprises a disease or disorder and the second phenotype comprises a different disease or disorder; or wherein the first phenotype comprises a stage or progression of a disease or disorder and the second phenotype comprises a different stage or progression of the same disease or disorder; or wherein
the first phenotype comprises a positive response to a therapy and the second phenotype comprises a negative response to the same therapy.

[00434] In yet another aspect, the invention provides a method of characterizing a disease or disorder, comprising: a) contacting a biological test sample with the oligonucleotide or plurality of oligonucleotides provided by the invention; b) detecting a presence or level of complexes formed in step (a) between the oligonucleotide or plurality of oligonucleotides provided by the invention and a target in the biological test sample; and c) comparing the presence or level detected in step (b) to a reference level from a biological control sample, thereby characterizing the disease or disorder. The step of detecting may comprise performing sequencing of all or some of the oligonucleotides in the complexes, amplification of all or some of the oligonucleotides in the complexes, and/or hybridization of all or some of the oligonucleotides in the complexes to an array. The sequencing may be high-throughput or next generation sequencing.

[00435] In the methods of the invention, the biological test sample and biological control sample may each comprise a tissue sample, a cell culture, or a biological fluid. In some embodiments, the biological fluid comprises a bodily fluid. Useful bodily fluids within the method of the invention comprise peripheral blood, sera, plasma, ascites, urine, cerebrospinal fluid (CSF), sputum, saliva, bone marrow, synovial fluid, aqueous humor, amniotic fluid, cerumen, breast milk, bronchialveolar 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 water, pancreatic juice, lavage fluids from sinus cavities, bronchopulmonary aspirates, blastocyl cavity fluid, or umbilical cord blood. In some preferred embodiments, the bodily fluid comprises blood, serum or plasma. The biological fluid may comprise microvesicles. In such case, the complexes may be formed between the oligonucleotide or plurality of oligonucleotides and at least one of the microvesicles.

[00436] The biological test sample and biological control sample may further comprise isolated microvesicles, wherein optionally the microvesicles are isolated using at least one of chromatography, filtration, ultrafiltration, centrifugation, ultracentrifugation, flow cytometry, affinity capture (e.g., to a planar surface, column or bead), polymer precipitation, and using microfluidics. The vesicles can also be isolated after contact with the oligonucleotide or plurality of oligonucleotides.

[00437] In various embodiments of the methods of the invention, the oligonucleotide or plurality of oligonucleotides binds a polypeptide or fragment thereof. The polypeptide or fragment thereof can be soluble or membrane bound, wherein optionally the membrane comprises a microvesicle membrane. The membrane could also be from a cell or a fragment of a cell of vesicle. In some embodiments, the polypeptide or fragment thereof comprises a biomarker in Table 3, Table 4 or any one of Tables 18-25. For example, the polypeptide or fragment thereof could be a general vesicle marker such as in Table 3 or a tissue-related or disease-related marker such as in Table 4, or a vesicle associated biomarker provided in any one of Tables 18-25. The oligonucleotide or plurality of oligonucleotides may bind a microvesicle.
surface antigen in the biological sample. For example, the oligonucleotide or plurality of oligonucleotides can be enriched from a naive library against microvesicles.

[00438] As noted above, the microvesicles may be isolated in whole or in part using polymer precipitation. In an embodiment, the polymer comprises polyethylene glycol (PEG). Any appropriate form of PEG may be used. For example, the PEG may be PEG 8000. The PEG may be used at any appropriate concentration. For example, the PEG can be used at a concentration of 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14% or 15% to isolate the microvesicles. In some embodiments, the PEG is used at a concentration of 6%.

[00439] The invention provides oligonucleotide probes that can be used to carry out the methods herein. See, e.g., Examples 26-32. In an aspect, the invention provides an oligonucleotide comprising a sequence according to any one of SEQ ID NOs 137-969 and 1072-4150. In a related aspect, the invention provides an oligonucleotide comprising a sequence according to any one of the SEQ ID NOs in Table 40. In another related aspect, the invention provides an oligonucleotide comprising a sequence according to any one of the SEQ ID NOs in the row "2000v1" in Table 43. In still another related aspect, the invention provides an oligonucleotide comprising a sequence according to any one of the SEQ ID NOs in the row "2000v2" in Table 43. In yet another related aspect, the invention provides an oligonucleotide comprising a sequence according to any one of the SEQ ID NOs in the row "Common" in Table 43.

[00440] The oligonucleotides of the invention can comprise flanking regions for various purposes, including without limitation amplification, capture, conjugation or spacing. For example, the invention provides an oligonucleotide comprising a sequence according to any one of the SEQ ID NOs above and further having a 5' region with sequence 5'-$C_{TAGCAGACTGAGTAC}$ (SEQ ID NO. 131) and/or a 3' region with sequence 5'-$CTGCTCTTTATACACATCTGACTGCACGAG$SEQ ID NO. 132).

[00441] The invention further provides oligonucleotides homologous to the SEQ ID NOs above. For example, the invention provides an oligonucleotide comprising a nucleic acid sequence or a portion thereof that is at least 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 96, 97, 98, 99 or 100 percent homologous to an oligonucleotide sequence of any one of the SEQ ID NOs above. The homologous sequences may comprise similar properties to the listed sequences, such as similar binding properties.

[00442] In an aspect, the invention provides a plurality of oligonucleotides comprising at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, or at least 10000 different oligonucleotide sequences as described in the paragraphs above. For example, the invention provides a plurality of oligonucleotides comprising member sequences having at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1500, 2000, 2500, 3000, 3500, 4000, or all variable regions according to SEQ ID NOs 137-969 and 1072-4150.

[00443] The plurality of oligonucleotides can comprise at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 40, 45, 50, 55, 60, 65, 70, 75, or all SEQ ID NOs listed in Table 40. In an
embodiment, the plurality of oligonucleotides comprises at least the first 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 40, 45, 50, 55, 60, 65, 70, 75, or SEQ ID NOs listed in Table 40.

[00444] The plurality of oligonucleotides can also comprise at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 200, 225, 250, 275, 300, or all SEQ ID NOs listed in row "2000vl" of Table 43. In an embodiment, the plurality of oligonucleotides comprises at least the first 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 200, 225, 250, 275, 300, or all SEQ ID NOs listed in row "2000vl" of Table 43.

[00445] The plurality of oligonucleotides can comprise at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145 or all SEQ ID NOs listed in row "2000v2" of Table 43. In an embodiment, the plurality of oligonucleotides comprises at least the first 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145 or all SEQ ID NOs listed in row "2000v2" of Table 43.

[00446] The plurality of oligonucleotides can also comprise at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17 or all variable regions listed in row "Common" of Table 43. In an embodiment, the plurality of oligonucleotides comprises at least the first 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17 or all SEQ ID NOs listed in row "Common" of Table 43.

[00447] The oligonucleotide or at least one member of the plurality of oligonucleotides can have least one functional modification selected from the group consisting of DNA, RNA, biotinylation, a non-naturally occurring nucleotides, a deletion, an insertion, an addition, and a chemical modification. Such modifications may provide additional or altered functions to the oligonucleotides, including without limitation capture, detection, stability, or binding properties.

[00448] Such oligonucleotides and plurality of oligonucleotides (pools) can be used to characterize a phenotype as described herein. In an aspect, the invention provides a method of characterizing a phenotype in a sample comprising: (a) contacting the sample with at least one oligonucleotide or plurality of oligonucleotides provided by the invention (see above); and (b) identifying a presence or level of a complex formed between the at least one oligonucleotide or plurality of oligonucleotides and the sample, wherein the presence or level is used to characterize the phenotype. Any useful technique for identifying can be used according to the invention. In various embodiments, the identifying comprises sequencing, amplification, hybridization, gel electrophoresis or chromatography. In an embodiment, identifying by hybridization comprises contacting the sample with at least one labeled probe that is configured to hybridize with at least one oligonucleotide. The at least one labeled probe can be directly or indirectly attached to a label. Any useful label can be used, including without limitation a fluorescent or magnetic label. In another embodiment, identifying by sequencing comprises next generation sequencing, dye termination sequencing, and/or pyrosequencing.

[00449] In the methods of the invention, the complex formed between the at least one oligonucleotide or the plurality of oligonucleotides and the sample can be a complex formed between a microvesicle
population in the sample and the at least one oligonucleotide or plurality of oligonucleotides. The microvesicle population can be isolated in whole or in part from other constituents in the sample before of after the contacting. In embodiments, the isolating uses affinity purification, filtration, polymer precipitation, PEG precipitation, ultracentrifugation, a molecular crowding reagent, affinity isolation, affinity selection, or any combination thereof.

[00450] In the methods of the invention, the phenotype can be any detectable phenotype. In some embodiments, the phenotype comprises a disease or disorder. In such cases, the characterizing can be a diagnosis, prognosis and/or theranosis for the disease or disorder. The theranosis can be any type of therapy-related such as described herein. The theranosis includes without limitation predicting a treatment efficacy or lack thereof, or monitoring a treatment efficacy.

[00451] The characterizing step of the methods of the invention may entail comparing the presence or level to a reference. Any useful reference can be used. In an embodiment wherein the phenotype comprises a disease or disorder, the reference can be the presence or level determined in a sample from an individual without a disease or disorder, or from an individual with a different state of the disease or disorder. In some embodiments, the comparison to the reference of at least one oligonucleotide comprising a sequence having a SEQ ID NO provided above indicates that the sample comprises a cancer sample or a non-cancer/normal sample.

[00452] The disease or disorder detected by the oligonucleotide, plurality of oligonucleotides, or methods provided here may comprise any appropriate disease or disorder of interest, including without limitation Breast Cancer, Alzheimer’s disease, bronchial asthma, Transitional cell carcinoma of the bladder, Giant cellular osteoblastoclastoma, Brain Tumor, Colorectal adenocarcinoma, Chronic obstructive pulmonary disease (COPD), Squamous cell carcinoma of the cervix, acute myocardial infarction (AMI) / acute heart failure, Chron’s Disease, diabetes mellitus type II, Esophageal carcinoma, Squamous cell carcinoma of the larynx, Acute and chronic leukemia of the bone marrow, Lung carcinoma, Malignant lymphoma, Multiple Sclerosis, Ovarian carcinoma, Parkinson disease, Prostate adenocarcinoma, psoriasis, Rheumatoid Arthritis, Renal cell carcinoma, Squamous cell carcinoma of skin, Adenocarcinoma of the stomach, carcinoma of the thyroid gland, Testicular cancer, ulcerative colitis, or Uterine adenocarcinoma.

[00453] In some embodiments, the disease or disorder comprises a cancer, a premalignant condition, an inflammatory disease, an immune disease, an autoimmune disease or disorder, a cardiovascular disease or disorder, neurological disease or disorder, infectious disease or pain. The cancer can include without limitation one of 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; carcinoind 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; cranioopharyngioma; 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; lung 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; Sezary 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; Waldenstrom macroglobulinemia; or Wilm's tumor. The premalignant condition can include without limitation Barrett's Esophagus. The autoimmune disease can include without limitation one of 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 erythematosus (SLE), Hashimoto's Thyroiditis, Grave's disease, Ankylosing Spondylitis Sjogrens Disease, CREST syndrome, Scleroderma, Rheumatic Disease, organ rejection, Primary Sclerosing Cholangitis, or sepsis. The cardiovascular disease can include without limitation one of atherosclerosis, congestive heart failure, vulnerable plaque, stroke, ischemia, high blood pressure, stenosis, vessel occlusion or a thrombotic event. The neurological disease can include without limitation one of Multiple Sclerosis (MS), Parkinson's
Disease (PD), Alzheimer’s Disease (AD), schizophrenia, bipolar disorder, depression, autism, Prion Disease, Pick's disease, dementia, Huntington disease (HD), Down's syndrome, cerebrovascular disease, Rasmussen's encephalitis, viral meningitis, neuropsychiatric systemic lupus erythematosus (NPSLE), amyotrophic lateral sclerosis, Creutzfeldt-Jacob disease, Gerstmann-Straussler-Scheinker disease, transmissible spongiform encephalopathy, ischemic reperfusion damage (e.g. stroke), brain trauma, microbial infection, or chronic fatigue syndrome. The pain can include without limitation one of fibromyalgia, chronic neuropathic pain, or peripheral neuropathic pain. The infectious disease can include without limitation one of a bacterial infection, viral infection, yeast infection, Whipple's Disease, Prion Disease, cirrhosis, methicillin-resistant staphylococcus aureus, HIV, HCV, hepatitis, syphilis, meningitis, malaria, tuberculosis, or influenza. One of skill will appreciate that the oligonucleotide or plurality of oligonucleotides or methods of the invention can be used to assess any number of these or other related diseases and disorders.

[00454] In some embodiments of the invention, the oligonucleotide or plurality of oligonucleotides and methods of use thereof are useful for characterizing certain diseases or disease states. As desired, a pool of oligonucleotides useful for characterizing various diseases is assembled to create a master pool that can be used to probe useful for characterizing the various diseases. One of skill will also appreciate that pools of oligonucleotides useful for characterizing specific diseases or disorders can be created as well. The sequences provided herein can also be modified as desired so long as the functional aspects are still maintained (e.g., binding to various targets or ability to characterize a phenotype). For example, the oligonucleotides may comprise DNA or RNA, incorporate various non-natural nucleotides, incorporate other chemical modifications, or comprise various deletions or insertions. Such modifications may facilitate synthesis, stability, delivery, labeling, etc, or may have little to no effect in practice. In some cases, some nucleotides in an oligonucleotide may be substituted while maintaining functional aspects of the oligonucleotide. Similarly, 5’ and 3' flanking regions may be substituted. In still other cases, only a portion of an oligonucleotide may be determined to direct its functionality such that other portions can be deleted or substituted. Numerous techniques to synthesize and modify nucleotides and polynucleotides are disclosed herein or are known in the art.

[00455] In an aspect, the invention provides a kit comprising a reagent for carrying out the methods of the invention provided herein. In a similar aspect, the invention contemplates use of a reagent for carrying out the methods of the invention provided herein. In embodiments, the reagent comprises an oligonucleotide or plurality of oligonucleotides. The oligonucleotide or plurality of oligonucleotides can be those provided herein. The reagent may comprise various other useful components including without limitation microRNA (miR) and messenger RNA (mRNA), a protein-nucleic acid complex, and various combinations, fragments and/or complexes of any of these. The one or more of: a) a reagent configured to isolate a microvesicle, optionally wherein the at least one reagent configured to isolate a microvesicle comprises a binding agent to a microvesicle antigen, a column, a substrate, a filtration unit, a polymer, polyethylene glycol, PEG4000, PEG8000, a particle or a bead; b) at least one oligonucleotide configured to act as a primer or probe in order to amplify, sequence, hybridize or detect the oligonucleotide or
plurality of oligonucleotides; and c) a reagent configured to remove one or more abundant protein from a sample, wherein optionally the one or more abundant protein comprises at least one of albumin, immunoglobulin, fibrinogen and fibrin.

[00456] **Recovery of oligonucleotide probes post-probing**

[00457] As described herein, the oligonucleotide probes of the invention can be used to probe a sample in order to characterize a phenotype. The methods may entail recovering the oligonucleotide probes that bound various biological entities in the sample in order to identify the bound probes. In an aspect, the invention provides a method of detecting at least one oligonucleotide in a sample, comprising: (a) providing the at least one oligonucleotide comprising a capture moiety; (b) contacting the sample with the at least one oligonucleotide provided in (a); (c) capturing the at least one oligonucleotide that formed a complex with a component in the sample in (b); and (d) identifying a presence or level of the at least one oligonucleotide captured in (c), wherein optionally the presence or level is used to characterize a phenotype. See, e.g., **Example 33** and **FIGs. 17A-E**. The at least one oligonucleotide may be captured to a substrate, including without limitation a bead or planar substrate. The capture moiety can be any useful capture moiety, including without limitation a biotin moiety. The capture moiety can be cleavable, e.g., photocleavable or chemically cleavable. In an embodiment, the at least one oligonucleotide is captured to a substrate coupled to avidin or streptavidin. Such configuration is particularly useful when the capture moiety comprises a biotin moiety. In some embodiments, the captured at least one oligonucleotide is released from the substrate by irradiation prior to the identifying. Any useful irradiation, e.g., ultra violet (UV) light may be used. Any useful technique for identifying can be used according to the invention. In various embodiments, the identifying comprises sequencing, amplification, hybridization, gel electrophoresis or chromatography. In an embodiment, identifying by hybridization comprises contacting the sample with at least one labeled probe that is configured to hybridize with at least one oligonucleotide. The at least one labeled probe can be directly or indirectly attached to a label. Any useful label can be used, including without limitation a fluorescent or magnetic label. In another embodiment, identifying by sequencing comprises next generation sequencing, dye termination sequencing, and/or pyrosequencing. The at least one oligonucleotide can be an oligonucleotide or plurality of oligonucleotides provided by the invention. See e.g., the oligonucleotides and plurality of oligonucleotides described above.

[00458] **Single strand DNA (ssDNA) library preparation**

[00459] In an embodiment, the invention provides a nucleic acid molecule comprising a 5′ leader region which is 5′ of a variable region, which is 5′ of a tail region, wherein the leader region comprises a lengthener region, a terminator region and a forward primer region, and the tail region comprises a reverse primer region. The nucleic acid molecule may be used for asymmetric or unequal length PCR applications as desired, e.g., to recover ssDNA. See, e.g., **Example 34** and **FIGs. 18A-C**. The lengthener region can be any desired length. In some embodiments, the lengthener region comprises at least 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 nucleotides. The lengthener region may comprise a poly-A sequence. Similarly, the terminator region can be any desired length. In some embodiments, the terminator region
comprises at least 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 30, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 nucleotides. The terminator region may comprise a non-nucleotide terminator. For example, the non-nucleotide terminator can be a polymer such as triethylene glycol or the like. The forward primer region can be any desired length. In some embodiments, the forward primer region comprises at least 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 30, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 nucleotides. The variable region can be any desired length. In some embodiments, the variable region comprises at least 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 30, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 nucleotides. In some embodiments, the variable region binds a target molecule or complex through non-Watson-Crick base pairing. For example, the variable region may act as an aptamer and bind proteins or other entities. Finally, the reverse primer region can be any desired length. In some embodiments, the reverse primer region comprises at least 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 30, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 nucleotides. 

[00460] In a related aspect, the invention provides a method of generating a single-stranded DNA (ssDNA) molecule comprising: a) providing a mixture comprising a nucleic acid molecule as described in the paragraph above, and forward and reverse primers configured to amplify the nucleic acid molecule from the forward primer region and reverse primer region, respectively; and b) performing asymmetric polymerase chain reaction (PCR) on the mixture in a) to favorably amplify the reverse strand of the nucleic acid molecule, wherein the forward and reverse primers in the mixture are at a ratio of at least about 1:5 (F/R) in favor of the reverse primers; thereby generating the ssDNA molecule. In an embodiment, the ratio is between about 1:20-1:50 (F/R) in favor of the reverse primers. For example, the ratio can be between about 1:37.5 (F/R) in favor of the reverse primers. The method may further comprise isolating the amplified reverse strand of the nucleic acid molecule on a native gel. The method may also further comprise: c) denaturing the amplified nucleic acid molecules from b); and d) isolating the denatured reverse strand of the nucleic acid molecules from c). In an embodiment, the denatured reverse strand of the nucleic acid molecules is isolated on a denaturing gel. The mixture in a) can comprise additional components as desired. For example, the mixture may further comprise at least one of an enrichment buffer, non-target molecules, proteins, microvesicles, and polyethylene glycol. 

[00461] In a related aspect, the invention provides a kit comprising a reagent for carrying out the methods herein. In still another related aspect, the invention provides for use of a reagent for carrying out the methods. In various embodiments, the reagent comprises at least one of a buffer, a nucleic acid molecule described above, and forward and/or reverse primers configured to amplify the nucleic acid molecule. 

[00462] Detecting Watson-Crick base pairing with an oligonucleotide probe 

[00463] The oligonucleotide probes provided by the invention can bind via non-Watson Crick base pairing. However, in some cases, the oligonucleotide probes provided by the invention can bind via Watson Crick base pairing. The oligonucleotide probe libraries of the invention, e.g., as described above, can query both types of binding events simultaneously. For example, some oligonucleotide probes may
bind the microvesicle protein antigens in the classical aptamer sense, whereas other oligonucleotide probes may bind microvesicles via nucleic acids associated with the microvesicles, e.g., nucleic acid (including without limitation microRNA and mRNA) on the surface of the microvesicles or as payload. Such surface bound nucleic acids can be associated with proteins. For example, they may comprise Argonaute-microRNA complexes. The argonaute protein can be Agol, Ago2, Ago3 and/or Ago4.

[00464] In addition to the oligonucleotide probe library approach described herein which relies on determining a sequence of the oligonucleotides (e.g., via sequencing, hybridization or amplification), assays can also be designed to detect Watson Crick base pairing. In some embodiments, these approaches rely on Ago2-mediated cleavage wherein an Ago2-microRNA complex can be used to detected using oligonucleotide probes. These approaches can use oligonucleotide probes to detect Ago2-microRNA complexes, which may or may not be associated with microvesicles. The detection can be used to characterize a phenotype as described herein. Such detection can be used along side the sequencing identification methods described herein (e.g., via sequencing, hybridization or amplification). For further details, see PCT/US 15/62184, filed November 23, 2015, which application is incorporated by reference herein in its entirety.

Therapeutics

[00465] As used herein "therapeutically effective amount" refers to an amount of a composition that relieves (to some extent, as judged by a skilled medical practitioner) one or more symptoms of the disease or condition in a mammal. Additionally, by "therapeutically effective amount" of a composition is meant an amount that returns to normal, either partially or completely, physiological or biochemical parameters associated with or causative of a disease or condition. A clinician skilled in the art can determine the therapeutically effective amount of a composition in order to treat or prevent a particular disease condition, or disorder when it is administered, such as intravenously, subcutaneously, intraperitoneally, orally, or through inhalation. The precise amount of the composition required to be therapeutically effective will depend upon numerous factors, e.g., such as the specific activity of the active agent, the delivery device employed, physical characteristics of the agent, purpose for the administration, in addition to many patient specific considerations. But a determination of a therapeutically effective amount is within the skill of an ordinarily skilled clinician upon the appreciation of the disclosure set forth herein.

[00466] The terms "treating," "treatment," "therapy," and "therapeutic treatment" as used herein refer to curative therapy, prophylactic therapy, or preventative therapy. An example of "preventative therapy" is the prevention or lessening the chance of a targeted disease (e.g., cancer or other proliferative disease) or related condition thereto. Those in need of treatment include those already with the disease or condition as well as those prone to have the disease or condition to be prevented. The terms "treating," "treatment," "therapy," and "therapeutic treatment" as used herein also describe the management and care of a mammal for the purpose of combating a disease, or related condition, and includes the administration of a composition to alleviate the symptoms, side effects, or other complications of the disease, condition. Therapeutic treatment for cancer includes, but is not limited to, surgery, chemotherapy, radiation therapy, gene therapy, and immunotherapy.
As used herein, the term "agent" or "drug" or "therapeutic agent" refers to a chemical compound, a mixture of chemical compounds, a biological macromolecule, or an extract made from biological materials such as bacteria, plants, fungi, or animal (particularly mammalian) cells or tissues that are suspected of having therapeutic properties. The agent or drug can be purified, substantially purified or partially purified. An "agent" according to the present invention, also includes a radiation therapy agent or a "chemotherapeutic agent."

As used herein, the term "diagnostic agent" refers to any chemical used in the imaging of diseased tissue, such as, e.g., a tumor.

As used herein, the term "chemotherapeutic agent" refers to an agent with activity against cancer, neoplastic, and/or proliferative diseases, or that has ability to kill cancerous cells directly.

As used herein, "pharmaceutical formulations" include formulations for human and veterinary use with no significant adverse toxicological effect. "Pharmaceutically acceptable formulation" as used herein refers to a composition or formulation that allows for the effective distribution of the nucleic acid molecules of the instant invention in the physical location most suitable for their desired activity.

As used herein the term "pharmaceutically acceptable carrier" is intended to include any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. The use of such media and agents for pharmaceutically active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active compound, use thereof in the compositions is contemplated.

**Aptamer-Toxin Conjugates as a Cancer Therapeutic**

Previous work has developed the concept of antibody-toxin conjugates ("immunoconjugates") as potential therapies for a range of indications, mostly directed at the treatment of cancer with a primary focus on hematological tumors. A variety of different payloads for targeted delivery have been tested in pre-clinical and clinical studies, including protein toxins, high potency small molecule cytotoxics, radioisotopes, and liposome-encapsulated drugs. While these efforts have successfully yielded three FDA-approved therapies for hematological tumors, immunoconjugates as a class (especially for solid tumors) have historically yielded disappointing results that have been attributable to multiple different properties of antibodies, including tendencies to develop neutralizing antibody responses to non-humanized antibodies, limited penetration in solid tumors, loss of target binding affinity as a result of toxin conjugation, and imbalances between antibody half-life and toxin conjugate half-life that limit the overall therapeutic index (reviewed by Reff and Heard, Critical Reviews in Oncology/Hematology, 40 (2001):25-35).

Aptamers are functionally similar to antibodies, except their absorption, distribution, metabolism, and excretion ("ADME") properties are intrinsically different and they generally lack many of the immune effector functions generally associated with antibodies (e.g., antibody-dependent cellular cytotoxicity, complement-dependent cytotoxicity). In comparing many of the properties of aptamers and antibodies previously described, several factors suggest that toxin-delivery via aptamers offers several concrete advantages over delivery with antibodies, ultimately affording them better potential as
therapeutics. Several examples of the advantages of toxin-delivery via aptamers over antibodies are as follows:

[00475] 1) Aptamer-toxin conjugates are entirely chemically synthesized. Chemical synthesis provides more control over the nature of the conjugate. For example, the stoichiometry (ratio of toxins per aptamer) and site of attachment can be precisely defined. Different linker chemistries can be readily tested. The reversibility of aptamer folding means that loss of activity during conjugation is unlikely and provides more flexibility in adjusting conjugation conditions to maximize yields.


[00477] 3) Tunable PK. Aptamer half-life/metabolism can be easily tuned to match properties of payload, optimizing the ability to deliver toxin to the tumor while minimizing systemic exposure. Appropriate modifications to the aptamer backbone and addition of high molecular weight PEGs should make it possible to match the half-life of the aptamer to the intrinsic half-life of the conjugated toxin/linker, minimizing systemic exposure to non-functional toxin-bearing metabolites (expected if ti/2(aptamer)<<ti/2(toxin)) and reducing the likelihood that persisting unconjugated aptamer will functionally block uptake of conjugated aptamer (expected if ti/2(aptamer)>>ti/2(toxin)).

[00478] 4) Relatively low material requirements. It is likely that dosing levels will be limited by toxicity intrinsic to the cytotoxic payload. As such, a single course of treatment will likely entail relatively small (<100 mg) quantities of aptamer, reducing the likelihood that the cost of oligonucleotide synthesis will be a barrier for aptamer-based therapies.

[00479] 5) Parenteral administration is preferred for this indication. There will be no special need to develop alternative formulations to drive patient/physician acceptance.

[00480] The invention provides a pharmaceutical composition comprising a therapeutically effective amount of an aptamer provided by the invention or a salt thereof, and a pharmaceutically acceptable carrier or diluent. The invention also provides a pharmaceutical composition comprising a therapeutically effective amount of the aptamer or a salt thereof, and a pharmaceutically acceptable carrier or diluent. Relatedly, the invention provides a method of treating or ameliorating a disease or disorder, comprising administering the pharmaceutical composition to a subject in need thereof. Administering a therapeutically effective amount of the composition to the subject may result in: (a) an enhancement of the delivery of the active agent to a disease site relative to delivery of the active agent alone; or (b) an enhancement of microvesicles clearance resulting in a decrease of at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% in a blood level of microvesicles targeted by the aptamer; or (c) an decrease in
biological activity of microvesicles targeted by the aptamer of at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90%. In an embodiment, the biological activity of microvesicles comprises immune suppression or transfer of genetic information. The disease or disorder can include without limitation those disclosed herein. For example, the disease or disorder may comprise a neoplastic, proliferative, or inflammatory, metabolic, cardiovascular, or neurological disease or disorder. See, e.g., section "Phenotypes."

[00481] Anti-target and multivalent oligonucleotides

[00482] As noted above, the target of oligonucleotide probes can be identified. For example, when the target comprises a protein or protein complex (e.g., a nucleoprotein or lipoprotein), identifying the target may comprise use of mass spectrometry (MS), peptide mass fingerprinting (PMF; protein fingerprinting), sequencing, N-terminal amino acid analysis, C-terminal amino acid analysis, Edman degradation, chromatography, electrophoresis, two-dimensional gel electrophoresis (2D gel), antibody array, or immunoassay. Such approaches can be applied to identify a number of targets recognized by an oligonucleotide probe library. For example, an oligonucleotide probe library can be incubated with a sample of interest, bound members of the library captured, and the targets bound to the captured members identified. See Example 23 herein for an example of such target identification using mass spectrometry.

[00483] The oligonucleotide aptamers to the various targets can be used for multiple purposes. In some embodiments, the aptamers are used as therapeutic agents. Immunotherapeutic approaches using antibodies that recognize foreign/misfolded antigens (e.g., anti-CD20, anti-CD30, anti-CD33, anti-CD52, anti-EGFR, anti-nucleolin, anti-nucleophosmin, etc.) can selectively kill target cells via linked therapeutic agents or by stimulating the immune system through activation of cell-mediated cytotoxicity. Aptamers or oligonucleotides are an attractive immunotherapeutic alternative for various reasons such as low cost, small size, ease and speed of synthesis, stability and low immunogenicity. In an embodiment, immunotherapeutic agents are conjugated to disease specific target oligonucleotide or antibody (Ab) for targeted cell killing via recruitment of complement proteins and the downstream membrane attack complex. See, e.g., Zhou and Rossi, Cell-type-specific, Aptamer-functionalyzed Agents for Targeted Disease Therapy, Mol Ther Nucleic Acids. 2014 Jun 17;3:el69. doi: 10.1038/mtna.2014.21; Pei et al., Clinical applications of nucleic acid aptamers in cancer, Mol Clin Oncol. 2014 May;2(3):341-348. Epub 2014 Feb 10. This approach can be applied to target diseased host cells such as cancer cells, gram negative bacteria, viral and/or parasitic infections, and the like.

[00484] In some embodiments, the invention provides a multipartite construct comprising a binding agent specific to a biological target with another binding agent specific to immunomodulatory entity. Examples of such constructs are shown in FIG. 23A. In Design 1 in the figure, the horizontal line indicates an oligonucleotide construct, which construct comprises a 5’ primer 2301 (Primer 1), a variable region 2302 that can be an aptamer to a target of interest, a 3’ primer 2303 (Primer 2), and an immunomodulatory domain region ("IMD") 2304. The complete Design 1 construct can be used to bring a target of interest in proximity with an immunomodulatory agent. The primers can be designed for any desired purpose, e.g., amplification, capture, modification, direct or indirect labeling, and the like. In some embodiments, the
target of the variable region is a disease marker and thus the construct is targeted to a disease cell or microvesicle. The immunomodulatory domain region can act as an immune stimulator or suppressor. Any appropriate immune stimulator or suppressor can be used, e.g., a small molecule, antibody or an aptamer. Thus, the construct can modulate the immune response at a target of interest, e.g., at a cell or microvesicle carrying the target. The basic construct can be modified as desired. For example, Design 2 in FIG. 23A shows the construct carrying a linker 2305 between Primer 2 2303 and the IMD 2304. Such linkers are explained further below and can be inserted between any components of the construct as desired. Linkers can provide a desired space between the regions of the construct and can be manipulated to influence other properties such as stability. Design 3 in FIG. 23A shows another example wherein the IMD 2304 is an oligonucleotide and the variable region 2302 and IMD 2304 lie between the primers 2301 and 2303.

One of skill will appreciate that one or more linker, such as 2305 of Design 2, can also be inserted into Design 3, e.g., between the variable region 2302 and IMD 2304. One of skill will further appreciate that the ordering of the oligonucleotide segments from 5’ to 3’ can be modified, e.g., reversed. As a concrete example which will be described further below, FIG. 23B illustrates Design 1 and Design 2 from FIG. 23A wherein the variable region comprises an anti-CD20 oligonucleotide 2311 and the IMD comprises an anti-C1q oligonucleotide 2312, e.g., an oligonucleotide provided herein. See, e.g., Example 35. This constructs of FIG. 23B can be used used to target a CD20+ cell population and stimulate C1q mediated cell killing.

[00485] As noted, the multipartite constructs may be synthesized and/or modified as desired. In some embodiments of the invention, the multipartite oligonucleotide construct is synthesized directly with or without a linker in between the oligonucleotide segments. See, e.g., FIG. 23A Design 3, which can be generated directly via amplification by Primer 1 2301 and Primer 2 2303. One or more linker can act as a spacer to create a desired spacing between the target of the variable region segment 2302 and the target of the IMD segment 2304. The spacing can be determined via computer modeling or via experimentation due to steric hindrance or other considerations. Following the example of FIG. 23B, the type and size of the linker may be dependent upon steric hindrance between the CD20 target protein and the Clq protein/MAC complex.

[00486] The multipartite constructs can be generated against any appropriate target. The targets can include without limitation diseased cells, cancer cells, circulating tumor cells (CTCs), immune cells (e.g., B-cells, T-cells, macrophages, dendritic cells), microvesicles, bacteria, viruses or other parasites. The target can be large biological complexes, e.g., protein complexes, ribonucleoprotein complexes, lipid complexes, or a combination thereof. It will be understood that the specific target of the multipartite constructs can be a certain member of the foregoing macromolecular targets. For example, consider that the desired target of the multipartite construct is a cell or microvesicle. In such case, the multipartite construct can be directed to a specific biomarker, e.g., a surface antigen, of the cell or microvesicle. As a non-limiting example, the target of interest can be B-cells and the specific target of the variable region of the multipartite construct can be CD20. CD20 is a cellular marker of B-cells targeted by the monoclonal antibodies (mAb) rituximab, obinutuzumab, ofatumumab, ibritumomab tiuxetan, and tositumomab, which
are used as agents in the treatment of B-cell lymphomas and leukemias. As another non-limiting example, the target of interest can be cancer cells and the specific target of the variable region of the multipartite construct can be c-MET. MET is a membrane receptor that is essential for embryonic development and wound healing. Abnormal MET activation in cancer correlates with poor prognosis, where aberrantly active MET triggers tumor growth, formation of new blood vessels (angiogenesis), and cancer spread to other organs (metastasis). MET has been observed to be deregulated in many types of human malignancies, including cancers of kidney, liver, stomach, breast, and brain. See FIG. 23D for illustration (discussed further below). Other biomarkers can be used as the specific target as desired. For example, the biomarker can be selected from Table 3, Table 4, Tables 18-26, or Table 45 herein. See FIG. 23C, which illustrates a construct of the invention 2331 having a segment that recognizes a biomarker 2332 ("Marker of Interest") on a cell or vesicle surface 2333 ("Membrane"), and another segment 2334 that attracts an immune response ("Complement"). The construct 2331 can be such as in FIGs. 23A-B or any other desired configuration. Binding of such a construct to a target can cause a complement cascade and induce apoptosis.

[00487] In some embodiments of the invention, the target biomarker is selected from the group consisting of CD19, CD20, CD21, CD22 (also known as LL2), CDIM, and Lym-1. The target biomarker can be a membrane associated protein. In embodiments, the membrane associated protein is selected from the group consisting of CD4, CD19, DC-SIGN/CD209, HIV envelope glycoprotein gp120, CCR5, EGFR/ErbBl, EGFR2/ErbB2/HER2, EGFR3/ErbB3, EGFR4/ErbB4, EGFRvIII, Transferrin Receptor, PSMA, VEGF, VEGF-2, CD25, CDlla, CD33, CD20, CD3, CD52, CEA, TAG-72, LDL receptor, insulin receptor, megalin receptor, LRP, mannose receptor, P63/CKAP4 receptor, arrestin, ASGP, CCK-B, HGFR, RON receptor, FGFR, ILR, AFP, CA125/MUC16, PDGFR, stem cell factor receptor, colony stimulating factor-1 receptor, integrins, TLR, BCR and BAFF-R. The target biomarker can also be a cellular receptor selected from the group consisting of: nucleolin, human epidermal growth factor receptor 2 (HER2), CD20, a transferrin receptor, an asialoglycoprotein receptor, a thyroid-stimulating hormone (TSH) receptor, a fibroblast growth factor (FGF) receptor, CD3, the interleukin 2 (IL-2) receptor, a growth hormone receptor, an insulin receptor, an acetylcholine receptor, an adrenergic receptor, a vascular endothelial growth factor (VEGF) receptor, a protein channel, cadherin, a desmosome, and a viral receptor. In various embodiments, the target biomarker is a cell surface molecule selected from the group consisting of IgM, IgD, IgG, IgA, IgE, CD 19, CD20, CD21, CD22, CD24, CD40, CD72, CD79a, CD79b, CDld, CD5, CD9, CD10, CDld, CD23, CD27, CD38, CD48, CD80, CD86, CD138, CD148; and combinations thereof. The target biomarker can be a lymphocyte-directing target such as one or more T-cell receptor motifs, T-cell α chains, T-cell β chains, T-cell γ chains, T-cell Δ chains, CCR7, CD3, CD4, CD5, CD7, CD8, CDldb, CDlle, CD16, CD19, CD20, CD21, CD22, CD25, CD28, CD34, CD35, CD40, CD45RA, CD45RO, CD52, CD56, CD62L, CD68, CD80, CD95, CD117, CD127, CD133, CD137 (4-1 BB), CD163, F4/80, IL-4Ra, Sca-1, CTLA-4, GITR, GARP, LAP, granzyme B, LFA-1, or transferrin receptor.
In some embodiments, the target biomarker comprises a growth factor, vascular endothelial growth factor (VEGF), TGF, TGFβ, PDGF, IGF, FGF, cytokine, lymphokine, hematopoietic factor, M-CSF, GM-CSF, TNF, interleukin, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IL-16, IL-17, IL18, IFN, TNFO, TNF1, TNF2, G-CSF, Meg-CSF, GM-CSF, thrombopoietin, stem cell factor, erythropoietin, hepatocyte growth factor/NK1, angiogenic factor, angiopoietin, Ang-1, Ang-2, Ang-4, Ang-Y, human angiopoietin-like polypeptide, angiogenin, morphogenic protein-1, bone morphogenic protein receptor, bone morphogenic protein receptor IA, bone morphogenic protein receptor IB, neurotrophic factor, chemotactic factor, CD proteins, CD3, CD4, CD8, CD19, CD20, erythropoietin, osteoinductive factors, immunotoxin, bone morphogenetic protein (BMP), interferon, interferon-alpha, interferon-beta, interferon-gamma, colony stimulating factor (CSF), M-CSF, GM-CSF, G-CSF, superoxide dismutase, T-cell receptor; surface membrane protein, decay accelerating factor, viral antigen, portion of the AIDS envelope, transport protein, homing receptor, addressin, regulatory protein, integrin, CD11a, CD11b, CD11c, CD18, ICAM, VLA-4, VCAM, tumor associated antigen, HER2, HER3, HER4, nucleosomomin, a heterogeneous nuclear ribonucleoproteins (hnRNPs), fibrillarin; or fragments or variants thereof.

In still other embodiments, the target biomarker is selected from the group consisting of epidermal growth factor receptor, transferrin receptor, platelet-derived growth factor receptor, Erb-B2, CD19, CD20, CD45, CD52, Ep-CAM, alpha ([alpha])-fetoprotein, carcinoembryonic antigen peptide-1, caspase-8, CDC27, CDK4, carcino-embryonic antigen, calcium-activated chloride channel-2, cyclophilin B, differentiation antigen melanoma, elongation factor 2, Ephrin type-A receptor 2, 3, Fibroblast growth factor-5, fibronectin, glycoprotein 250, G antigen, N-acetylglucosaminyltransferase V, glycoprotein 100 kD, helicase antigen, human epidermal receptor-2/neurological, heat shock protein 70-2 mutated, human signet ring tumor-2, human telomerase reverse transcriptase, intestinal carboxyl esterase, interleukin 13 receptor [alpha]2 chain, [beta]-D-galactosidase 2-[alpha]-L-fucosyltransferase, melanoma antigen, melanoma antigen recognized by T cells-1/Melanoma antigen A, melanocortin 1 receptor, macrophage colony-stimulating factor, mucin 1, 2, melanoma ubiquitous mutated 1, 2, 3, New York-esoophagous 1, ocular albinism type 1 protein, O-linked N-acetyl glucosamine transferase gene, protein 15, promyelocytic leukemia/retinoic acid receptor [alpha], prostate-specific antigen, prostate-specific membrane antigen, receptor-type protein-tyrosinephosphatase kappa, renal antigen, renal ubiquitous 1, 2, sarcoma antigen, squamous antigen rejecting tumor 1, 2, 3, synovial sarcoma, Survivin-2B, synaptotagmin I/synovial sarcoma, X fusion protein, translocation Ets-family leukemia/acute myeloid leukemia 1, transforming growth factor [beta] receptor 2, triosephosphate isomerase, taxol resistant associated protein 3, testis-related gene, tyrosinase related protein 1, and tyrosinase related protein 2.

The target biomarker can be a cancer-associated or tumor associated antigen. The cancer-associated antigen may include without limitation one or more of human Her2/neu, Herl/EGF receptor (EGFR), HER2 (ERBB2), Her3, Her4, A33 antigen, B7H3, CD5, CD19, CD20, CD22, CD23 (IgE Receptor), C242 antigen, 5T4, IL-6, IL-13, vascular endothelial growth factor VEGF (e.g., VEGF-A), VEGFR-1, VEGFR-2, CD30, CD33, CD37, CD40, CD44, CD51, CD52, CD56, CD74, CD80, CD152,
CD200, CD221, CCR4, HLA-DR, CTLA-4, N PC-1C, tenascin, vimentin, insulin-like growth factor 1 receptor (IGF-1R), alpha-fetoprotein, insulin-like growth factor 1 (IGF-1), carbonic anhydrase 9 (CAIX), carcinoembryonic antigen (CEA), integrin α/β3, integrin α5β1, folate receptor 1, transmembrane glycoprotein NMB, fibroblast activation protein alpha (FAP), glypican 3, glycoprotein 75, TAG-72, MUC1, MUC16 (also known as CA-125), phosphatidylserine, prostate-specific membrane antigen (PMSA), NR-LU-13 antigen, TRAIL-R1, tumor necrosis factor receptor superfamily member 10b (TNFRSF10B or TRAIL-R2), SLAM family member 7 (SLAM F7), EGP40 pancarcinoma antigen, B-cell activating factor (BAFF), platelet- derived growth factor receptor, glycoprotein EpCAM (17-1A), Programmed Death-1 (PD1), Programmed Death Ligand 1 (PD-L1), protein disulfide isomerase (PDI), Phosphatase of Regenerating Liver 3 (PRL-3), prostatic acid phosphatase, Lewis-Y antigen, GD2 (a disialoganglioside expressed on tumors of neuroectodermal origin), or mesothelin. For example, the target can be one or more of human Her2/neu, Herl/EGFR, TNF-a, B7H3 antigen, CD20, VEGF, CD52, CD33, CTLA-4, tenascin, alpha-4 (a4) integrin, IL-23, amyloid-β, Huntingtin, CD25, nerve growth factor (NGF), TrkA, and a-synuclein. In some embodiments, the target biomarker is a tumor antigen selected from the group consisting of PSMA, BRCA1, BRCA2, alpha-actinin-4, BCR-ABL fusion protein (b3a2), CASP-8, β-catenin, Cdc27, CDK4, dek-can fusion protein, Elongation factor 2, ETV6-AML1 fusion protein, LDLR-fucosyltransferase AS fusion protein, hsp70-2, KIAAO205, MART2, MUM- If, MUM-2, MUM-3, neo-PAP, Myosin class 1, OS-9g, pml-RAR alpha fusion protein, PTPRK, K-ras, N-ras, CEA, gpl00/Pmell7, Kallikrein 4, mammaglobin-A, Melan-A/MART- 1, PSA, TRP-l/gp75, TRP-2, tyrosinase, CPSF, EphA3, G250/MN/CAIX, HER-2/neu, Intestinal carboxyl esterase, alpha-fetoprotein, M-CSF, MUC1, p53, PRAME, RAGE-1, RU2AS, survivin, Telomerase, WT1, or CA125. In still other embodiments, the target biomarker is a tumor antigen selected from the group consisting of 4-1BB, ST4, AGS-5, AGS-16, Angiopoietin 2, B7.1, B7.2, B7DC, B7H1, B7H2, B7H3, BT-062, BTLA, CAIX, Carcinoembryonic antigen, CTLA4, Cripto, ED-B, ErbB1, ErbB2, ErbB3, ErbB4, EpCAM, EphA2, EphA3, EphB2, EphB3, FAP, Fibronectin, Folate Receptor, Ganglioside GM3, GD2, glucocorticoid-induced tumor necrosis factor receptor (GITR), gplOO, gpa33, GPMB, ICOS, IGFIR, Integrin av, Integrin ανβ, KIR, LAG-3, Lewis Y, Mesothelin, c-MET, MN Carbonic anhydrase IX, MUC1, MUC16, Nectin-4, NKG2D, NOTCH, OX40, OX40L, PD-1, PDL1, PSCA, PSMA, RANKL, ROR1, ROR2, SLC44A4, Syndecan-1, TACI, TAG-72, Tenascin, TIM3, TRAILR1, TRAILR2, VEGFR-1, VEGFR-2, VEGFR-3, and variants thereof. In still other embodiments, the target biomarker is a tumor-associated antigen selected from the group consisting of Lewis Y, Muc-1, erbB-2,-3 and-4, Ep-CAM, EGF-receptor (e.g., EGFRTYPE I or EGFRTYPE II), EGFR deletion neoepitope, CA19-9, Muc-1, LeY, TF-, Tn-and STn-antigen, TAG-72, PSMA, STEAP, Cora antigen, CD7, CD19 and CD20, CD22, CD25, Ig-a and Ig-β, A33 and G250, CD30, MCSP and gp100, CD44-v6, MT-MMPs, (MIS) receptor type II, carboanhydrase 9, F19-antigen, Ly6, desmoglein 4, PSCA, Wue-1, GD2 and GD3 as well as TM4SF-antigens (CD63, L6, CO-29, SAS) and the alpha and/or gamma subunit of the fetal type acetylcholinreceptor (AChR). The target biomarker can be a cancer antigen selected from A33, BAGE, Bcl-2, β-catenin, CA125, CA19-9, CD5, CD19, CD20, CD21, CD22, CD33, CD37, CD45, CD123, CEA,
c-Met, CS-1, cyclin B1, DAGE, EBNA, EGFR, ephrinB2, estrogen receptor, FAP, ferritin, folate-binding protein, GAGE, G250, GD-2, GM2, gp75, gp100 (Pmel 17), HER-2/neu, HPV E6, HPV E7, Ki-67, LRP, mesothelin, p53, PRAME, progesterone receptor, PSA, PSMA, MAGE, MART, mesothelin, MUC, MUM-1 -B, myc, NYESO-1, ras, RORI, survivin, tenascin, TSTA tyrosinase, VEGF, and WT1. The target biomarker can also be a tumor antigen selected from carcinoembryonic antigen (CEA), alpha-fetoprotein (AFP), prostate specific antigen (PSA), prostate specific membrane antigen (PSMA), CA-125 (epithelial ovarian cancer), soluble Interleukin-2 (IL-2) receptor, RAGE-1, tyrosinase, MAGE-2, NY-ESO-1, Melan- A/MART-1, glycoprotein (gp) 75, gp100, beta-catenin, PRAME, MUM-1, ZFP161, Ubiquitin-1, HOX-B6, YB-1, Osteonectin, ILF3, or IGF-1. In some embodiments, the cancer-related antigen is one or more of CD2, CD4, CD19, CD20, CD22, CD23, CD30, CD33, CD37, CD40, CD44v6, CD52, CD56, CD70, CD74, CD79a, CD80, CD98, CD138, EGFR (Epidermal growth factor receptor), VEGF (Vascular endothelial growth factor), VEGFRI (Vascular endothelial growth factor receptor 1), PDGFR (Platelet-derived growth factor receptor), RANKL (Receptor activator of nuclear factor kappa-B ligand), GPNMB (Transmembrane glycoprotein Neuromedin B), EphA 2 (Ephrin type-A receptor 2), PSMA (Prostate-specific membrane antigen), Cripto (Crypotic family protein IB), EpCAM (Epithelial cell adhesion molecule), CTLA 4 (Cytotoxic T-Lymphocyte Antigen 4), IGF-IR (Type 1 insulin-like growth factor receptor), GP3 (M13 bacteriophage), GP9 (Glycoprotein IX (platelet), CD42a, GP 40 (Glycoprotein 40kDa), GPC3 (glypican-3), GPC1 (glypican-1), TRAILRI (Tumor necrosis factor-related apoptosis-inducing ligand receptor 1), TRAILRII (Tumor necrosis factor-related apoptosis-inducing ligand receptor II), FAS (Type II transmembrane protein), PS (phosphatidyl serine) lipid, Gal GalNac Gal N-linked, Mucl (Mucin 1, cell surface associated, PEM), Mucl8, CD146, A5B1 integrin (α5β1), α4β1 integrin, αv integrin (Vitronectin Receptor), Chondrolectin, CAIX (Carbonic anhydrase IX, gene G250/MN-encoded transmembrane protein), GD2 gangloside, GD3 gangloside, GM1 gangloside, Lewis Y, Mesothelin, HER2 (Human Epidermal Growth factor 2), HER3, HER4, FN14 (Fibroblast Growth Factor Inducible 14), CS1 (Cell surface glycoprotein, CD2 subset 1, CRACC, SLAMF7, CD319), 41BB CD137, SIP (Siah-1 Interacting Protein), CTGF (Connective tissue growth factor), HLADR (MHC class II cell surface receptor), PD-1 (Programmed Death 1, Type I membrane protein, PD-L1 (Programmed Death Ligand 1), PD-L2 (Programmed Death Ligand 2), IL-2 (Interleukin-2), IL-8 (Interleukin-8), IL-13 (Interleukin-13), PIGF (Phosphatidylinositol-glycan biosynthesis class F protein), NRP1 (Neuropilin-1), ICAM1, CD54, GC182 (Claudin 18.2), Claudin, HGF (Hepatocyte growth factor), CEA (Carcinoembryonic antigen), LTβR (lymphotoxin β receptor), Kappa Myeloma, Folate Receptor alpha, GRP78 (BIP, 78 kDa Glucose-regulated protein), A33 antigen, PSA (Prostate-specific antigen), CA 125 (Cancer antigen 125 or carbohydrate antigen 125), CA19.9, CA15.3, CA242, leptin, prolactin, osteopontin, IGF-II (Insulin-like growth factor 2), fascin, sPIgR (secreted chain of polymorphic immunoglobulin receptor), 14-3-3 protein eta, 5T4 oncofetal protein, ETA (epithelial tumor antigen), MAGE (Melanoma-associated antigen), MAPG (Melanoma-associated proteoglycan, NG2), vimentin, EPCA-1 (Early prostate cancer antigen-2), TAG-72 (Tumor-associated glycoprotein 72), factor VIII, Nephrilysin (Membrane metallo-endopeptidase) and 17-1 A (Epithelial cell surface antigen 17-1A). The
cancer antigen can be selected from the group consisting of carbonic anhydrase IX, alpha-fetoprotein, A3, antigen specific for A33 antibody, Ba 733, BrE3-antigen, CA125, CD1, CD1a, CD3, CD5, CD15, CD16, CD19, CD20, CD21, CD22, CD23, CD25, CD30, CD33, CD38, CD45, CD74, CD79a, CD80, CD138, colon-specific antigen-p (CSAp), CEA (CEACAM5), CEACAM6, CSAp, EGFR, EGP-1, EGP-2, Ep-CAM, Fit-1, Fli-3, folate receptor, HLA-DR, human chorionic gonadotropin (HCG) and its subunits, HER2/neu, hypoxia inducible factor (HIF-1), la, IL-2, IL-6, IL-8, insulin growth factor-1 (IGF-1), KC4-antigen, KS-1-antigen, KS1-4, Le-Y, macrophage inhibition factor (MIF), MAGE, MUC1, MUC2, MUC3, MUC4, MUC16, NCA66, NCA95, NCA90, antigen specific for PAM-4 antibody, placental growth factor, p53, prostatic acid phosphatase, PSA, PSMA, RS5, S100, TAC, TAG-72, tenascin, TRAIL receptors, Tn antigen, Thomson-Friedenreich antigens, tumor necrosis antigens, VEGF, ED-B fibronectin, 17-1A-antigen, an angiogenesis marker, an oncogene marker and an oncogene product.

The tumor marker can be a generic tumor marker or be associated with certain tumor types, such as those originating from different anatomical origins. In an embodiment, the tumor marker can be chosen to correspond to a certain tumor type. For example, exemplary tumor markers and associated tumor types include without limitation the following, listed as antigen (optional name) (cancer types): Alpha fetoprotein (AFP) (germ cell tumor, hepatocellular carcinoma); CA15-3 (breast cancer); CA27-29 (breast cancer); CA19-9 (mainly pancreatic cancer, but also colorectal cancer and other types of gastrointestinal cancer); CA-125 (ovarian cancer, endometrial cancer, fallopian tube cancer, lung cancer, breast cancer and gastrointestinal cancer); Calcitonin (medullary thyroid carcinoma); Calretinin (mesothelioma, sex cord-gonadal stromal tumour, adrenocortical carcinoma, synovial sarcoma); Carcinoembryonic antigen (gastrointestinal cancer, cervix cancer, lung cancer, ovarian cancer, breast cancer, urinary tract cancer); CD34 (hemangiopericytoma/solitary fibrous tumor, pleomorphic lipoma, gastrointestinal stromal tumor, dermatofibrosarcoma protuberans); CD99 (MIC2) (Ewing sarcoma, primitive neuroectodermal tumor, hemangiopericytoma/solitary fibrous tumor, synovial sarcoma, lymphoma, leukemia, sex cord-gonadal stromal tumour); CD1 17 (gastrointestinal stromal tumor, mastocytosis, seminoma); Chromogranin (neuroendocrine tumor); Chromosomes 3, 7, 17, and 9p21 (bladder cancer); Cytokeratin (various types) (various carcinoma, some types of sarcoma); Desmin (smooth muscle sarcoma, skeletal muscle sarcoma, endometrial stromal sarcoma); Epithelial membrane antigen (EMA) (many types of carcinoma, meningioma, some types of sarcoma); Factor VIII (CD31, FL1) (vascular sarcoma); Glial fibrillary acidic protein (GFAP) (glioma (astrocytoma, ependymoma)); Gross cystic disease fluid protein (GCDFP-15) (breast cancer, ovarian cancer, salivary gland cancer); HMB-45 (melanoma, PEComa (for example angiomyolipoma), clear cell carcinoma, adrenocortical carcinoma); Human chorionic gonadotropin (hCG) (gestational trophoblastic disease, germ cell tumor, chorioicarcinoma); Immunoglobulin (lymphoma, leukemia); Inhibin (sex cord-gonadal stromal tumour, adrenocortical carcinoma, hemangioblastoma); keratin (various types) (carcinoma, some types of sarcoma); lymphocyte marker (various types, lymphoma, leukemia); MART-1 (Melan-A) (melanoma, steroid-producing tumors (adrenocortical carcinoma, gonadal tumor)); Myo D1 (rhabdomyosarcoma, small, round, blue cell tumour); muscle-specific actin (MSA) (myosarcoma (leiomyosarcoma, rhabdomyosarcoma); neurofilament
(neuroendocrine tumor, small-cell carcinoma of the lung); neuron-specific enolase (NSE) (neuroendocrine tumor, small-cell carcinoma of the lung, breast cancer); placental alkaline phosphatase (PLAP) (seminoma, dysgerminoma, embryonal carcinoma); prostate-specific antigen (prostate); PTPRC (CD45) (lymphoma, leukemia, histiocytic tumor); SI100 protein (melanoma, sarcoma (neurosarcoma, lipoma, chondrosarcoma), astrocytoma, gastrointestinal stromal tumor, salivary gland cancer, some types of adenocarcinoma, histiocytic tumor (dendritic cell, macrophage)); smooth muscle actin (SMA) (gastrointestinal stromal tumor, leiomyosarcoma, PEComa); synaptophysin (neuroendocrine tumor); thyroglobulin (thyroid cancer but not typically medullary thyroid cancer); thyroid transcription factor-1 (all types of thyroid cancer, lung cancer); Tumor M2-PK (colorectal cancer, Breast cancer, renal cell carcinoma, Lung cancer, Pancreatic cancer, Esophageal Cancer, Stomach Cancer, Cervical Cancer, Ovarian Cancer); Vimentin (sarcoma, renal cell carcinoma, endometrial cancer, lung carcinoma, lymphoma, leukemia, melanoma). Additional tumor types and associated biomarkers comprise the following, listed as tumor type (markers): Colorectal (M2-PK, CEA, CA 19-9, CA 125); Breast (CEA, CA 15-3, Cyfra 21-1); Ovary (CEA, CA 19-9, CA 125, AFP, BHCG); Uterine (CEA, CA 19-9, CA 125, Cyfra 21-1, SCC); Prostate (PSA); Testicle (AFP, BHCG); Pancreas/Stomach (CEA, CA 19-9, CA 72-4); Liver (CEA, AFP); Oesophagus (CEA, Cyfra 21-1); Thyroid (CEA, NSE); Lung (CEA, CA 19-9, CA 125, NSE, Cyfra 21-1); Bladder (CEA, Cyfra 21-1, TPA). One or more of these markers can be used as the target biomarker recognized by the variable region of the multipartite construct of the invention.

[00492] In some embodiments of the invention, the target biomarker recognized by the variable region comprises one or more of PDGF, IgE, IgE Fee Rl, PSMA, CD22, TNF-alpha, CTLA4, PD-1, PD-L1, PD-L2, FcRIIB, BTLA, TIM-3, CD11c, BAFF, B7-X, CD19, CD20, CD25, and CD33. The target biomarker can also be a protein comprising one or more of insulin-like growth factor 1 receptor (IGF1R), IGF2R, insulin-like growth factor (IGF), mesenchymal epithelial transition factor receptor (c-met), hepatocyte growth factor (HGF), epidermal growth factor receptor (EGFR), ErbB2, ErbB3, epidermal growth factor (EGF), heregulin, fibroblast growth factor receptor (FGFR), platelet-derived growth factor receptor (PDGFR), platelet-derived growth factor (PDGF), vascular endothelial growth factor receptor (VEGFR), vascular endothelial growth factor (VEGF), tumor necrosis factor receptor (TNFR), tumor necrosis factor alpha (TNF-a), folate receptor (FOLR), folate, transferrin receptor (TIR), mesothelia, Fc receptor, c-kit receptor, c-kit, a4 integrin, P-selectin, sphingosine-1-phosphate receptor-1 (S1PR), hyaluronate receptor, leukocyte function antigen-1 (LFA-1), CD4, CD11, CD18, CD20, CD25, CD27, CD52, CD70, CD80, CD85, CD95 (Fas receptor), CD106 (vascular cell adhesion molecule 1 (VCAM1)), CD166 (activated leukocyte cell adhesion molecule (ALCAM)), CD 178 (Fas ligand), CD253 (TNF-related apoptosis-inducing ligand (TRAIL)), inducible costimulator (ICOS) ligand, CCR2, CXCR3, CCR5, CXCL12 (stromal cell-derived factor 1 (SDF-1)), interleukin 1 (IL-1), cytotoxic T-lymphocyte antigen 4 (CTLA-4), MART-1, gliO0, MAGE-1, ephrin (Eph) receptor, mucosal addressin cell adhesion molecule 1 (MAdCAM-1), carcinoembryonic antigen (CEA), LewisY, MUC-1, epithelial cell adhesion molecule (EpCAM), cancer antigen 125 (CA125), prostate specific membrane antigen (PSMA), TAG-72 antigen, and fragments thereof. In various embodiments, the target biomarker comprises one or more of PSMA,
The target biomarker can be a microvesicle antigen, such as a microvesicle antigen selected from any of Table 3, Table 4, Tables 18-26, or Table 45. For example, the target biomarker can be one or more microvesicle antigen selected from CD9, EphA2, EGFR, B7H3, PSMA, PCSA, CD63, STEAP, CD81, B7H3, STEAP1, ICAM1 (CD54), A33, DR3, CD66e, MFG-e8, Hepsin, TMEM211, TROP-2, EGFR, Mammoglobin, Hepsin, NPGP/NPF2, PSCA, 5T4, NGAL, NK-2, EpCam, NK-1R, 5T4, PAI-1, and CD45. The target biomarker can be one or more microvesicle antigen selected from SPB, SPC, NSE, PGP9.5, CD9, P2RX7, NDUFB7, NSE, Gal3, Osteopontin, CHI3L1, EGFR, B7H3, iC3b, MUC1, Mesothelin, SPA, TPA, PCSA, CD63, AQP5, DLL4, CD81, DR3, PSMA, GPCR 110 (GPR1 10), EPHA2, CEACAM, PTP, CABYR, TMEM211, ADAM28, UNC93a, A33, CD24, CD10, NGAL, EpCam, MUC17, TROP2 and MUC2. In some embodiments, the target biomarker comprises one or more microvesicle antigen selected from CD9, CD63, CD81, B7H3, PRO GRP, CYTO 18, FTHI, TGM2, CENPH, ANNEXIN I, ANNEXIN V, ERBB2, EGFR, CRP, VEGF, CYTO 19, CCL2, Osteopontin (OST19), Osteopontin (OST22), BTUB, CD45, TIMP, NACC1, MMP9, BRCA1, P27, NSE, M2PK, HCG, MUC1, CEA, CEACAM, CYTO 7, EPCAM, MS4A1, MUC1, MUC2, PGP9, SPA, SPA, SPD, P53, GPCR (GPR1 10), SFTP, UNCR2, NSE, INGA3, INTO b4, MMP1, PNT, RACK1, NAP2, HLA, BMP2, PTH1R, PAN ADH, NCAM, CD151, CKS1, FSHR, HIF, KRAS, LAMP2, SNAIL, TRIM29, TSPAN1, TWIST1, ASPH and AURKB. In another embodiment, the target biomarker is selected from the group of proteins consisting of CD9, PSMA, PCSA, CD63, CD81, B7H3, IL 6, OPG-13, IL6R, PA2G4, EZH2, RUNX2, SERPINB3, and EpCam. In another embodiment, a target biomarker is selected from the group of proteins consisting of A33, a33 n15, AFP, ALA, ALIX, ALP, AnnexinV, APC, ASCA, ASPH (246-260), ASPH (666-680), ASPH (A-10), ASPH (DO IP), ASPH (D03), ASPH (G-20), ASPH (H-300), AURKA, AURKB, B7H3, B7H4, BCA-225, BCNP1, BDNF, BRCA, CA125 (MUC16), CA-19-9, C-Bir, CD1.1, CD10, CD174 (Lewis y), CD24, CD44, CD46, CD59 (MEM-43), CD63, CD66e CEA, CD73, CD81, CD9, CDA, CDAC1 la2, CEA, C-Erb2, C-erbB2, CRMP-2, CRP, CXCL12, CYFRA21-1, DLL4, DR3, EGFR, Epcam, Epha2, Epha2 (H-77), ER, ErbB4, EZH2, FASL, FRT, FRT c.f23, GDF15, GPCR, GPR30, Gro-alpha, HAP, HBD 1, HBD2, HER 3 (ErbB3), HSP, HSP70, hVEGFR2, ic3b, IL 6 Unc, IL-1B, IL6 Unc, IL6R, IL8, IL-8, INSIG-2, KLK2, LICAM, LAMN, LDH, MACC-1, MAP4K, MART-1, MCP-1, M-CSF, MFG-E8, MCI, MIF, MIS RII, MMG, MMP26, MMP7, MMP9, MS4A1, MUC1, MUC1 seqI, MUC1 seqIIA, MUC17, MUC2, Ncam, NGAL, NPGP/NPF2, OPG, OPN, p53, p53, PA2G4, PBP, PCSA, PDGFRB, PGP9.5, PIM1, PR (B), PRL, PSA, PSMA, PSME3, PTEN, R5-C9 Tube 1, Reg IV, RUNX2, SCRNI, seprase, SERPINB3, SPARC, SPB, SPDEF, SRVN, STAT 3, STEAP1, TF (FL-295), TFF3, TIMG2, TIMP-1, TIMP1, TIMP2, TMEM211, TMPRSS2, TNF-alpha, Trail-R2, Trail-R4, TrkB, TROP2, Tsg 101, TWEAK, UNC93A, VEGF A, and YPSMA-1. The target biomarker can be selected from the group of proteins consisting of 5T4, A33, ACTG1, ADAM10, ADAM 15, AFP, ALA, ALDOA, ALIX, ALP, ALX4, ANCA, Annexin V, ANXA2, ANXA6, APC,
APOA1, ASCA, ASPH, ATP1A1, AURKA, AURKB, B7H3, B7H4, BANK1, BASP1, BCA-225, BCNPI, BDNF, BRCA, Clorf58, C20orf114, C8B, CA125 (MUC16), CA-19-9, CAPZAI, CAV1, C-Bir, CCSA-2, CCSA-3&4, CD1.1, CD10, CD151, CD174 (Lewis y), CD24, CD2AP, CD37, CD44, CD46, CD53, CD59, CD63, CD66 CEA, CD73, CD81, CD82, CD9, CDA, CADAC1 la2, CEA, C-Erbb2, CFL1, CFP, CHMP4B, CLTC, COTL1, CRMP-2, CRP, CRTN, CTNNCD1, CTSB, CTSZ, CXCL12, CYCS, CYFRA21-1, DcR3, DLLL4, DPP4, DR3, EEF1A1, EGF, EHD1, ENO1, EpCAM, EphA2, ER, ErbB4, EZH2, F1I1R, F2, F5, FAM125A, FASL, Ferritin, FNBPL, FOLH1, FRT, GAL3, GAPDH, GDF15, GLB1, GPCR (GPR1 10), GPR30, GPX3, GRO-1, Gro-alphal a, HAP, HBD 1, HBD2, HER 3 (ErbB3), HIST1H1C, HIST1H2AB, HNP1-3, HSP, HSP70, HSP90AB1, HSPA1B, HSPA8, hVEGFR2, ic3b, ICAM, IGSF8, IL 6, IL-1B, IL6R, IL8, IMP3, INSIG-2, ITGB1, ITIH3, JUP, KLK2, L1CAM, LAMN, LDH, LDHA, LDHB, LUM, LYZ, MACC-1, MAPK4, MART-1, MCP-1, M-CSF, MFE8, MGAM, MGC20553, MIC1, MIF, MIS RII, MMG, MMP26, MMP7, MMP9, MSAI1, MUC1, MUC17, MUC2, MYH2, MYL6B, Ncam, NGAL, NME1, NME2, NNMT, NPG/NNPF2, OP, OPG, OPG-13, OPN, p53, PA2G4, PABPC1, PABPC4, PACSIN2, PBP, PCBP2, PCSA, PDCD6IP, PDGFRB, PGP9.5, PIM1, PR (B), PRDX2, PRL, PSA, PSCA, PSMA, PSA1, PSA2, PSA4, PSA6, PSA7, PSMB1, PSMB2, PSMB3, PSMB4, PSMB5, PSMB6, PSMB8, PTEN, PTGFRN, Rab-5b, Reg IV, RPS27A, RUNX2, SCRNI, SDCBP, seprase, Sept-9, SERINC5, SERPINB3, SERPINB3, SH3GL1, SLC3A2, SMPLD3B, SNX9, SPARC, SPB, SPDEF, SPON2, SPR, SRVN, SXX2, SXX4, STAT 3, STEAP, STEAP1, TACSTD1, TCN2, tetraspanin, TF (FL-295), TFF3, TGM2, THBS1, TIMP, TIMP1, TIMP2, TMEM211, TMPRSS2, TFN-alpha, TPA, TPI1, TPS, Trail-R2, Trail-R4, TrKb, TROP2, TROP2, Tsg 101, TUBB, TWEAK, UNC93A, VDAC2, VEGF A, VPS37B, YPSMA-1, YWHAG, YWHAQ, and YWHAZ.

In another embodiment, the target biomarker is selected from the group of proteins consisting of 5T4, ACTG1, ADAM 10, ADAM 15, ALDOA, ANXA2, ANXA6, APOA1, ATP1A1, BASP1, Clorf58, C20orf114, C8B, CAPZAI, CAV1, CD151, CD2AP, CD59, CD9, CFLI, CFP, CHMP4B, CLTC, COTL1, CTNNCD1, CTSB, CTSZ, CYCS, DPP4, EEF1A1, EHD1, ENO1, F1I1R, F2, F5, FAM125A, FNBPL, FOLH1, GAPDH, GLB1, GPX3, HIST1H1C, HIST1H2AB, HSP90B1, HSPA1B, HSPA8, IGSF8, ITGB1, ITIH3, JUP, LDHA, LDHB, LUM, LYZ, MFG-E8, MGAM, MMP9, MYH2, MYL6B, NME1, NME2, PABPC1, PABPC4, PACSIN2, PCBP2, PDCD6IP, PRDX2, PSA, PSMA, PSA1, PSA2, PSA4, PSA6, PSA7, PSMB1, PSMB2, PSMB3, PSMB4, PSMB5, PSMB6, PSMB8, PTGFRN, RPS27A, SDCBP, SERINC5, SH3GL1, SLC3A2, SMPLD3B, SNX9, TACSTD1, TCN2, THBS1, TPII, TSG101, TUBB, VDAC2, VPS37B, YWHAG, YWHAQ, and YWHAZ. In another embodiment, the target biomarker is selected from the group of proteins consisting of CD9, CD63, CD81, PSMA, PCSA, B7H3 and EpCam. In another embodiment, the target biomarker is selected from the group of proteins consisting of a tetraspanin, CD9, CD63, CD81, CD63, CD9, CD81, CD82, CD37, CD53, Rab-5b, Annexin V, MFG-E8, Mucl, GPCR 110, TMEM211 and CD24. In another embodiment, the target biomarker is selected from the group of proteins consisting of A33, AFP, ALIX, ALX4, ANCA, APC, ASCA, AURKA, AURKB, B7H3, BANK1, BCNPI, BDNF, CA-19-9, CCA-2, CCSA-3, CD10, CD12, CD24, CD44, CD63, CD66 CEA, CD66 CE, CD81, CD9, CDA, C-Erbb2, CRMP-
WO 2017/004243

PCT/US2016/040157

(scaffolding), TRAIL-R2, TRAP1, TrkB, Tsg 101, TXNDC16, U2AF2, UEVLD, UFC1, UNCI93a, USP14, VASP, VCP, VDAC1, VEGFA, VEGFR1, VEGFR2, VPS37C, WIZ, XRCC5, XRCC6, YB-1, YWHAZ, or any combination thereof. In other embodiments, the target biomarker is selected from the group consisting of p53, p63, p73, mdm-2, procathepsin-D, B23, C23, PLAP, CA125, MUC-1, HER2, NY-ESO-1, SCPI, SSX-1, SSX-2, SSX-4, HSP27, HSP60, HSP90, GRP78, TAG72, HoxA7, HoxB7, EpCAM, ras, mesothelin, survivin, EGFK, MUC-1, or c-myc. The microvesicle antigen can be from any of Tables 18-26 or Table 45.

[00494] One of skill will appreciate that the above biomarker listings are not intended to be mutually exclusive. For example, a single target biomarker can have one or more of the following attributes: cancer/tumor antigen, cell antigen, microvesicle antigen, membrane antigen, and any combination thereof.

In some embodiments, the target biomarker will have all of these attributes.

[00495] As noted above, the IDM domain can be constructed to illicit a complement mediated immune response that can induce apoptosis. Such IDM can include but are not limited to Clq, Clr, Cls, CI, C3a, C3b, C3d, C5a, C2, C4, and cytokines. The IDM region may comprise an oligonucleotide sequence including without limitation Toll-Like Receptor (TLR) agonists like CpG sequences which are immunostimulatory and/or polyG sequences which can be anti-proliferative or pro-apoptotic. The moiety can be vaccine like moiety or antigen that stimulates an immune response. In an embodiment, the immune stimulating moiety comprises a superantigen. In some embodiments, the superantigen can be selected from the group consisting of staphylococcal enterotoxins (SEs), a Streptococcus pyogenes exotoxin (SPE), a Staphylococcus aureus toxic shock-syndrome toxin (TSST-1), a streptococcal mitogenic exotoxin (SME), a streptococcal superantigen (SSA), a hepatitis surface antigen, or a combination thereof. Other bacterial antigens that can be used with the invention comprise bacterial antigens such as Freund's complete adjuvant, Freund's incomplete adjuvant, monophosphoryl-lipid A/trehalose dicorynomycolate (Ribi's adjuvant), BCG (Calmette-Guerin Bacillus; Mycobacterium bovis), and Corynebacterium parvum. The immune stimulating moiety can also be a non-specific immunostimulant, such as an adjuvant or other non-specific immunomodulator. Useful adjuvants comprise without limitation aluminium salts, alum, aluminium phosphate, aluminium hydroxide, squalene, oils, MF59, and AS03 (“Adjuvant System 03”).

The adjuvant can be selected from the group consisting of Cationic liposome-DNA complex JVRS-100, aluminum hydroxide vaccine adjuvant, aluminum phosphate vaccine adjuvant, aluminum potassium sulfate adjuvant, Alhydrogel, ISCOM(s)™, Freund's Complete Adjuvant, Freund's Incomplete Adjuvant, CpG DNA Vaccine Adjuvant, Cholera toxin, Cholera toxin B subunit, Liposomes, Saponin Vaccine Adjuvant, DDA Adjuvant, Squalene-based Adjuvants, Etx B subunit Adjuvant, IL-12 Vaccine Adjuvant, LTK63 Vaccine Mutant Adjuvant, TiterMax Gold Adjuvant, Ribi Vaccine Adjuvant, Montanide ISA 720 Adjuvant, Corynebacterium-derived P40 Vaccine Adjuvant, MPL™ Adjuvant, AS04, AS02, Lipopolysaccharide Vaccine Adjuvant, Muramyl Dipeptide Adjuvant, CRL1005, Killed Corynebacterium parvum Vaccine Adjuvant, Montanide ISA 51, Bordetella pertussis component Vaccine Adjuvant, Cationic Liposomal Vaccine Adjuvant, Adamantylamide Dipeptide Vaccine Adjuvant, Arlacel A, VSA-3 Adjuvant, Aluminum vaccine adjuvant, Polygen Vaccine Adjuvant, Adjumer™, Algal Glucan, Bay...
R1005, Theramide®, Stearyl Tyrosine, Specol, Algammulin, Avridine®, Calcium Phosphate Gel, CTA1-DD gene fusion protein, DOC/Alum Complex, Gamma Inulin, Gerbu Adjuvant, GM-CSF, GMDP, Recombinant hIFN-gamma/Interferon-g, Interleukin-1β, Interleukin-2, Interleukin-7, Sclavo peptide, Rehydragel LV, Rehydragel HPA, Loxoribine, MF59, MTP-PE Liposomes, Murametide, Murapalmitine, D-Murapalmitine, NAGO, Non-Ionic Surfactant Vesicles, PMMA, Protein Cochleates, QS-21, SPT (Antigen Formulation), nanoemulsion vaccine adjuvant, AS03, Quil-A vaccine adjuvant, RC529 vaccine adjuvant, LTR192G Vaccine Adjuvant, E. coli heat-labile toxin, LT, amorphous aluminum hydroxyphosphate sulfate adjuvant, Calcium phosphate vaccine adjuvant, Montaniide Incomplete Seppic Adjuvant, Imiquimod, Resiquimod, AF03, Flagellin, Poly(I:C), ISCOMATRIX®, Abisco-100 vaccine adjuvant, Albumin-heparin microparticles vaccine adjuvant, AS-2 vaccine adjuvant, B7-2 vaccine adjuvant, DHEA vaccine adjuvant, Immunoliposomes Containing Antibodies to Costimulatory Molecules, SAF-1, Sendai Proteoliposomes, Sendai-containing Lipid Matrices, Threonyl muramyl dipeptide (TMDP), Ty Particles vaccine adjuvant, Bupivacaine vaccine adjuvant, DL-PGL (Polyester poly (DL-lactide-co-glycolide)) vaccine adjuvant, IL-15 vaccine adjuvant, LTK72 vaccine adjuvant, MPL-SE vaccine adjuvant, non-toxic mutant E1 12K of Cholera Toxin mCT-El 12K, and Matrix-S. Additional adjuvants that can be used with the multipartite constructs of the invention can be identified using the Vaxjo database. See Sayers S, Ulysse G, Xiang Z, and Ye Y. Vaxjo: a web-based vaccine adjuvant database and its application for analysis of vaccine adjuvants and their uses in vaccine development. Journal of Biomedicine and Biotechnology. 2012;2012:831486. Epub 2012 Mar 13. PMID: 22505817; www.violinet.org/vaxjo/. Other useful non-specific immunostimulators comprise histamine, interferon, transfer factor, tuftsin, interleukin-1, female sex hormones, prolactin, growth hormone vitamin D, deoxycholic acid (DCA), tetrachlorodecaoxide (TCDO), and imiquimod or resiquimod, which are drugs that activate immune cells through the toll-like receptor 7. A multipartite construct can be created that comprises more than one immunomodulating moiety, e.g., using segments that span CpG sequences which are immunostimulatory with complement directed segments that can stimulate apoptosis.

[0046] Modifications

[0047] Modifications to the multipartite constructs can be made to alter desired characteristics, including without limitation in vivo stability, specificity, affinity, avidity or nuclease susceptibility. Alterations to the half life may improve stability in vivo or may reduce stability to limit in vivo toxicity. Such alterations can include mutations, truncations or extensions. The 5' and/or 3' ends of the multipartite oligonucleotide constructs can be protected or deprotected to modulate stability as well. Modifications to improve in vivo stability, specificity, affinity, avidity or nuclease susceptibility or alter the half life to influence in vivo toxicity may be at the 5' or 3' end and include but are not limited to the following: LNA incorporation, phosphorothioate backbone instead of phosphodiester backbone, amino modifiers (i.e. C6-dT), dye conjugates (Cy dyes, Fluorophores, etc), Biotinylation, PEG linkers, Click chemistry linkers, dideoxynucleotide end blockers, inverted end bases, cholesterol TEG or other lipid based labels.

[0048] Linkage options for segments of the multipartite construct can be on the 5' or 3' end of an oligonucleotide or to a primary amine, sulphydryl or carboxyl group of an antibody and include but are not
limited to the following: Biotin-target oligonucleotide /Ab, streptavidin-complement oligonucleotide or vice versa, amino modified-target Ab/ oligonucleotide, thiol/carboxy-complement oligonucleotide or vice versa, Click chemistry-target Ab/ oligonucleotide, corresponding Click chemistry partner-complement oligonucleotide or vice versa. The linkages may be covalent or non-covalent and may include but are not limited to monovalent, multivalent (i.e., bi, tri or tetra-valent) assembly, to a DNA scaffold (i.e., DNA origami structure), drug/chemotherapeutic agent, nanoparticle, microparticle or a micelle or liposome.

A linker region can comprise a spacer with homo- or multifunctional reactive groups that can vary in length and type. These include but are not limited to the following: spacer C18, PEG4, PEG6, PEG8, and PEG12.

The multipartite oligonucleotide of the invention can further comprise additional elements to add desired biological effects. For example, the multipartite construct may comprise a membrane disruptive moiety. The multipartite construct may also be conjugated to one or more chemical moiety that provides such effects. For example, the multipartite construct may be conjugated to a detergent-like moiety to disrupt the membrane of a target cell or microvesicle. 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]-l-propanesulfonate (CHAPS), CHAPSO, deoxycholic acid, sodium deoxycholate, NP-40, glycosides, octyl-thio-glucosides, maltosides, and the like. One of skill will appreciate that functional fragments, such as membrane disruptive moieties, can be covalently or non-covalently attached to the multipartite construct.

Oligonucleotide segments of a multipartite construct can include any desirable base modification known in the art. In certain embodiments, oligonucleotide segments are 10 to 50 nucleotides in length. One having ordinary skill in the art will appreciate that this embodies oligonucleotides of 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 nucleotides in length, or any range derivable there within.

In certain embodiments, a multipartite construct comprises a chimeric oligonucleotide that contains two or more chemically distinct regions, each made up of at least one nucleotide. Such chimeras can be referred to using terms such as multipartite, multivalent, or the like. The oligonucleotides portions may contain at least one region of modified nucleotides that confers one or more beneficial properties, e.g., increased nuclease resistance, bioavailability, increased binding affinity for the target. Chimeric nucleic acids of the invention may be formed as composite structures of two or more oligonucleotides, two or more types of oligonucleotides (e.g., both DNA and RNA segments), modified oligonucleotides, oligonucleosides and/or oligonucleotide mimetics. Such compounds have also been referred to in the art as hybrids. Representative United States patents that teach the preparation of such hybrid structures comprise, but are not limited to, US patent nos: 5,013,830; 5,149,797; 5,220,007; 5,256,775; 5,366,878;
In certain embodiments, a multipartite construct comprises at least one nucleotide modified at the 2’ position of the sugar, including without limitation a 2’-0-alkyl, 2’-0-alkyl-0-alkyl or 2’- fluoro-modified nucleotide. In other embodiments, RNA modifications include 2’- fluoro, 2’-amino and 2’ O-methyl modifications on the ribose of pyrimidines, a basic residue or an inverted base at the 3’ end of the RNA. Such modifications are routinely incorporated into oligonucleotides and these oligonucleotides have been shown to have higher target binding affinity in some cases than 2’-deoxyoligonucleotides against a given target.

A number of nucleotide and nucleoside modifications have been shown to make an oligonucleotide more resistant to nuclease digestion, thereby prolonging in vivo half-life. Specific examples of modified oligonucleotides include those comprising backbones comprising, for example, phosphorothioates, phosphotriesters, methyl phosphonates, short chain alkyl or cycloalkyl intersugar linkages or short chain heteroatomic or heterocyclic intersugar linkages. The constructs of the invention can comprise oligonucleotides with phosphorothioate backbones and/or heteroatom backbones, e.g., CH2-NH-0-CH2, CH2-N(CH3)-0-CH2 (known as a methylene(methylimino) or MMI backbone), CH2-O-N(CH3)-CH2, CH2-O-N (CH3)-CH2 and O-N (CH3)-CH2 backbones, wherein the native phosphodiester backbone is represented as O-P—O-CH2; amide backbones (De Mesmaeker et al., 1995); morpholino backbone structures (Summerton and Weller, U.S. Pat. No. 5,034,506); peptide nucleic acid (PNA) backbone (wherein the phosphodiester backbone of the oligonucleotide is replaced with a polyamide backbone, the nucleotides being bound directly or indirectly to the aza nitrogen atoms of the polyamide backbone (Nielsen, et al., 1991), each of which is herein incorporated by reference in its entirety. Phosphorus- containing linkages include, but are not limited to, phosphorothioates, chiral phosphorothioates, phosphorodithioates, phosphotriesters, aminoalkylphosphotriesters, methyl and other alkyl phosphonates comprising 3’alkylene phosphonates and chiral phosphonates, phosphinates, phosphoramidates comprising 3’-amino phosphoramidate and aminoalkylphosphoramidates, thionophosphoramidates, thioalkylphosphonates, thioalkylphosphotriesters, and boranophosphates having normal 3’-5’ linkages, 2’-5’ linked analogs of these, and those having inverted polarity wherein the adjacent pairs of nucleoside units are linked 3*-5* to 5*-3* or 2*-5* to 5*-2*; see U.S. Patent Nos. 3,687,808; 4,469,863; 4,476,301; 5,023,243; 5,177,196; 5,188,897; 5,264,423; 5,276,019; 5,278,302; 5,286,717; 5,321, 131; 5,399,676; 5,405,939; 5,453,496; 5,455,233; 5,466,677; 5,476,925; 5,519,126; 5,536,821; 5,541,306; 5,550,111; 5,563,253; 5,571,799; 5,587,361; and 5,625,050, each of which is herein incorporated by reference in its entirety. Morpholino-based oligomeric compounds are known in the art described in Braasch & Corey, Biochemistry vol. 41, no. 14, 2002, pages 4503 - 4510; Genesis vol. 30, 2001, page 3; Heasman, J. Dev. Biol. vol. 243, 2002, pages 209 - 214; Nasevicius et al. Nat. Genet., vol. 26, 2000, pages 216 - 220; Lacerra et al. Proc. Natl. Acad. Sci. vol. 97, 2000, pages 9591 - 9596 and U.S. Pat. No. 5,034,506, issued Jul. 23, 1991, each of which is herein incorporated by reference in its entirety. Cyclohexenyl nucleic acid oligonucleotide mimetics are described in Wang et al, J. Am. Chem.
Soc. Vol. 122, 2000, pages 8595 - 8602, the contents of which is incorporated herein in its entirety. A multipartite construct of the invention can comprise at least such modification as desired.

Modified oligonucleotide backbones that do not include a phosphorus atom therein have backbones that can be formed by short chain alkyl or cycloalkyl internucleoside linkages, mixed heteroatom and alkyl or cycloalkyl internucleoside linkages, or one or more short chain heteroatomic or heterocyclic internucleoside linkages. These comprise those having morpholino linkages (formed in part from the sugar portion of a nucleoside); siloxane backbones; sulfide, sulfite, and sulfone backbones; formacetyl and thioformacetyl backbones; methylene formacetyl and thioformacetyl backbones; alkene containing backbones; sulfamate backbones; methyleneimino and methylenehydrazino backbones; sulfonate and sulfonamide backbones; amide backbones; and others having mixed N, O, S and CH2 component parts; see U.S. Patent Nos. 5,034,506; 5,166,315; 5,185,444; 5,214,134; 5,216,141; 5,235,033; 5,264,562; 5,264,564; 5,405,938; 5,434,257; 5,466,677; 5,470,967; 5,489,677; 5,541,307; 5,561,225; 5,596,086; 5,602,240; 5,610,289; 5,602,240; 5,608,046; 5,610,289; 5,618,704; 5,623,070; 5,663,312; 5,633,360; 5,677,437; and 5,677,439, each of which is herein incorporated by reference in its entirety. A multipartite construct of the invention can comprise at least such modification as desired.

In certain embodiments, a multipartite construct comprises one or more substituted sugar moieties, e.g., one of the following at the 2' position: OH, SH, SCH3, F, OCN, OCH3, OCH3

0 (CH2)n CH3, 0 (CH2)n NH2 or 0 (CH2)n CH3 where n is from 1 to about 10; Cl to ClO lower alkyl, alkoxyalkoxy, substituted lower alkyl, alkaryl or aralkyl; Cl; Br; CN; CF3; OCF3; 0-, S-, or N-alkyl; 0-, S-, or N-alkenyl; SOCH3; SO2 CH3; ON02; N 0 2; N2; NH2; heterocycloalkyl; heterocycloalkaryl; aminoalkylamino; polyalkylamino; substituted silyl; an RNA cleaving group; a reporter group; an intercalator; a group for improving the pharmacokinetic properties of an oligonucleotide; or a group for improving the pharmacokinetic/pharmacodynamic properties of an oligonucleotide and other substituents having similar properties. A preferred modification includes 2'-methoxyethoxy [2'-0-CH2CH2OCH3, also known as 2'-0-(2-methoxyethyl)] . Other preferred modifications include 2*-methoxy (2*-0-CH3), 2*-propoxy (2*-OCH2 CH2CH3) and 2*-fluoro (2*-F). Similar modifications may also be made at other positions on the oligonucleotide, e.g., the 3' position of the sugar on the 3' terminal nucleotide and the 5' position of 5' terminal nucleotide. Oligonucleotides may also have sugar mimetics such as cyclobutyls in place of the pentofuranosyl group.

In certain embodiments, a multipartite construct of the invention comprises one or more base modifications and/or substitutions. As used herein, "unmodified" or "natural" bases include adenine (A), guanine (G), thymine (T), cytosine (C) and uracil (U). Modified bases include, without limitation, bases found only infrequently or transiently in natural nucleic acids, e.g., hypoxanthine, 6-methyladenine, 5-Me pyrimidines, particularly 5-methylcytosine (also referred to as 5-methyl-2'- deoxy cytosine and often referred to in the art as 5-Me-C), 5- hydroxymethylcytosine (HMC), glycosyl HMC and gentobiosyl HMC, as well as synthetic bases, e.g., 2-amino adenine, 2-(methylamino)adenine, 2-(imidazolylalkyl)adenine, 2- (aminoalkylamino)adenine or other heterosubstituted alkyladenines, 2-thiouracil, 2-thiothymine, 5-bromouracil, 5-hydroxymethyluracil, 8-azaguanine, 7-deazaguanine, N6 (6-
aminohexyl)adenine and 2,6-diaminopurine (Kornberg, 1980; Gebeyehu, et al., 1987). A "universal" base known in the art, e.g., inosine, can also be included. 5-Me-C substitutions can also be included. These have been shown to increase nucleic acid duplex stability by 0.6-1.20°C. See, e.g., Sanghvi et al., 'Antisense Research & Applications', 1993, CRC PRESS pages 276 - 278. Further suitable modified bases are described in U.S. Patent Nos. 3,687,808, as well as 4,845,205; 5,130,302; 5,134,066; 5,175,273; 5,367,066; 5,432,272; 5,457,187; 5,459,255; 5,484,908; 5,502,177; 5,525,711; 5,552,540; 5,587,469; 5,596,091; 5,614,617; 5,750,692, and 5,681,941, each of which is herein incorporated by reference.

[00508] It is not necessary for all positions in a given oligonucleotide to be uniformly modified, and in fact more than one of the aforementioned modifications may be incorporated in a single oligonucleotide or even at within a single nucleoside within an oligonucleotide.

[00509] In certain embodiments, both a sugar and an internucleoside linkage, i.e., the backbone, of one or more nucleotides units within a multipartite construct are replaced with novel groups. The base can be maintained for hybridization with an appropriate nucleic acid target compound. One such oligomeric compound, an oligonucleotide mimic that has been shown to retain hybridization properties, is referred to as a peptide nucleic acid (PNA). In PNA compounds, the sugar-backbone of an oligonucleotide is replaced with an amide containing backbone, for example, an aminomethylglycine backbone. The nucleobases are retained and are bound directly or indirectly to aza nitrogen atoms of the amide portion of the backbone. Representative patents that teach the preparation of PNA compounds comprise, but are not limited to, U.S. Patent Nos. 5,539,082; 5,714,331; and 5,719,262, each of which is herein incorporated by reference. Further teaching of PNA compounds can be found in Nielsen et al. Science vol. 254, 1991, page 1497, which is herein incorporated by reference.

The multipartite construct of the invention can be modified to incorporate a wide variety of modified nucleotides as desired. For example, the construct may be synthesized entirely of modified nucleotides or with a subset of modified nucleotides. The modifications can be the same or different. Some or all nucleotides may be modified, and those that are modified may contain the same modification. For example, all nucleotides containing the same base may have one type of modification, while nucleotides containing other bases may have different types of modification. All purine nucleotides may have one type of modification (or are unmodified), while all pyrimidine nucleotides have another, different type of modification (or are unmodified). Thus, the construct may comprise any combination of desired modifications, including for example, ribonucleotides (2'-OH), deoxyribonucleotides (2'-deoxy), 2'-amino nucleotides (2'-NH2), 2'-fluoro nucleotides (2'-F) and 2'-0-methyl (2'-OMe) nucleotides.

In some embodiments, the multipartite construct is synthesized using a transcription mixture containing modified nucleotides in order to generate a modified construct. For example, a transcription mixture may contain only 2'-OMe A, G, C and U and/or T triphosphates (2'-OMe ATP, 2'-OMe UTP and/or 2*-OMe TTP, 2*-OMe CTP and 2*-OMe GTP), referred to as an MNA or mRmY mixture. Oligonucleotides generated therefrom are referred to as MNA oligonucleotides or mRmY oligonucleotides and contain only 2'-0-methyl nucleotides. A transcription mixture containing all 2'-OH nucleotides is referred to as an "rN" mixture, and oligonucleotides generated therefrom are referred to as "rN", "rRrY" or RNA oligonucleotides. A transcription mixture containing all deoxy nucleotides is referred to as a "dN" mixture, and oligonucleotides generated therefrom are referred to as "dN", "dRdY" or DNA oligonucleotides. Alternatively, a subset of nucleotides (e.g., C, U and/or T) may comprise a first modified nucleotide (e.g, 2'-OMe) nucleotides and the remainder (e.g., A and G) comprise a second modified nucleotide (e.g., 2'-OH or 2'-F). For example, a transcription mixture containing 2'-F U and 2'-OMe A, G and C is referred to as a "fUmV" mixture, and oligonucleotides generated therefrom are referred to as "fUmV" oligonucleotides. A transcription mixture containing 2'-F A and G, and 2'-OMe C and U and/or T is referred to as an "fRmY" mixture, and oligonucleotides generated therefrom are referred to as "fRmY" oligonucleotides. A transcription mixture containing 2'-F A and 2'-OMe C, G and U and/or T is referred to as "fAmB" mixture, and oligonucleotides generated therefrom are referred to as "fAmB" oligonucleotides.

One of skill in the art can improve pre-identified aptamer segments (e.g., variable regions or immunomodulatory regions that comprise an aptamer to a biomarker target or other entity) using various process modifications. Examples of such process modifications include, but are not limited to, truncation,
deletion, substitution, or modification of a sugar or base or internucleotide linkage, capping, and PEGylation. In addition, the sequence requirements of an aptamer may be explored through doped reselections or aptamer medicinal chemistry. Doped reselections are carried out using a synthetic, degenerate pool that has been designed based on the aptamer of interest. The level of degeneracy usually varies from about 70-85% from the aptamer of interest. In general, sequences with neutral mutations are identified through the doped reselection process. Aptamer medicinal chemistry is an aptamer improvement technique in which sets of variant aptamers are chemically synthesized. These variants are then compared to each other and to the parent aptamer. Aptamer medicinal chemistry is used to explore the local, rather than global, introduction of substituents. For example, the following modifications may be introduced: modifications at a sugar, base, and/or internucleotide linkage, such as 2'-deoxy, 2'-ribo, or 2'-0-methyl purines or pyrimidines, phosphorothioate linkages may be introduced between nucleotides, a cap may be introduced at the 5' or 3' end of the aptamer (such as 3' inverted dT cap) to block degradation by exonucleases, or a polyethylene glycol (PEG) element may be added to the aptamer to increase the half-life of the aptamer in the subject.

[00514] Additional compositions comprising a multipartite construct of the invention and uses thereof are further described below.

[00515] **Personalized multipartite constructs**

[00516] The oligonucleotide probe libraries of the invention can be used to construct personalized multipartite constructs. As an example, an oligonucleotide probe library can be used to probe a sample from a patient. The biomarker targets are identified for library members that preferentially recognize diseased cells, e.g., cancer cells, from the individual. Such methodology is described herein. See, e.g.,

**Example 23.** The variable region of one or more oligonucleotide probe to the patient's diseased cells is used to synthesize a multipartite construct of the invention. Such construct can be used as a personalized treatment for the patient. FIG. 23E illustrates a flow chart of this approach. An oligonucleotide probe library 2351 is contacted with patient sample 2352. The sample can be any useful patient sample, including without limitation tissue such as biopsy or tissue removed during surgical or other procedures, bodily fluids, frozen sections taken for histological purposes, cell cultures, and various embodiments, fractions, and components of any thereof as described herein. See, e.g., section entitled "Samples" above. Oligonucleotide probes that bind the patient sample are identified 2353. Methodology for identifying oligonucleotide probe members that bind a biological sample is described herein, and includes without limitation sequence analysis such as next generation sequencing, amplification and/or hybridization approaches. In an embodiment, flow sorting is used to separate diseased cells from the contacted patient sample and sorted and bound oligonucleotides are identified 2353. In an optional step, the biomarkers recognized by the oligonucleotide probe binders are identified 2354. This step can be optional as the target biomarker of a disease-specific oligonucleotide probe may not be necessary to synthesize a multipartite construct specific for the patient's disease. However, the identification of the target may assist in design of the multipartite construct. For example, identification of the target may allow computer modeling of the constructs _in vivo_ interactions or help to select the most promising probe from multiple
candidates, e.g., by avoiding toxicity associated with non-disease specific targets. The oligonucleotide probe binders are used to create a multipartite construct of the invention 2355. The multipartite construct can be as shown in FIG. 23A or variants thereof. In an embodiment, the variable region of an oligonucleotide probe binder is used as the variable region as shown in FIG. 23A. The multipartite construct can be administered to the patient 2356, thereby providing a personalized therapy for the patient. Further as noted in the figure, constructs comprising the variable region of an oligonucleotide probe binder can be used for other purposes, such as disease monitoring 2357. As an example of such methodology, the oligonucleotide probe binder can be used to detect the presence or absence of disease markers in the patient over time such as in an immunoassay format.

[00517] **Anti-Clq oligonucleotides**

[00518] The complement system is a part of the immune system that enhances (complements) the ability of antibodies and phagocytic cells to clear microbes and damaged cells from an organism. It is part of the innate immune system, which is not adaptable and does not change over the course of an individual's lifetime. However, it can be recruited and brought into action by the adaptive immune system. Complement activation or fixation can stimulate phagocytes to clear foreign and damaged material, induce inflammation to attract additional phagocytes, and activate the cell-killing membrane attack complex. The "classical" complement pathway is triggered by activation of the C1-complex, which occurs when Clq binds to IgM or IgG complexed with antigens. The Cl-complex is composed of 1 molecule of Clq, 2 molecules of Clr and 2 molecules of Cls, or Clqr 52. Such immunoglobulin-mediated binding of the complement uses the ability of the immunoglobulin system to detect and bind to non-self antigens. Clq can also directly identify various structures and ligands on microbial surfaces and apoptotic cells, and binds additional self proteins including C-reactive protein (CRP), HIV-1, phosphatidylserine (PS), HTLV-1, and others. Because the complement system has the potential to be extremely damaging to host tissues, its activation must be tightly regulated. The classical pathway is inhibited by C1-inhibitor, which binds to CI to prevent its activation. Clq also performs a number of non-complement functions, including without limitation such diverse functions as clearance of bacterial pathogens, induction of angiogenesis during wound healing, tolerance induction, anti-inflammatory responses and inhibiting T cell response. As a result of these diverse functions, complement and Clq play a role in diverse diseases and disorders, including without limitation autoimmune settings, pregnancy disorders, pathogen infection, aggregated proteins leading to neurodegenerative diseases, inflammation, and cancer. Deficiencies have been associated with autoimmune disease (e.g., systemic lupus erythematosus), pathogen infection and cancer. However, the tumor microenvironment may also hijack Clq to promote cell adhesion, migration and proliferation. See, e.g., Kouser et al., Emerging and Novel Functions of Complement Protein Clq. Front Immunol. 2015; 6: 317. Published online 2015 Jun 29; Son et al., Fundamental role of Clq in autoimmunity and inflammation, Immunol Res. 2015 Dec; 63(1-3): 101-106; Ghebrehiwet et al., The Clq Family of Proteins: Insights into the Emerging Non-Traditional Functions, Front Immunol. 2012; 3: 52; Nayak et al., Complement and non-complement activating functions of Clq: a prototypical innate immune molecule. Innate Immun. 2012 Apr;18(2):350-63.
[00519] Clq is a Ca²⁺ dependent hexameric complex comprised of 18 polypeptide chains, 6 of three different subunits (Clq A chain (P02745), Clq B chain (P02746), and Clq C chain (P02747)), that binds C1r and C1s to form the CI complex, the first component in classical pathway of complement. Clq globular heads form a pattern recognition complex that binds to various targets, including without limitation clustered antigen-antibody Fc immune complexes (e.g., IgG, IgM), C-reactive protein (CRP), abnormal proteins (e.g., prion and beta-amyloid), apoptotic and secondary necrotic cells, phosphatidylserine and the surface of a subpopulation of microparticles in human plasma. Recognition of IgG and IgM on a cell surface can induce a complement cascade and lead to apoptosis. See, e.g., Kishore et al., Clq and tumor necrosis factor superfamily: modularity and versatility, TRENDS in Immunology 25 (2004) 551-561; Nayak et al., Complement and non-complement activating functions of Clq: a prototypical innate immune molecule, Innate Immunity 18 (2012) 350-363. Aptamer-biotin-C1q protein conjugates have been used to induce complement mediated cell death. See, e.g., Bruno, Aptamer-biotin-streptavidin-C1q complexes can trigger the classical complement pathway to kill cancer cells, In Vitro Cell Dev Biol —Animal (2010) 46: 107-113.


[00521] Clq protein quantification has been used for disease monitoring and monoclonal antibody (mAb) production. For example, Clq mAb is used to coat ELISA plates to capture and quantify immune complexes in clinical samples. Various companies sell diagnostic kits for immune complex detection and quantitation which are based on the ability of Clq to bind well to immune complexes, but to not bind significantly to monomeric immunoglobulins. Because the DNA recognition domain of Clq does not overlap with the Fc-recognition domain, a DNA based ELISA may further allow a more accurate quantitation of immune complex detection.

[00522] Example 35 herein presents identification of an anti-Clq oligonucleotide aptamers and describes various uses thereof. The aptamers to Clq were identified via oligonucleotide probe analysis of plasma microvesicles followed by identification of oligonucleotide probe targets using gel electrophoresis and mass spectrometry analysis.

[00523] Anti-Clq aptamers of the invention can be used for multiple purposes. As described above, the invention provides a multipartite construct having a disease specific target oligonucleotide or antibody (Ab) that can recognize a target of interest and an immunomodulatory region. In an embodiment of the invention, the immunomodulatory region comprises the Clq aptamer. Such construct can act as an immunotherapeutic agent for targeted cell killing via recruitment of complement proteins and the
downstream membrane attack complex (MAC). By linking the Clq aptamer segment to another segment that specifically binds to a target of interest (e.g., a biomarker present on a cell or microvesicle of interest), the construct can bring Clq into proximity of a target. See FIG. 23C, which illustrates a construct 2331 having a segment that recognizes a Marker of Interest 2332 on a Membrane 2333, and another segment that attracts the Complement system 2334. Such binding can cause a complement cascade and induce complement mediated cell killing. This approach can be applied in multiple setting, e.g., to recognize cancer cells, gram negative bacteria, and/or viral and/or parasitic infections. For example, an anti-CD20 specific oligonucleotide can be linked with an anti-Clq specific oligonucleotide. The linkage to create the oligonucleotide - oligonucleotide construct can include but is not limited to direct synthesis with a spacer between the two oligonucleotide recognition sites. Different biomarkers can be used as the target of interest, thereby directing the complement cascade to the various targets as desired. The spacer type and size can be configured based on steric hindrance between the target protein and the Clq protein/MAC complex. As noted above, the target specific oligonucleotides/Ab can be chosen to specifically recognize various targets of interest, including but not limited to cancer cells, circulating tumor cells, immune cells (e.g., B-cells, T-cells, neutrophils, macrophage, dendritic cells) microvesicles, bacteria, viruses or parasites. In addition to Clq, the target of the complement specific oligonucleotide segment can include without limitationClr, Cls, CI, C3a, C3b, C3d, C5a, C2, C4, and cytokines.

[00524] The multipartite construct of the invention can comprise a linear molecule, a circular molecule, and/or adopt various secondary structures. FIG. 23D illustrates a construct 2340 having a Clq recruiting domain 2341 (nucleotides 37-77) and a C-met targeting domain 2342 (nucleotides 1-36). As desired, the Clq recruiting domain 2341 can comprise an anti-Clq oligonucleotide sequence of the invention. See e.g., Example 35. As shown in the illustration, the Clq recruiting domain 2341 comprises a single stranded hairpin and a complementary base pairing region, and the C-met targeting domain 2342 comprises a complementary base pairing region and a region having a more complex secondary structure. Such structures can be estimated using available software programs such as Vienna or mfold (available at mfold.rit.albany.edu). Such structural estimates can also be used to design derivatives of the sequences, e.g., by substituting, adding or deleting nucleotides in order to increase or decrease melting temperature, facilitate additions of non-natural nucleotide analogs, direct chemical modification, and/or manipulate structure or other parameters.

[00525] The invention further provides a method of molecular profiling of patient specific autoantigens by identifying autoantigens bound to complement 1 (Cl) in plasma. The invention also provides immunoassays that detect levels of Clq protein. Such assays can be any applicable immunoassay format using the anti-Clq oligonucleotide of the invention, including without limitation an oligonucleotide based ELISA, Western analysis, flow cytometry, or affinity isolation. The immunoassay can be applied to various settings, including without limitation: 1) monitor cancer patient specific immune responses before, during and after administration of immunosuppressing drugs for optimal treatment with chemotherapeutic agents; 2) monitor immune responses in patients with autoimmune disorders in response to administration.
of immunosuppressing drugs such as TNF blockers; 3) detect levels of C1q and/or anti-C1q autoantibodies in patients with systemic lupus erythematosus (SLE); 4) quantitatative C1q assay for mAb biosimilar production to satisfy the EMA biosimilar antibody guidance measures; 5) a WHO secondary test as a companion test to mAb based ELISAs; 6) as a marker for apoptosis/secondary necrosis; and 7) a C1q test for research purposes.

[00526] **Pharmaceutical Compositions**

[00527] In an aspect, the invention provides pharmaceutical compositions comprising the multipartite constructs of the invention as described above. The invention further provides methods of administering such compositions.

[00528] The term "condition," as used herein means an interruption, cessation, or disorder of a bodily function, system, or organ. Representative conditions include, but are not limited to, diseases such as cancer, inflammation, diabetes, and organ failure.

[00529] The phrase "treating," "treatment of," and the like include the amelioration or cessation of a specified condition.

[00530] The phrase "preventing," "prevention of," and the like include the avoidance of the onset of a condition.

[00531] The term "salt," as used herein, means two compounds that are not covalently bound but are chemically bound by ionic interactions.

[00532] The term "pharmaceutically acceptable," as used herein, when referring to a component of a pharmaceutical composition means that the component, when administered to an animal, does not have undue adverse effects such as excessive toxicity, irritation, or allergic response commensurate with a reasonable benefit/risk ratio. Accordingly, the term "pharmaceutically acceptable organic solvent," as used herein, means an organic solvent that when administered to an animal does not have undue adverse effects such as excessive toxicity, irritation, or allergic response commensurate with a reasonable benefit/risk ratio. Preferably, the pharmaceutically acceptable organic solvent is a solvent that is generally recognized as safe ("GRAS") by the United States Food and Drug Administration ("FDA"). Similarly, the term "pharmaceutically acceptable organic base," as used herein, means an organic base that when administered to an animal does not have undue adverse effects such as excessive toxicity, irritation, or allergic response commensurate with a reasonable benefit/risk ratio.

[00533] The phrase "injectable" or "injectable composition," as used herein, means a composition that can be drawn into a syringe and injected subcutaneously, intraperitoneally, or intramuscularly into an animal without causing adverse effects due to the presence of solid material in the composition. Solid materials include, but are not limited to, crystals, gummy masses, and gels. Typically, a formulation or composition is considered to be injectable when no more than about 15%, preferably no more than about 10%, more preferably no more than about 5%, even more preferably no more than about 2%, and most preferably no more than about 1% of the formulation is retained on a 0.22 µm filter when the formulation is filtered through the filter at 98° F. There are, however, some compositions of the invention, which are gels, that can be easily dispensed from a syringe but will be retained on a 0.22 µm filter. In one embodiment, the
term "injectable," as used herein, includes these gel compositions. In one embodiment, the term "injectable," as used herein, further includes compositions that when warmed to a temperature of up to about 40°C and then filtered through a 0.22 μm filter, no more than about 15%, preferably no more than about 10%, more preferably no more than about 5%, even more preferably no more than about 2%, and most preferably no more than about 1% of the formulation is retained on the filter. In one embodiment, an example of an injectable pharmaceutical composition is a solution of a pharmaceutically active compound (for example, a multipartite construct) in a pharmaceutically acceptable solvent. One of skill will appreciate that injectable solutions have inherent properties, e.g., sterility, pharmaceutically acceptable excipients and free of harmful measures of pyrogens or similar contaminants.

[00534] The term "solution," as used herein, means a uniformly dispersed mixture at the molecular or ionic level of one or more substances (solute), in one or more other substances (solvent), typically a liquid.

[00535] The term "suspension," as used herein, means solid particles that are evenly dispersed in a solvent, which can be aqueous or non-aqueous.

[00536] The term "animal," as used herein, includes, but is not limited to, humans, canines, felines, equines, bovines, ovines, porcines, amphibians, reptiles, and avians. Representative animals include, but are not limited to a cow, a horse, a sheep, a pig, an ungulate, a chimpanzee, a monkey, a baboon, a chicken, a turkey, a mouse, a rabbit, a rat, a guinea pig, a dog, a cat, and a human. In one embodiment, the animal is a mammal. In one embodiment, the animal is a human. In one embodiment, the animal is a non-human. In one embodiment, the animal is a canine, a feline, an equine, a bovine, an ovine, or a porcine.

[00537] The phrase "drug depot," as used herein means a precipitate, which includes the multipartite construct, formed within the body of a treated animal that releases the multipartite construct over time to provide a pharmaceutically effective amount of the multipartite construct.

[00538] The phrase "substantially free of," as used herein, means less than about 2 percent by weight. For example, the phrase "a pharmaceutical composition substantially free of water" means that the amount of water in the pharmaceutical composition is less than about 2 percent by weight of the pharmaceutical composition.

[00539] The term "effective amount," as used herein, means an amount sufficient to treat or prevent a condition in an animal.

[00540] The nucleotides that make up the multipartite construct can be modified to, for example, improve their stability, i.e., improve their in vivo half-life, and/or to reduce their rate of excretion when administered to an animal. The term "modified" encompasses nucleotides with a covalently modified base and/or sugar. For example, modified nucleotides include nucleotides having sugars which are covalently attached to low molecular weight organic groups other than a hydroxyl group at the 3' position and other than a phosphate group at the 5' position. Modified nucleotides may also include 2' substituted sugars such as 2'-0-methyl-; 2'-0-alkyl; 2'-0-alkyl; 2'-S-alkyl; 2'-S-alkyl; 2'-fluoro-; 2'-halo or 2'-azido-ribose; carbocyclic sugar analogues; α-anomeric sugars; and epimeric sugars such as arabinose, xyloses or lyxoses, pyranose sugars, furanose sugars, and sedoheptulose.
Modified nucleotides are known in the art and include, but are not limited to, alkylated purines and/or pyrimidines; acylated purines and/or pyrimidines; or other heterocycles. These classes of pyrimidines and purines are known in the art and include, pseudoisocytosine; N4,N4-ethanocytosine; 8-hydroxy-N6-methyladenine; 4-acetylcytosine, 5-(carboxyhydroxymethyl) uracil; 5-fluorouracil; 5-bromouracil; 5-carboxymethylaminomethyl-2-thiouracil; 5-carboxymethylaminomethyl uracil; dihydouracil; inosine; N6-isopentyl-adenine; 1-methyladenine; 1-methylpsuedouracil; 1-methylguanine; 2,2-dimethylguanine; 2-methyladenine; 2-methylguanine; 3-methylcytosine; 5-methylcytosine; N6-methyladenine; 7-methylguanine; 5-methylaminomethyl uracil; 5-methoxy amino methyl-2-thiouracil; β-D-mannosylqueosine; 5-methoxycarbonylmethyluracil; 5-methoxyuracil; 2 methylthio-N6-isopentenyladenine; uracil-5-oxy acetic acid methyl ester; psueouracil; 2-thiouracine; 5-methyl-2 thiouracil, 2-thiouracil; 4-thiouracil; 5-methyluracil; N-uracil-5-oxyacetic acid methylester; uracil 5-oxyacetic acid; queosine; 2-thiouracine; 5-propyluracil; 5-propylcytosine; 5-ethyluracil; 5-ethylcytosine; 5-butyluracil; 5-pentyluracil; 5-pentylcytosine; and 2,6-diaminopurine; methylpsuedouracil; 1-methylguanine; and 1-methylcytosine.

The multipartite construct can also be modified by replacing one or more phosphodiester linkages with alternative linking groups. Alternative linking groups include, but are not limited to embodiments wherein P(0)0 is replaced by P(0)S, P(S)S, P(0)NR2, P(0)R, P(0)OR’, CO, or CH2, wherein each R or R’ is independently H or a substituted or unsubstituted C1-C20 alkyl. A preferred set of R substitutions for the P(0)NR2 group are hydrogen and methoxyethyl. Linking groups are typically attached to each adjacent nucleotide through an —O— bond, but may be modified to include —N— or —S— bonds. Not all linkages in an oligomer need to be identical.

The multipartite construct can also be modified by conjugating the multipartite construct to a polymer, for example, to reduce the rate of excretion when administered to an animal. For example, the multipartite construct can be "PEGylated," i.e., conjugated to polyethylene glycol ("PEG"). In one embodiment, the PEG has an average molecular weight ranging from about 20 kD to 80 kD. Methods to conjugate an oligonucleotide with a polymer, such PEG, are well known to those skilled in the art (See, e.g., Greg T. Hermanson, Bioconjugate Techniques, Academic Press, 1966).

The multipartite constructs of the invention, e.g., such as described above, can be used in the pharmaceutical compositions disclosed herein or known in the art.

In one embodiment, the pharmaceutical composition further comprises a solvent.

In one embodiment, the solvent comprises water.

In one embodiment, the solvent comprises a pharmaceutically acceptable organic solvent. Any useful and pharmaceutically acceptable organic solvents can be used in the compositions of the invention.

In one embodiment, the pharmaceutical composition is a solution of the salt in the pharmaceutically acceptable organic solvent.

In one embodiment, the pharmaceutical composition comprises a pharmaceutically acceptable organic solvent and further comprises a phospholipid, a sphingomyelin, or phosphatidyl choline. Without wishing to be bound by theory, it is believed that the phospholipid, sphingomyelin, or phosphatidyl
choline facilitates formation of a precipitate when the pharmaceutical composition is injected into water
and can also facilitate controlled release of the multipartite construct from the resulting precipitate.
Typically, the phospholipid, sphingomyelin, or phosphatidyl choline is present in an amount ranging from
greater than 0 to 10 percent by weight of the pharmaceutical composition. In one embodiment, the
phospholipid, sphingomyelin, or phosphatidyl choline is present in an amount ranging from about 0.1 to
10 percent by weight of the pharmaceutical composition. In one embodiment, the phospholipid,
sphingomyelin, or phosphatidyl choline is present in an amount ranging from about 1 to 7.5 percent by
weight of the pharmaceutical composition. In one embodiment, the phospholipid, sphingomyelin, or phosphatidyl choline is present in an amount ranging from about 1.5 to 5 percent by weight of the
pharmaceutical composition. In one embodiment, the phospholipid, sphingomyelin, or phosphatidyl
choline is present in an amount ranging from about 2 to 4 percent by weight of the pharmaceutical
composition.

[00550] The pharmaceutical compositions can optionally comprise one or more additional excipients or
additives to provide a dosage form suitable for administration to an animal. When administered to an
animal, the multipartite construct containing pharmaceutical compositions are typically administered as a
component of a composition that comprises a pharmaceutically acceptable carrier or excipient so as to
provide the form for proper administration to the animal. Suitable pharmaceutical excipients are described
incorporated herein by reference. The pharmaceutical compositions can take the form of solutions,
suspensions, emulsion, tablets, pills, pellets, capsules, capsules containing liquids, powders, suppositories,
emulsions, aerosols, sprays, suspensions, or any other form suitable for use.

[00551] In one embodiment, the pharmaceutical compositions are formulated for intravenous or parenteral
administration. Typically, compositions for intravenous or parenteral administration comprise a suitable
sterile solvent, which may be an isotonic aqueous buffer or pharmaceutically acceptable organic solvent.
Where necessary, the compositions can also include a solubilizing agent. Compositions for intravenous
administration can optionally include a local anesthetic such as lidocaine to lessen pain at the site of the
injection. Generally, the ingredients are supplied either separately or mixed together in unit dosage form,
for example, as a dry lyophilized powder or water free concentrate in a hermetically sealed container such
as an ampoule or sachette indicating the quantity of active agent. Where multipartite construct containing
pharmaceutical compositions are to be administered by infusion, they can be dispensed, for example, with
an infusion bottle containing, for example, sterile pharmaceutical grade water or saline. Where the
pharmaceutical compositions are administered by injection, an ampoule of sterile water for injection,
saline, or other solvent such as a pharmaceutically acceptable organic solvent can be provided so that the
ingredients can be mixed prior to administration.

[00552] In another embodiment, the pharmaceutical compositions are formulated in accordance with
routine procedures as a composition adapted for oral administration. Compositions for oral delivery can
be in the form of tablets, lozenges, aqueous or oily suspensions, granules, powders, emulsions, capsules,
syrups, or elixirs, for example. Oral compositions can include standard excipients such as mannitol,
lactose, starch, magnesium stearate, sodium saccharin, cellulose, and magnesium carbonate. Typically, the excipients are of pharmaceutical grade. Orally administered compositions can also contain one or more agents, for example, sweetening agents such as fructose, aspartame or saccharin; flavoring agents such as peppermint, oil of wintergreen, or cherry; coloring agents; and preserving agents, to provide a pharmaceutically palatable preparation. Moreover, when in tablet or pill form, the compositions can be coated to delay disintegration and absorption in the gastrointestinal tract thereby providing a sustained action over an extended period of time. Selectively permeable membranes surrounding an osmotically active driving compound are also suitable for orally administered compositions. A time-delay material such as glycerol monostearate or glycerol stearate can also be used.

[00553] The pharmaceutical compositions further comprising a solvent can optionally comprise a suitable amount of a pharmaceutically acceptable preservative, if desired, so as to provide additional protection against microbial growth. Examples of preservatives useful in the pharmaceutical compositions of the invention include, but are not limited to, potassium sorbate, methylparaben, propylparaben, benzoic acid and its salts, other esters of parahydroxybenzoic acid such as butylparaben, alcohols such as ethyl or benzyl alcohol, phenolic compounds such as phenol, or quaternary compounds such as benzalkonium chlorides (e.g., benzethonium chloride).

[00554] In one embodiment, the pharmaceutical compositions of the invention optionally contain a suitable amount of a pharmaceutically acceptable polymer. The polymer can increase the viscosity of the pharmaceutical composition. Suitable polymers for use in the compositions and methods of the invention include, but are not limited to, hydroxypropylcellulose, hydroxypropylmethylcellulose (HPMC), chitosan, polyacrylic acid, and polymethacrylic acid.

[00555] Typically, the polymer is present in an amount ranging from greater than 0 to 10 percent by weight of the pharmaceutical composition. In one embodiment, the polymer is present in an amount ranging from about 0.1 to 10 percent by weight of the pharmaceutical composition. In one embodiment, the polymer is present in an amount ranging from about 1 to 7.5 percent by weight of the pharmaceutical composition. In one embodiment, the polymer is present in an amount ranging from about 1.5 to 5 percent by weight of the pharmaceutical composition. In one embodiment, the polymer is present in an amount ranging from about 2 to 4 percent by weight of the pharmaceutical composition. In one embodiment, the pharmaceutical compositions of the invention are substantially free of polymers.

[00556] In one embodiment, any additional components added to the pharmaceutical compositions of the invention are designated as GRAS by the FDA for use or consumption by animals. In one embodiment, any additional components added to the pharmaceutical compositions of the invention are designated as GRAS by the FDA for use or consumption by humans.

[00557] The components of the pharmaceutical composition (the solvents and any other optional components) are preferably biocompatible and non-toxic and, over time, are simply absorbed and/or metabolized by the body.

[00558] As described above, the pharmaceutical compositions of the invention can further comprise a solvent.
In one embodiment, the solvent comprises water.

In one embodiment, the solvent comprises a pharmaceutically acceptable organic solvent.

In an embodiment, the multipartite constructs are available as the salt of a metal cation, for example, as the potassium or sodium salt. These salts, however, may have low solubility in aqueous solvents and/or organic solvents, typically, less than about 25 mg/mL. The pharmaceutical compositions of the invention comprising (i) an amino acid ester or amino acid amide and (ii) a protonated multipartite construct, however, may be significantly more soluble in aqueous solvents and/or organic solvents. Without wishing to be bound by theory, it is believed that the amino acid ester or amino acid amide and the protonated multipartite construct form a salt, such as illustrated above, and the salt is soluble in aqueous and/or organic solvents.

Similarly, without wishing to be bound by theory, it is believed that the pharmaceutical compositions comprising (i) a multipartite construct; (ii) a divalent metal cation; and (iii) optionally a carboxylate, a phospholipid, a phosphatidyl choline, or a sphingomyelin form a salt, such as illustrated above, and the salt is soluble in aqueous and/or organic solvents.

In one embodiment, the concentration of the multipartite construct in the solvent is greater than about 2 percent by weight of the pharmaceutical composition. In one embodiment, the concentration of the multipartite construct in the solvent is greater than about 5 percent by weight of the pharmaceutical composition. In one embodiment, the concentration of the multipartite construct in the solvent is greater than about 7.5 percent by weight of the pharmaceutical composition. In one embodiment, the concentration of the multipartite construct in the solvent is greater than about 12 percent by weight of the pharmaceutical composition. In one embodiment, the concentration of the multipartite construct in the solvent is greater than about 15 percent by weight of the pharmaceutical composition. In one embodiment, the concentration of the multipartite construct in the solvent is ranges from about 2 percent to 5 percent by weight of the pharmaceutical composition. In one embodiment, the concentration of the multipartite construct in the solvent is ranges from about 2 percent to 7.5 percent by weight of the pharmaceutical composition. In one embodiment, the concentration of the multipartite construct in the solvent ranges from about 2 percent to 10 percent by weight of the pharmaceutical composition. In one embodiment, the concentration of the multipartite construct in the solvent is ranges from about 2 percent to 12 percent by weight of the pharmaceutical composition. In one embodiment, the concentration of the multipartite construct in the solvent is ranges from about 2 percent to 15 percent by weight of the pharmaceutical composition. In one embodiment, the concentration of the multipartite construct in the solvent is ranges from about 2 percent to 20 percent by weight of the pharmaceutical composition.

Any pharmaceutically acceptable organic solvent can be used in the pharmaceutical compositions of the invention. Representative, pharmaceutically acceptable organic solvents include, but are not limited to, pyrrolidone, N-methyl-2-pyrrolidone, polyethylene glycol, propylene glycol (i.e., 1,3-propylene glycol), glycerol formal, isosorbid dimethyl ether, ethanol, dimethyl sulfoxide, tetraglycol,
tetrahydrofurfuryl alcohol, triacetin, propylene carbonate, dimethyl acetamide, dimethyl formamide, dimethyl sulfoxide, and combinations thereof.

[00565] In one embodiment, the pharmaceutically acceptable organic solvent is a water soluble solvent. A representative pharmaceutically acceptable water soluble organic solvents is triacetin.

[00566] In one embodiment, the pharmaceutically acceptable organic solvent is a water miscible solvent. Representative pharmaceutically acceptable water miscible organic solvents include, but are not limited to, glycerol formal, polyethylene glycol, and propylene glycol.

[00567] In one embodiment, the pharmaceutically acceptable organic solvent comprises pyrrolidone. In one embodiment, the pharmaceutically acceptable organic solvent is pyrrolidone substantially free of another organic solvent.

[00568] In one embodiment, the pharmaceutically acceptable organic solvent comprises N-methyl-2-pyrrolidone. In one embodiment, the pharmaceutically acceptable organic solvent is N-methyl-2-pyrrolidone substantially free of another organic solvent.

[00569] In one embodiment, the pharmaceutically acceptable organic solvent comprises polyethylene glycol. In one embodiment, the pharmaceutically acceptable organic solvent is polyethylene glycol substantially free of another organic solvent.

[00570] In one embodiment, the pharmaceutically acceptable organic solvent comprises propylene glycol. In one embodiment, the pharmaceutically acceptable organic solvent is propylene glycol substantially free of another organic solvent.

[00571] In one embodiment, the pharmaceutically acceptable organic solvent comprises glycerol formal. In one embodiment, the pharmaceutically acceptable organic solvent is glycerol formal substantially free of another organic solvent.

[00572] In one embodiment, the pharmaceutically acceptable organic solvent comprises isosorbid dimethyl ether. In one embodiment, the pharmaceutically acceptable organic solvent is isosorbid dimethyl ether substantially free of another organic solvent.

[00573] In one embodiment, the pharmaceutically acceptable organic solvent comprises ethanol. In one embodiment, the pharmaceutically acceptable organic solvent is ethanol substantially free of another organic solvent.

[00574] In one embodiment, the pharmaceutically acceptable organic solvent comprises dimethyl sulfoxide. In one embodiment, the pharmaceutically acceptable organic solvent is dimethyl sulfoxide substantially free of another organic solvent.

[00575] In one embodiment, the pharmaceutically acceptable organic solvent comprises tetraglycol. In one embodiment, the pharmaceutically acceptable organic solvent is tetraglycol substantially free of another organic solvent.

[00576] In one embodiment, the pharmaceutically acceptable organic solvent comprises tetrahydrofurfuryl alcohol. In one embodiment, the pharmaceutically acceptable organic solvent is tetrahydrofurfuryl alcohol substantially free of another organic solvent.
In one embodiment, the pharmaceutically acceptable organic solvent comprises triacetin. In one embodiment, the pharmaceutically acceptable organic solvent is triacetin substantially free of another organic solvent.

In one embodiment, the pharmaceutically acceptable organic solvent comprises propylene carbonate. In one embodiment, the pharmaceutically acceptable organic solvent is propylene carbonate substantially free of another organic solvent.

In one embodiment, the pharmaceutically acceptable organic solvent comprises dimethyl acetamide. In one embodiment, the pharmaceutically acceptable organic solvent is dimethyl acetamide substantially free of another organic solvent.

In one embodiment, the pharmaceutically acceptable organic solvent comprises dimethyl formamide. In one embodiment, the pharmaceutically acceptable organic solvent is dimethyl formamide substantially free of another organic solvent.

In one embodiment, the pharmaceutically acceptable organic solvent comprises at least two pharmaceutically acceptable organic solvents.

In one embodiment, the pharmaceutically acceptable organic solvent comprises N-methyl-2-pyrrolidone and glycerol formal. In one embodiment, the pharmaceutically acceptable organic solvent is N-methyl-2-pyrrolidone and glycerol formal. In one embodiment, the ratio of N-methyl-2-pyrrolidone to glycerol formal ranges from about 90:10 to 10:90.

In one embodiment, the pharmaceutically acceptable organic solvent comprises propylene glycol and glycerol formal. In one embodiment, the pharmaceutically acceptable organic solvent is propylene glycol and glycerol formal. In one embodiment, the ratio of propylene glycol to glycerol formal ranges from about 90:10 to 10:90.

In one embodiment, the pharmaceutically acceptable organic solvent is a solvent that is recognized as GRAS by the FDA for administration or consumption by animals. In one embodiment, the pharmaceutically acceptable organic solvent is a solvent that is recognized as GRAS by the FDA for administration or consumption by humans.

In one embodiment, the pharmaceutically acceptable organic solvent is substantially free of water. In one embodiment, the pharmaceutically acceptable organic solvent contains less than about 1 percent by weight of water. In one embodiment, the pharmaceutically acceptable organic solvent contains less than about 0.5 percent by weight of water. In one embodiment, the pharmaceutically acceptable organic solvent contains less about 0.2 percent by weight of water. Pharmaceutically acceptable organic solvents that are substantially free of water are advantageous since they are not conducive to bacterial growth. Accordingly, it is typically not necessary to include a preservative in pharmaceutical compositions that are substantially free of water. Another advantage of pharmaceutical compositions that use a pharmaceutically acceptable organic solvent, preferably substantially free of water, as the solvent is that hydrolysis of the multipartite construct is minimized. Typically, the more water present in the solvent the more readily the multipartite construct can be hydrolyzed. Accordingly, multipartite construct containing pharmaceutical compositions that use a pharmaceutically acceptable organic solvent as the solvent can be
more stable than the multipartite construct containing pharmaceutical compositions that use water as the solvent.

[00586] In one embodiment, comprising a pharmaceutically acceptable organic solvent, the pharmaceutical composition is injectable.

[00587] In one embodiment, the injectable pharmaceutical compositions are of sufficiently low viscosity that they can be easily drawn into a 20 gauge and needle and then easily expelled from the 20 gauge needle. Typically, the viscosity of the injectable pharmaceutical compositions are less than about 1,200 cps. In one embodiment, the viscosity of the injectable pharmaceutical compositions are less than about 1,000 cps. In one embodiment, the viscosity of the injectable pharmaceutical compositions are less than about 800 cps. In one embodiment, the viscosity of the injectable pharmaceutical compositions are less than about 500 cps. Injectable pharmaceutical compositions having a viscosity greater than about 1,200 cps and even greater than about 2,000 cps (for example gels) are also within the scope of the invention provided that the compositions can be expelled through an 18 to 24 gauge needle.

[00588] In one embodiment, comprising a pharmaceutically acceptable organic solvent, the pharmaceutical composition is injectable and does not form a precipitate when injected into water.

[00589] In one embodiment, comprising a pharmaceutically acceptable organic solvent, the pharmaceutical composition is injectable and forms a precipitate when injected into water. Without wishing to be bound by theory, it is believed, for pharmaceutical compositions that comprise a protonated multipartite construct and an amino acid ester or amide, that the a-amino group of the amino acid ester or amino acid amide is protonated by the multipartite construct to form a salt, such as illustrated above, which is soluble in the pharmaceutically acceptable organic solvent but insoluble in water. Similarly, when the pharmaceutical composition comprises (i) a multipartite construct; (ii) a divalent metal cation; and (iii) optionally a carboxylate, a phospholipid, a phosphatidyl choline, or a sphingomyelin, it is believed that the components of the composition form a salt, such as illustrated above, which is soluble in the pharmaceutically acceptable organic solvent but insoluble in water. Accordingly, when the pharmaceutical compositions are injected into an animal, at least a portion of the pharmaceutical composition precipitates at the injection site to provide a drug depot. Without wishing to be bound by theory, it is believed that when the pharmaceutically compositions are injected into an animal, the pharmaceutically acceptable organic solvent diffuses away from the injection site and aqueous bodily fluids diffuse towards the injection site, resulting in an increase in concentration of water at the injection site, that causes at least a portion of the composition to precipitate and form a drug depot. The precipitate can take the form of a solid, a crystal, a gummy mass, or a gel. The precipitate, however, provides a depot of the multipartite construct at the injection site that releases the multipartite construct over time. The components of the pharmaceutical composition, i.e., the amino acid ester or amino acid amide, the pharmaceutically acceptable organic solvent, and any other components are biocompatible and non-toxic and, over time, are simply absorbed and/or metabolized by the body.

[00590] In one embodiment, comprising a pharmaceutically acceptable organic solvent, the pharmaceutical composition is injectable and forms liposomal or micellar structures when injected into
water (typically about 500 µL are injected into about 4 mL of water). The formation of liposomal or micellar structures are most often formed when the pharmaceutical composition includes a phospholipid. Without wishing to be bound by theory, it is believed that the multipartite construct in the form of a salt, which can be a salt formed with an amino acid ester or amide or can be a salt with a divalent metal cation and optionally a carboxylate, a phospholipid, a phosphatidyl choline, or a sphingomyelin, that is trapped within the liposomal or micellar structure. Without wishing to be bound by theory, it is believed that when these pharmaceutically compositions are injected into an animal, the liposomal or micellar structures release the multipartite construct over time.

[00591] In one embodiment, the pharmaceutical composition further comprising a pharmaceutically acceptable organic solvent is a suspension of solid particles in the pharmaceutically acceptable organic solvent. Without wishing to be bound by theory, it is believed that the solid particles comprise a salt formed between the amino acid ester or amino acid amide and the protonated multipartite construct wherein the acidic phosphate groups of the multipartite construct protonates the amino group of the amino acid ester or amino acid amide, such as illustrated above, or comprises a salt formed between the multipartite construct; divalent metal cation; and optional carboxylate, phospholipid, phosphatidyl choline, or sphingomyelin, as illustrated above. Pharmaceutical compositions that are suspensions can also form drug depots when injected into an animal.

[00592] By varying the lipophilicity and/or molecular weight of the amino acid ester or amino acid amide it is possible to vary the properties of pharmaceutical compositions that include these components and further comprise an organic solvent. The lipophilicity and/or molecular weight of the amino acid ester or amino acid amide can be varied by varying the amino acid and/or the alcohol (or amine) used to form the amino acid ester (or amino acid amide). For example, the lipophilicity and/or molecular weight of the amino acid ester can be varied by varying the R1 hydrocarbon group of the amino acid ester. Typically, increasing the molecular weight of R1 increase the lipophilicity of the amino acid ester. Similarly, the lipophilicity and/or molecular weight of the amino acid amide can be varied by varying the R3 or R4 groups of the amino acid amide.

[00593] For example, by varying the lipophilicity and/or molecular weight of the amino acid ester or amino acid amide it is possible to vary the solubility of the multipartite construct in water, to vary the solubility of the multipartite construct in the organic solvent, vary the viscosity of the pharmaceutical composition comprising a solvent, and vary the ease at which the pharmaceutical composition can be drawn into a 20 gauge needle and then expelled from the 20 gauge needle.

[00594] Furthermore, by varying the lipophilicity and/or molecular weight of the amino acid ester or amino acid amide (i.e., by varying R1 of the amino acid ester or R3 and R4 of the amino acid amide) it is possible to control whether the pharmaceutical composition that further comprises an organic solvent will form a precipitate when injected into water. Although different multipartite constructs exhibit different solubility and behavior, generally the higher the molecular weight of the amino acid ester or amino acid amide, the more likely it is that the salt of the protonated multipartite construct and the amino acid ester of the amide will form a precipitate when injected into water. Typically, when R1 of the amino acid ester is a
hydrocarbon of about C16 or higher the pharmaceutical composition will form a precipitate when injected into water and when R1 of the amino acid ester is a hydrocarbon of about C12 or less the pharmaceutical composition will not form a precipitate when injected into water. Indeed, with amino acid esters wherein R1 is a hydrocarbon of about C12 or less, the salt of the protonated multipartite construct and the amino acid ester is, in many cases, soluble in water. Similarly, with amino acid amides, if the combined number of carbons in R3 and R4 is 16 or more the pharmaceutical composition will typically form a precipitate when injected into water and if the combined number of carbons in R3 and R4 is 12 or less the pharmaceutical composition will not form a precipitate when injected into water. Whether or not a pharmaceutical composition that further comprises a pharmaceutically acceptable organic solvent will form a precipitate when injected into water can readily be determined by injecting about 0.05 mL of the pharmaceutical composition into about 4 mL of water at about 98°F and determining how much material is retained on a 0.22 μm filter after the composition is mixed with water and filtered. Typically, a formulation or composition is considered to be injectable when no more than 10% of the formulation is retained on the filter. In one embodiment, no more than 5% of the formulation is retained on the filter. In one embodiment, no more than 1% of the formulation is retained on the filter.

[00595] Similarly, in pharmaceutical compositions that comprise a protonated multipartite construct and a diester or diamide of aspartic or glutamic acid, it is possible to vary the properties of pharmaceutical compositions by varying the amount and/or lipophilicity and/or molecular weight of the diester or diamide of aspartic or glutamic acid. Similarly, in pharmaceutical compositions that comprise a multipartite construct; a divalent metal cation; and a carboxylate, a phospholipid, a phosphatidyl choline, or a sphingomyelin, it is possible to vary the properties of pharmaceutical compositions by varying the amount and/or lipophilicity and/or molecular weight of the carboxylate, phospholipid, phosphatidyl choline, or sphingomyelin.

[00596] Further, when the pharmaceutical compositions that further comprises an organic solvent form a depot when administered to an animal, it is also possible to vary the rate at which the multipartite construct is released from the drug depot by varying the lipophilicity and/or molecular weight of the amino acid ester or amino acid amide. Generally, the more lipophilic the amino acid ester or amino acid amide, the more slowly the multipartite construct is released from the depot. Similarly, when the pharmaceutical compositions that further comprises an organic solvent and also further comprise a carboxylate, phospholipid, phosphatidyl choline, sphingomyelin, or a diester or diamide of aspartic or glutamic acid and form a depot when administered to an animal, it is possible to vary the rate at which the multipartite construct is released from the drug depot by varying the amount and/or lipophilicity and/or molecular weight of the carboxylate, phospholipid, phosphatidyl choline, sphingomyelin, or the diester or diamide of aspartic or glutamic acid.

[00597] Release rates from a precipitate can be measured injecting about 50 μL of the pharmaceutical composition into about 4 mL of deionized water in a centrifuge tube. The time that the pharmaceutical composition is injected into the water is recorded as T=0. After a specified amount of time, T, the sample
is cooled to about -9° C. and spun on a centrifuge at about 13,000 rpm for about 20 min. The resulting supernatant is then analyzed by HPLC to determine the amount of multipartite construct present in the aqueous solution. The amount of multipartite construct in the pellet resulting from the centrifugation can also be determined by collecting the pellet, dissolving the pellet in about 10 µL of methanol, and analyzing the methanol solution by HPLC to determine the amount of multipartite construct in the precipitate. The amount of multipartite construct in the aqueous solution and the amount of multipartite construct in the precipitate are determined by comparing the peak area for the HPLC peak corresponding to the multipartite construct against a standard curve of multipartite construct peak area against concentration of multipartite construct. Suitable HPLC conditions can be readily determined by one of ordinary skill in the art.

[00598] **Methods of Treatment**

[00599] The pharmaceutical compositions of the invention are useful in human medicine and veterinary medicine. Accordingly, the invention further relates to a method of treating or preventing a condition in an animal comprising administering to the animal an effective amount of the pharmaceutical composition of the invention.

[00600] In one embodiment, the invention relates to methods of treating a condition in an animal comprising administering to an animal in need thereof an effective amount of a pharmaceutical composition of the invention.

[00601] In one embodiment, the invention relates to methods of preventing a condition in an animal comprising administering to an animal in need thereof an effective amount of a pharmaceutical composition of the invention.

[00602] Methods of administration include, but are not limited to, intradermal, intramuscular, intraperitoneal, intravenous, subcutaneous, intranasal, epidural, oral, sublingual, intracerebral, intravaginal, transdermal, rectal, by inhalation, or topical. The mode of administration is left to the discretion of the practitioner. In some embodiments, administration will result in the release of the multipartite construct into the bloodstream.

[00603] In one embodiment, the method of treating or preventing a condition in an animal comprises administering to the animal in need thereof an effective amount of a multipartite construct by parenterally administering the pharmaceutical composition of the invention. In one embodiment, the pharmaceutical compositions are administered by infusion or bolus injection. In one embodiment, the pharmaceutical composition is administered subcutaneously.

[00604] In one embodiment, the method of treating or preventing a condition in an animal comprises administering to the animal in need thereof an effective amount of a multipartite construct by orally administering the pharmaceutical composition of the invention. In one embodiment, the composition is in the form of a capsule or tablet.

[00605] The pharmaceutical compositions can also be administered by any other convenient route, for example, topically, by absorption through epithelial or mucocutaneous linings (e.g., oral, rectal, and intestinal mucosa, etc.).
The pharmaceutical compositions can be administered systemically or locally. The pharmaceutical compositions can be administered together with another biologically active agent.

In one embodiment, the animal is a mammal. In one embodiment the animal is a human.

In one embodiment, the animal is a non-human animal.

In one embodiment, the animal is a canine, a feline, an equine, a bovine, an ovine, or a porcine.

The effective amount administered to the animal depends on a variety of factors including, but not limited to the type of animal being treated, the condition being treated, the severity of the condition, and the specific multipartite construct being administered. A treating physician can determine an effective amount of the pharmaceutical composition to treat a condition in an animal.

In one embodiment, the multipartite construct comprises an anti-EpCAM aptamer segment. For example, the target of interest comprises EpCAM. In another embodiment, the target is selected from the group of proteins consisting of a EGFR, PBP, EpCAM, and KLK2. In another embodiment, the target is selected from the group of proteins consisting of a tetraspanin, EpCam, CD9, PCSA, CD63, CD81, PSMA, B7H3, PSCA, ICAM, STEAP, KLK2, SSX2, SSX4, PBP, SPDEF, and EGFR. In another embodiment, the target is selected from the group of proteins consisting of CD9, PSMA, PCSA, CD63, CD81, B7H3, IL 6, OPG-13, IL6R, PA2G4, EZH2, RUNX2, SERPINB3, and EpCam. In another embodiment, a target is selected from the group of proteins consisting of A33, A33 nl5, AFP, ALA, ALIX, ALP, AnnexinV, APC, ASCA, ASPH (246-260), ASPH (666-680), ASPH (A-10), ASPH (D01P), ASPH (D03), ASPH (G-20), ASPH (H-300), AURKA, AURKB, B7H3, B7H4, BCA-225, BCNP1, BDNF, BRCA, CA 125 (MUC16), CA-19-9, C-Bir, CD 1.1, CD 10, CD 174 (Lewis y), CD24, CD44, CD46, CD59 (MEM-43), CD63, CD66e, CEACD73, CD81, CD9, CDA, CDAC1 la2, CEA, C-Erb2, CerB2, CRMP-2, CRP, CXCL12, CYFRA2 1-1, DLL4, DR3, EGFR, Epcam, EphA2, EphA2 (H-77), ER, ErbB4, EZH2, FASL, FRT, FRT c.f23, GDF15, GPCR, GPR30, Gro-alpha, HAP, HBD 1, HBDB, HER 3 (ErbB3), HSP, HSP70, hVEGF2R2, iC3b, IL 6 Unc, IL1B, IL6 Unc, IL6R, IL8, IL-8, INSIG-2, KLK2, L1CAM, LAMN, LDH, MACC-1, MAPK4, MART-1, MCP-1, M-CSF, MEG-E8, MCL1, MIF, MIS RII, MMG, MMP26, MMP7, MMP9, MS4A1, MUC 1, MUC1 seq1, MUC1 seq1 la, MUC17, MUC2, Ncam, NGAL, NPGP/NPF22, OPG, OPN, p53, p53, PA2G4, PBP, PCSA, PDGFRB, PGP9.5, PIM1, PR (B), PRL, PSA, PSMA, PSME3, PTEN, R5-CD9 Tube 1, Reg IV, RUNX2, SCRN1, seprase, SERPINB3, SPARC, SPB, SPDEF, SRYN, STAT 3, STEAP 1, TF (FL-295), TFF3, TGMI, TIMP-1, TIMP1, TIMP2, TMEM2 11, TIPRSS2, TNF-alpha, Trail-R2, Trail-R4, TrKB, TROP2, Tsg 101, TWEAK, UNC93A, VEGF A, and YPSMA-1. In another embodiment, the target is selected from the group of proteins consisting of 5T4, ACTG1, ADAM 10, ADAM 15, ALDOA, ANXA2, ANXA6, APOA1, ATP 1A1, BASP1, Clorf58, C20orf1 l4, C8B, CAPZA1, CAV1, CD 151, CD2AP, CD59, CD9, CD9, CFL1, CFP, CHMP4B, CLTC, COTL1, CTNND1, CTBS, CTTSZ, CYCS, DPP4, EEF 1A1, EHD1, ENO1, FJ1R, F2, F5, FAM 125A, FNBP1L, FOLH1, GAPDH, GLBI, GPX3, HIST1H1C, HIST1H2AB, HSP90AB1, HSPA1B, HSPA8, IGSF8, ITGB 1, ITIH3, JUP, LDHA, LDHB, LUM, LYZ, MFGE8, MGAM, MMP9,
MYH2, MYL6B, NME1, NME2, PABPC1, PABPC4, PACSIN2, PCBP2, PDCD6IP, PRDX2, PSA, PSMA, PSMA1, PSMA2, PSMA4, PSMA6, PSMA7, PSMB1, PSMB2, PSMB3, PSMB4, PSMB5, PSMB6, PSMB8, PTGFRN, RPS27A, SDCBP, SERINC5, SH3GL1, SLC3A2, SMPDL3B, SNX9, TACSTD1, TCN2, THBS1, TPI1, TSG101, TUBB, VDAC2, VPS37B, YWHAQ, YWHAQ, and YWHAZ. In another embodiment, the target is selected from the group of proteins consisting of CD9, CD63, CD81, PSMA, PCSA, B7H3 and EpCam. CD9, CD63, CD81, PSMA, PCSA, B7H3 and EpCam. In another embodiment, the target is selected from the group of proteins consisting of a tetraspanin, CD9, CD63, CD81, CD63, CD9, CD81, CD82, CD37, CD53, Rab-5b, Annexin V, MFG-E8., Mucl, GPCR 110, TMEM211 and CD24. In another embodiment, the target is selected from the group of proteins consisting of A33, AFP, ALIX, ALX4, ANCA, APC, ASCA, AURKA, AURKB, B7H3, BANK1, BCNP1, BDNF, CA-19-9, CCSA-2, CCSA-3&4, CD10, CD24, CD44, CD63, CD66 CE, CD66e CE, CD81, CD9, C-Erb2, CRMP-2, CRP, CRTN, CXCL12, CYFRA21-1, DcR3, DLL4, DR3, EGFR, Epcam, EphA2, FASL, FRT, GAL3, GDF15, GPCR (GPR1 10), GPR30, GRO-1, HBD 1, HBD2, HNPI- 3, IL-1B, IL8, IMP3, LICAM, LAMN, MACC-1, MGC20553, MCP-1, M-CSF, MIG, MIF, MMP7, MMP9, MS4A1, MUC1, MUC17, MUC2, Ncam, NGAL, NNMT, OPN, p53, PCSA, PDGFRB, PRL, PSMA, PSME3, Reg IV, SCRNI, Sept-9, SPARC, SPON2, SPR, SRVN, TFF3, TGM2, TIMP-1, TMEM211, TNF-alpha, TPA, TPS, Trail-R2, Trail-R4, TrKB, TROP2, Tsg 101, TWEAK, UNC93A, and VEGFA. In another embodiment, the target is selected from the group of proteins consisting of CD9, EGFR, NGAL, CD81, STEAP, CD24, A33, CD66E, EPHA2, Ferritin, GPR30, GPR110, MMP9, OPN, p53, TMEM211, TROP2, TGM2, TIMP, EGFR, DR3, UNC93A, MUC17, EpCAM, MUC1, MUC2, TSG101, CD63, B7H3, CD24, and a tetraspanin.

[00614] The immunosuppressive target can be a tumor-derived protein found on cMVs and/or cancer cells, including without limitation TGF-β, CD39, CD73, IL10, Fasl or TRAIL.

[00615] In one embodiment, the multipartite construct can inhibit angiogenesis. In one embodiment, the multipartite construct can inhibit angiogenesis and the disease being treated is cancer. In one embodiment, the aptamer can inhibit angiogenesis and the disease being treated is a solid tumor. The multipartite construct can be a multipartite construct that inhibits a neoplastic growth or a cancer. 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; 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; medullopithelioma;
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; Sezary 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;
Waldenstrom macroglobulinemia; or Wiln'm's tumor. The compositions and methods of the invention
can be used to treat these and other cancers.

Kits

[00616] The invention also provides a kit comprising one or more reagent to carry out the methods of the
invention. For example, the one or more reagent can be the one or more aptamer, a buffer, blocker,
enzyme, or combination thereof. The one or more reagent may comprise any useful reagents for carrying
out the subject methods, including without limitation aptamer libraries, substrates such as microbeads or
planar arrays or wells, reagents for biomarker and/or microvesicle isolation (e.g., via chromatography,
filtration, ultrafiltration, centrifugation, ultracentrifugation, flow cytometry, affinity capture (e.g., to a
planar surface, column or bead), polymer precipitation, and/or using microfluidics), aptamers directed to
specific targets, aptamer pools that facilitate detection of a biomarker/microvesicle population, reagents
such as primers for nucleic acid sequencing or amplification, arrays for nucleic acid hybridization,
detectable labels, solvents or buffers and the like, various linkers, various assay components, blockers,
and the like. The one or more reagent may also comprise various compositions provided by the invention. In an embodiment, the one or more reagent comprises one or more aptamer of the invention. The one or more reagent can comprise a substrate, such as a planar substrate, column or bead. The kit can contain instructions to carry out various assays using the one or more reagent.

In an embodiment, the kit comprises an oligonucleotide probe or composition provided herein. The kit can be configured to carry out the methods provided herein. For example, the kit can include an aptamer of the invention, a substrate, or both an aptamer of the invention and a substrate.

In an embodiment, the kit is configured to carry out an assay. For example, the kit can contain one or more reagent and instructions for detecting the presence or level of a biological entity in a biological sample. In such cases, the kit can include one or more binding agent to a biological entity of interest. The one or more binding agent can be bound to a substrate.

In an embodiment, the kit comprises a set of oligonucleotides that provide a particular oligonucleotide profile for a biological sample. An oligonucleotide profile can include, without limitation, a profile that can be used to characterize a particular disease or disorder. For example, the disease or disorder can be a proliferative disease or disorder, including without limitation a cancer. In some embodiments, the cancer comprises a breast cancer.

**EXAMPLES**

**Example 1: Identification of DNA oligonucleotides that bind a target**

The target is affixed to a solid substrate, such as a glass slide or a magnetic bead. For a magnetic bead preparation, beads are incubated with a concentration of target protein ranging from 0.1 to 1 mg/ml. The target protein is conjugated to the beads according to a chemistry provided by the particular bead manufacturer. Typically, this involves coupling via an N-hydroxysuccinimide (NHS) functional group process. Unoccupied NHS groups are rendered inactive following conjugation with the target.

Randomly generated oligonucleotides (oligos) of a certain length, such as 32 base pairs long, are added to a container holding the stabilized target. Each oligo contains 6 thymine nucleotides (a "thymine tail") at either the 5 or 3 prime end, along with a single molecule of biotin conjugated to the thymine tail. Additional molecules of biotin could be added. Each oligo is also manufactured with a short stretch of nucleotides on each end (5-10 base pairs long) corresponding to amplification primers for PCR ("primer tails"). The sequences are shown absent the thymine tails or primer tails.

The oligonucleotides are incubated with the target at a specified temperature and time in phosphate-buffered saline (PBS) at 37 degrees Celsius in 500 microliter reaction volume.

The target/oligo combination is washed 1-10 times with buffer to remove unbound oligo. The number of washes increases with each repetition of the process (as noted below).

The oligos bound to the target are eluted using a buffer containing a chaotropic agent such as 7 M urea or 1% SDS and collected using the biotin tag. The oligos are amplified using the polymerase chain reaction using primers specific to 5' and 3' sequences added to the randomized region of the oligos. The amplified oligos are added to the target again for another round of selection. This process is repeated as desired to observe binding enrichment.
Example 2: Competitive assay

The process is performed as in Example 1 above, except that a known ligand to the target, such as an antibody, is used to elute the bound oligo species (as opposed to or in addition to the chaotropic agent). In this case, anti-EpCAM antibody from Santa Cruz Biotechnology, Inc. was used to elute the aptamers from the target EpCAM.

Example 3: Screening and Affinity Analysis

All aptamers generated from the binding assays described above are sequenced using a high-throughput sequencing platform, such as the Ion Torrent from Life Technologies:

Library Preparation - Aptamers were pooled after ligating barcodes and adapter sequences (Life Technologies) according to manufacturer protocols. In brief, equimolar pools of the aptamers were made using the following steps: Analyzed an aliquot of each library with a Bioanalyzer™ instrument and Agilent DNA 1000 Kit or Agilent High Sensitivity Kit, as appropriate for the final library concentration. The molar concentration (nmol/L) of each amplicon library was determined using the commercially available software (Agilent).

An equimolar pool of the library was prepared at the highest possible concentration.

The combined concentration of the pooled library stock was calculated.

The template dilution factor of the library pool was determined using the following equation:
Template Dilution Factor = (Library pool concentration [pM]) / 26 pM.

Template Preparation - Using a freshly diluted library, the aptamer pool resulting from binding assays provided above were sequenced using conventional sequencing protocols. High throughput (NextGen) sequencing methods can be used as desired.

Twenty aptamers were selected based on direct or competitive assays assessing binding to EpCAM (as described above).

Affinity Measurements - These twenty aptamers were then tested for binding affinity using an in vitro binding platform. SPR can be used for this step, e.g., a Biacore SPR machine using the T200 control software, as follows:

Dilute the antigen to a concentration of 32 nM.

Prepare necessary dilutions for kinetics, starting at 32nM prepare two-fold dilutions of antigen down to 0.5nM.

The Biacore 200 control software is programmed with the following conditions: Solution: HBS-EP+ Buffer; Number of cycles: 3; Contact time: 120s; Flow rate: 30μL/minute; Dissociation time: 300s; Solution: Glycine-HCl pH 2.5; Contact time: 120s; Flow rate: 20μL/minute; Stabilization period: 60s. The binding affinities of these aptamers are then measured using the SPR assay above, or an alternate in vitro assay assessing the aptamer for a desired function.

FIG. 5 shows the SPR data for aptamer BTX176881 (SEQ ID NO: 3). The figure comprises an association and dissociation graph of 1:1 fitting model of the biotinylated aptamers to EpCAM protein at the indicated concentrations (nM). Table 5 shows the calculated Kd values from the SPR measurements.
that are illustrated in FIG. 5. In addition, Table 5 shows the SPR data and calculated $K_d$ values for BTX 187269 (SEQ ID NO: 6) and Aptamer 4 (SEQ ID NO: 1).

<table>
<thead>
<tr>
<th>Immobilized aptamer</th>
<th>Analyte</th>
<th>Conc (nM)</th>
<th>Response</th>
<th>$K_d$ (nM)</th>
<th>Full R²</th>
<th>Full Chi²</th>
</tr>
</thead>
<tbody>
<tr>
<td>BTX176881 (SEQ ID No: 3)</td>
<td>EpCAM protein</td>
<td>500</td>
<td>0.2434</td>
<td>8.40</td>
<td>0.989322</td>
<td>0.179008</td>
</tr>
<tr>
<td>BTX187269 (SEQ ID NO: 6)</td>
<td>EpCAM protein</td>
<td>250</td>
<td>0.136</td>
<td>8.40</td>
<td>0.989322</td>
<td>0.179008</td>
</tr>
<tr>
<td>BTX187269 (SEQ ID NO: 6)</td>
<td>EpCAM protein</td>
<td>100</td>
<td>0.0776</td>
<td>8.40</td>
<td>0.989322</td>
<td>0.179008</td>
</tr>
<tr>
<td>Aptamer 4 (SEQ ID NO: 1)</td>
<td>EpCAM protein</td>
<td>500</td>
<td>0.2742</td>
<td>10.10</td>
<td>0.986276</td>
<td>0.299279</td>
</tr>
<tr>
<td>Aptamer 4 (SEQ ID NO: 1)</td>
<td>EpCAM protein</td>
<td>250</td>
<td>0.1618</td>
<td>10.10</td>
<td>0.986276</td>
<td>0.299279</td>
</tr>
<tr>
<td>Aptamer 4 (SEQ ID NO: 1)</td>
<td>EpCAM protein</td>
<td>100</td>
<td>0.0809</td>
<td>10.10</td>
<td>0.986276</td>
<td>0.299279</td>
</tr>
</tbody>
</table>

[00638] *$K_d$, $r^2$ and $Chi^2$ values by Global fitting for single reference method.

Example 4: Motif analysis

[00639] The process of Example 3 is followed to identify a high affinity aptamer to a target of interest. Once a high affinity aptamer is identified, its sequence is then analyzed using a software program to estimate its two-dimensional folding structure. Well-known sequence alignment programs and algorithms for motif identification can be used to identify sequence motifs and reduce the dimensionality of even large data sets of sequences. Further, software programs such as Vienna and mfold are well-known to those skilled in the art of aptamer selection and can be used to further group sequences based on secondary structure motifs (shared shapes). See FIG. 3A and FIG. 3B for example structure predictions. Shared secondary structure of course, does not guarantee identical three-dimensional structure. Therefore "wet-lab" validation of aptamers is still useful as no one set of in silico tools has yet been able to fully predict the optimal aptamer among a set of aptamer candidates.

Example 5: Microvesicle-based aptamer subtraction assay

[00640] Circulating microvesicles are isolated from normal plasma (e.g., from individuals without cancer) using one of the following methods: 1) Isolation using the ExoQuick reagent according to manufacturer's protocol; 2) Ultracentrifugation comprising spin at 50,000 to 150,000g for 1 to 20 hours then resuspending the pellet in PBS; 3) Isolation using the TExIS reagent from Life Technologies according to manufacturer's protocol; and 4) filtration methodology. The filtration method is described in more detail as follows:

[00641] Place syringe and filter (1.2 µm Acrodisc Syringe Filter Versapor Membrane Non-Pyrogenic Ref: 4190, Pall Life Sciences) on open 7 ml 150K MWCO column (Pierce concentrators, 150K MWCO (molecular weight cut off) 7 ml. Part number: 89922). Fill open end of syringe with 5.2 ml of filtered IX PBS prepared in sterile molecular grade water.

[00642] Pipette patient plasma (900-1000 µl) into the PBS in the syringe, pipette mix twice

[00643] Filter the plasma into the 7 ml 150K MWCO column.

[00644] Centrifuge 7 ml 150K MWCO columns at 2000 x g at 20°C (16°C to 24°C) for 1 hour.

[00645] After 1 hour spin, pour the flow-through into 10% bleach to be discarded.
Visually inspect sample volume. If plasma concentrate is above the 8.5 ml graduation on the concentrator tube, continue to spin plasma sample at 10 minute increments at 2000 x g at 20°C (16°C to 24°C) checking volume after each spin until plasma concentrate is between 8.0 and 8.5 ml.

Pipette mix slowly on the column a minimum of 6 times and adjust pipette to determine plasma concentrate volume. If volume is between 100 µl and Target Volume, transfer plasma concentrate to previously labeled co-polymer 1.5 ml tube. If volume is still greater than Target Volume, repeat the above centrifugation step.

Pour -45 ml of filtered IX PBS prepared in sterile molecular grade water into 50 ml conical tube for use in the next step.

Add the appropriate amount of filtered IX PBS to reconstitute the sample to the Target Volume.

The microvesicles produced using any of the isolation methods will comprise a mixture of vesicle types and will be various sizes with the possible exception of ultracentrifugation methods, which may favor isolating exosome size particles.

Randomly generated oligonucleotides (produced as described in Example 1 above) are incubated with the isolated normal vesicles in PBS overnight at room temperature or at 4 degrees Celsius.

The aptamers that do not bind to these vesicles are isolated by spinning down the vesicles at 50,000 to 150,000 X g for 1 to 20 hours and collecting the supernatant.

The aptamer oligonucleotides are collected from the supernatant by running the mixture over a column containing streptavidin-coated beads. These aptamers are then added to vesicles isolated from diseased patients (using the same methods as above) and incubated overnight in PBS at room temperature or 4 degrees Celsius.

The vesicles are then spun at 50,000 to 150,000 X g for 1 to 20 hours and the supernatant is discarded. The vesicles are resuspended in PBS and lysed using SDS or some similar detergent.

The aptamers are then captured by running the lysis mixture over a column of streptavidin-coated beads. The isolated aptamers are then subjected to a round of PCR to amplify the products.

The process is then repeated for a set number of times, e.g., 5 times. The remaining aptamer pool has been depleted of aptamers that recognize microvesicles found in "normal" plasma. Accordingly, this method can be used to enrich the pool in aptamers that recognize cancer vesicles. See FIG. 4.

**Example 6: Detection of Microvesicles using anti-EpCAM aptamers**

Aptamers can be used as binding agents to detect a biomarker. In this Example, aptamers are used as binding agents to detect EpCAM protein associated with microvesicles.

FIGs. 6A-D illustrate the use of an anti-EpCAM aptamer (Aptamer 4; SEQ ID NO. 1) to detect a microvesicle population in plasma samples. Plasma samples were obtained from three men with prostate cancer and three men without prostate cancer (referred to as controls or normals). Antibodies to the following microvesicle surface protein antigens of interest were conjugated to microbeads (Luminex Corp., Austin, TX): FIG. 6A) EGFR (epidermal growth factor receptor); FIG. 6B) PBP (prostatic binding protein; also known as PEBP1 (phosphatidylethanolamine binding protein 1)); FIG. 6C) EpCAM (epithelial cell adhesion molecule); and FIG. 6D) KLK2 (kallikrein-related peptidase 2). Microvesicles in
the plasma samples were captured using the bead-conjugated antibodies. Fluorescently labeled Aptamer 4 was used as a detector in the microbead assay. FIGS. 6A-D show the average median fluorescence values (MFI values) detected for the bead-captured and Aptamer 4 detected microvesicles. Each plot individually shows the three cancer (C1-C3) and three normal samples (N1-N3). These data show that, on average, the prostate cancer samples have higher levels of microvesicles containing the target proteins than the normals.

Example 7: Negative and Positive Selection of Aptamers

[00659] Aptamers can be used in various biological assays, including numerous types of assays which rely on a binding agent. For example, aptamers can be used instead of antibodies in immune-based assays. This Example provides an aptamer screening method that identifies aptamers that do not bind to any surfaces (substrates, tubes, filters, beads, other antigens, etc.) throughout the assay steps and bind specifically to an antigen of interest. The assay relies on negative selection to remove aptamers that bind non-target antigen components of the final assay. The negative selection is followed by positive selection to identify aptamers that bind the desired antigen.

[00660] Preliminary experiments were done with five DNA aptamer libraries with 10^{15} sequences each and variable lengths (60, 65, 70, 75, 80-mers) were pre-amplified and strand separated so that forward strand (non-biotinylated) serves as an aptamer. Multiple rounds of negative selection and positive selection were performed. Before each round, the recovered aptamer products were PCR amplified and strand separated using standard methodology. Selections were performed as follows:

[00661] Negative selection

[00662] 1. Prepare bead negative Selection Mix: Incubate 1200 non-magnetic beads with standard blocking agent for 20 min.

[00663] 2. Add 50 μl of aptamer library (5 libraries total) to a PCR strip tube with 4.5 μl of each bead mixture. Incubate for 2 h at 37°C with agitation at 550 rpm.

[00664] 3. Pre-wet filter plate (1.2 μl, Millipore) with PBS-BN buffer. Add 150 μl PBS-BN.

[00665] 4. Transfer samples from the PCR strip tubes to the filter plate, incubate for 1 h at room temperature with agitation at 550 rpm.

[00666] 5. Collect flow-through from filter plate into a collection (NBS) plate using a vacuum manifold.

[00667] 6. Concentrate and clean samples to remove excess materials as desired.

[00668] The negative selection process is repeated up to 6-7 times.

[00669] Positive selection

[00670] Before starting, conjugate the protein biomarkers of interest (here, SSX4, SSX2, PBP, KLK2, SPDEF) to desired non-magnetic microbeads using conditions known in the art. The recombinant purified starting material included: SPDEF recombinant protein from Novus Biologicals (Littleton, CO, USA), catalog number H00025803-P01; KLK2 recombinant protein from Novus, catalog number H00003817-P02; SSX2 recombinant protein from Novus, catalog number H00006757-P01; PBP recombinant protein
from Fitzgerald Industries International (Action, MA, USA), catalog number 30R-1382; SSX4 recombinant protein from GenWay Biotech, Inc. (San Diego, CA, USA), catalog number GWB-E219AC.

1. Bead blocking: Incubate a desired number of each bead (8400 x number of aptamer libraries) x an average factor of (1.2)) with a starting block for 20 min.

2. Mix 50 µl of each aptamer library sample to PCR strip tubes add 2.3 µl of bead sample with particular antigen. Incubate for 2 h at 37°C with agitation at 550 rpm.

3. Pre-wet filter plate (1.2 µl, Millipore) with PBS-BN buffer. Add 150 µl PBS-BN.

4. Transfer samples from the PCR strip tubes to the filter plate, incubate for 1 h at room temperature with agitation at 550 rpm.

5. Wash 3x with PBS-BN, add 50 µl of PBS and collect samples from the top of the filter to the 1.5 ml tubes.

The positive selection is repeated up to 16 times. Certain rounds of positive selection have additional steps to treat the recovered RNA (i.e., remaining aptamer candidates) as follows:

Round 8 of positive selection was modified as follows:

1. After the third wash (PBS-BN) 25 µl of sample were collected from the top of the filter into 1.5 ml tubes.

2. The filter plate was incubated at 45°C for -10 min and washed immediately using vacuum. The plate was washed three more times with PBS-BN.

3. 50 µl of PBS were added to the plate and step 2 was repeated.

4. After the last wash, 25 µl of PBS was added to the wells. The samples were mixed well and collected from the top of the filter into 1.5 ml tubes.

Round 9 of positive selection was modified as follows:

1. After the final wash in step 5), 5 µg/ml Streptavidin-PE was added to the aptamer mixture and incubated for 30 min at room temperature with agitation at 550 rpm.

2. Samples on filter plate were washed 3x with PBS-BN (+ additional 500 mM NaCl).

3. One additional wash with regular PBS-BN was performed.

4. 50 µl of PBS was added to the samples followed by collection as above into 1.5 ml tubes.

5. Samples stored at -20°C.

Round 14 of positive selection was modified as follows:

Before start this round, the antigens of interest (SSX4, SSX2, PBP, KLK2, SPDEF) were conjugated to carboxylated magnetic beads using methods known in the art.

1. Bead blocking: take desired number of each non-magnetic bead (3000 x number of aptamer libraries) x an average factor of (1.2), add starting block (3:1, blocking per 1200 beads), make 5 mixes of 4 antigens and supplement each with different target antigen conjugated to magnetic beads (see Table 6 below, wherein the antigens are conjugated to non-magnetic beads except as indicated), incubate 20 min.

<table>
<thead>
<tr>
<th>Blocking Mix</th>
<th>Non-magnetic bead antigens</th>
<th>Magnetic bead antigens</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 6: Bead blocking mixtures
Add 50 μl of aptamer libraries to PCR strip tubes, add bead mixtures with target antigen on magnetic beads to the tubes with pre-selected corresponding aptamer library and incubate for 2 h at 37°C with agitation at 550 rpm.

Pre-wet filter plate with PBS-BN buffer, add 150 μl PBS-BN.

Transfer samples from PCR strip tubes to filter plate, incubate 1h room temperature with agitation at 550 rpm.

After last (standard) wash, add 5μg/ml Streptavidin-PE, incubate for 30 min room temperature with agitation at 550 rpm;

Wash 3x with PBS-BN (+ additional 500 mM NaCl).

Perform one additional wash with regular PBS-BN.

50 μl of PBS was added to the samples followed by collection as above into 1.5 ml tubes.

Remove the magnetic beads using a magnetic stand, and replace with fresh PBS buffer.

Samples stored at -20°C for subsequent DNA extraction and strand separation.

Optional steps implemented in the later round of positive selection are intended to increase stringency of aptamer binding (e.g., increased heat or salt concentration).

**Example 8: Discovery and Characterization of anti-EpCAM aptamers**

In this Example, an aptamer to EpCAM identified using the technique in the Example above is characterized. After selection for a pool of EpCAM binding aptamers as described above, the aptamer library was sequenced using the Ion Torrent standard protocol (Life Technologies, Carlsbad, CA). Lead candidates were selected as those having (a) high abundant motifs across all read sequences with full expected length product and (b) strong secondary structure (**FIG. 7B**).

Aptamers were selected for EPCAM protein conjugated to MicroPlex beads in competition with SSX4, SSX2, PBP, KLK2, and SPDEF recombinant proteins. A portion of the aptamers was selected in initial rounds against EpCAM that was attached to an Fc tag, and after round 8 the selection was switched to EPCAM with a Histidine tag. Another portion of the aptamers was selected in initial rounds against EpCAM that was attached to a Histidine tag, and after round 8 the selection was switched to EPCAM with an Fc tag. Methods of using Fc and histidine tags for protein purification and capture are known to those of skill in the art.

**Aptamer Characterization**

CAR003 is an aptamer candidate identified using the above methodology. As an RNA aptamer, CAR003 with alternate tail sequence has the following RNA sequence (**SEQ ID NO. 5**):

5`-auccagagug acgcagcagu cuuuucugau ggacacgugg uggucuagua ucacuaagcc accgugucca-3`
CAR003 was further characterized. EpCAM aptamer CAR003 is modified as desired by attachment of a biotin moiety on the 5’end or 3’ end. The biotin can be used to bind the aptamer using a streptavidin-biotin system, e.g., for labeling, capture and/or anchoring. FIG. 7B illustrates the optimal secondary structure of CAR003 with a minimum free energy (ΔG) of -30.00 kcal/mol. For purposes of illustration, the aptamer is shown as an RNA aptamer (SEQ ID NO. 5) corresponding to the CAR003 DNA sequence (SEQ ID NO. 4).

Synthesis and purification. The selected CAR003 aptamer was re-synthesized using AKTA OligoPilot 100 Synthesizer (GE Healthcare Life Sciences Corp., Piscataway, NJ) with a 3'-Biotin and final detritylation. The product was purified with anion exchange chromatography by FPLC. Several fractions after FPLC were combined as shown as the indicated Pools 1-3 in FIG. 7C. The figure comprises an FPLC chromatogram with all product and fractions assigned in pools after checking quality on gel. FIG. 7D illustrates a SYBR GOLD stained gel with different FPLC fractions of CAR003 aptamer after synthesis. Different fractions were combined in pools based on amount of unfinished chains in order high to low (pool 1 - pool 3). The pools 1-3 correspond to those indicated in FIG. 7C.

CAR003 aptamer characterization. Purified CAR003 aptamer was tested for binding to recombinant EPCAM protein with a polyhistidine tag (“His tagged”) using the following internally developed assay. Anti-His tag conjugated beads were mixed with EPCAM-His tagged protein. The aptamer to be tested was labeled with streptavidin-phycoerythrin (SA-PE). The EpCAM-beads and SA-PE labeled aptamers were mixed. Binding was determined as median fluorescent value in a bead assay as described herein. MFI values (FIG. 7E-F) increase with increased binding of the SA-PE labeled aptamer to the recombinant EpCam. FIG. 7E-F illustrate binding of CAR003 to EPCAM protein in 25 mM HEPES with PBS-BN (PBS, 1% BSA, 0.05% Azide, pH 7.4) (FIG. 7E) or in 25 mM HEPES with 1 mM MgCl₂ (FIG. 7F). EPCAM aptamer Aptamer 4 (see above) was used for comparison. As shown in the figures, CAR003 pool 3 more efficiently binds its target in the presence of MgCl₂ (FIG. 7F) than in the presence of BSA (FIG. 7E).

To understand its performance further, CAR003 binding was tested in the presence of both BSA and MgCl₂ in various buffers. FIG. 7G illustrates CAR003 binding to EpCAM in the indicated salts with and without addition of bovine serum albumin (BSA). Again, CAR003 binding to EpCAM is more efficient when BSA is not present. Additionally, 150 mM NaCl was tested but did not appear to improve CAR003 performance over MgCl₂.

Another factor which might influence performance of aptamer is denaturing with different salt compositions. FIG. 7H illustrates the effect of denaturing on CAR003 binding to EPCAM protein. As seen from the chart, denaturing of the aptamer has a positive effect on CAR003 binding to Epcam similar as the effect on CAR003 from MgCl₂. However, denaturing in the presence of MgCl₂ may not synergistically improve binding of CAR003 to EpCAM. Interestingly, CAR003 appeared more stable compared to control Aptamer 4 in the conditions tested.

CAR003 affinity to Epcam in the bead assay environment was assessed in the same assay as above with aptamer titrated across a constant input of antigen. FIG. 7I illustrates titration of aptamers
against EPCAM recombinant protein (constant input 5 µg). Under the conditions tested, Aptamer 4 had a higher affinity to EPCAM protein compared to CAR003 as suggested from saturation level starting at 5 µg of aptamer input.

In order to evaluate specificity of CAR003, it was tested using Western Blot against EPCAM recombinant protein, and controls comprising bovine serum albumin (BSA) and human serum albumin (HSA). **FIG. 7J** illustrates a Western blot with CAR003 aptamer versus EPCAM his-tagged protein, BSA, and HSA (5 µg each). The gel was blocked 0.5% F127 and probed with -50 µg/ml CAR003 biotinylated aptamer, fraction 3. The blot was visualized with NeutrAvidin-HRP followed by SuperSignal West Femto Chemiluminescent Substrate. The Western blot probed with CAR003 aptamer showed a clear preference of the aptamer to EPCAM protein over the albumins.

**[00713]** CAR003 test with plasma samples. Plasma samples from five prostate cancer and five normal subjects were tested with CAR003 to detect microvesicles using bead-conjugated proteins to capture the microvesicles and SA-PE labeled aptamer to detect the vesicles. SA-PE labeled Aptamer 4 detector was used as control. Fold changes of Cancer over Normal are shown in **Table 7**. The fold changes are shown without normalization ("Raw") or with normalization to a negative control. The vesicles were captured with bead conjugated antibodies to SSX4, PBP, SPDEF, EPCAM, KLK2 and SSX2 as indicated.

**Table 7: CAR003 to detect microvesicles**

<table>
<thead>
<tr>
<th></th>
<th>SSX4</th>
<th>PBP</th>
<th>SPDEF</th>
<th>EPCAM</th>
<th>KLK2</th>
<th>SSX2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Raw</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standard protocol</td>
<td>0.87</td>
<td>0.39</td>
<td>0.71</td>
<td>0.63</td>
<td>0.93</td>
<td>0.87</td>
</tr>
<tr>
<td>Incubation in presence of 1mM MgCl2 and absence of PBS-BN</td>
<td>0.77</td>
<td>0.39</td>
<td>0.69</td>
<td>0.6</td>
<td>0.91</td>
<td>0.81</td>
</tr>
<tr>
<td>Aptamer 4 control (standard protocol)</td>
<td>0.78</td>
<td>0.67</td>
<td>0.81</td>
<td>0.72</td>
<td>1.19</td>
<td>0.79</td>
</tr>
<tr>
<td><strong>Normalized to Negative control</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standard protocol</td>
<td>1.49</td>
<td>0.84</td>
<td>1.13</td>
<td>1.17</td>
<td>1.5</td>
<td>1.38</td>
</tr>
<tr>
<td>Incubation in presence of 1mM MgCl2 and absence of PBS-BN</td>
<td>1.27</td>
<td>0.83</td>
<td>1.08</td>
<td>1.1</td>
<td>1.46</td>
<td>1.29</td>
</tr>
<tr>
<td>Aptamer 4 control (standard protocol)</td>
<td>1.18</td>
<td>0.96</td>
<td>1.11</td>
<td>1.04</td>
<td>1.82</td>
<td>1.1</td>
</tr>
</tbody>
</table>

Under the conditions tested, the samples detected with CAR003 had lower MFI values as compared to detection with Aptamer 4, whereas CAR003 had a better signal-to-noise ratio and showed better separation between cancer and normal samples with SSX4, SPDEF, EPCAM and SSX2 capturing markers.

**[00715]** Control Aptamer
The characteristics of the aptamers (size, stability, binding affinity and specificity, etc) can be compared against control aptamers specific to EpCAM or other targets. For example, the aptamers are compared to the anti-VEGF aptamer 5’ biotin-CA ATT GGG CCC GTC CGT ATG GTG GGT (SEQ ID NO. 7) as described in Kaur and Yung, 2012.

References:

Example 9: Aptamer Target Identification
[00718] In this Example, aptamers conjugated to microspheres are used to assist in determining the target of two aptamers identified by library screening methods as described above. The general approach is shown in FIG. 9. The approach is used to verify the targets of CAR003, an aptamer identified by library screening to recognize EpCAM. See description above for CAR003. In this approach, the sequence of CAR003 is randomly rearranged before linkage to the microspheres. The microspheres are used as controls to bind to targets that are similar but not identical to the intended target molecule.
[00719] The protocol used is as follows:
[00720] 1) The candidate aptamers (here, CAR003) and negative control aptamers (here, randomly arranged CAR003) are synthesized with modifications to allow capture (here, the aptamers are biotinylated) and crosslinking (here, using the Sulfo-SBED Biotin Label Transfer Reagent and Kit, Catalog Number 33073 from Thermo Fisher Scientific Inc., Rockford, IL, to allow photocrosslinking).
[00721] 2) Each of the aptamers is individually mixed with microvesicles having the target of interest (here, BrCa cell line microvesicles).
[00722] 3) After incubation to allow the aptamers to bind target, ultraviolet light is applied to the mixtures to trigger crosslinking of the aptamers with the microvesicle targets.
[00723] 4) The microvesicles are lysed, thereby releasing the crosslinked aptamer-target complex into solution.
[00724] 5) The crosslinked aptamer-target complexes are captured from solution using a streptavidin coated substrate.
6) The crosslinked aptamer-target complexes for each aptamer are run individually on SDS-PAGE gel electrophoresis. The captured protein targets are visualized with Coomassie Blue staining.

7) The crosslinking and binding steps may be promiscuous so that multiple bands including the intended target but also random proteins will appear on each of the gels. The intended target will be found in a band that appears on the gel with the candidate aptamer (here, CAR003) but not the related negative control aptamers (here, randomly arranged CAR003). The bands corresponding to the target are excised from the gel.

8) Mass spectrometry (MS) is used to identify the aptamer target from the excised bands.

**Example 10: Aptamers to Breast Cancer (BrCa) derived microvesicles**

In this Example, an aptamer library is screened to identify aptamers that distinguish between microvesicles circulating in the blood of breast cancer patients and microvesicles circulating in the blood of healthy, control individuals (i.e., without breast cancer).

Microvesicles were isolated from plasma of a pool of 60 breast cancer patients (BrCa+). Microvesicles were also isolated from pool of 60 non-cancer samples (BrCa-). Microvesicles were isolated from the plasma using ultracentrifugation (120,000 x g). Microvesicles were in the pellet from the ultracentrifugation. The supernatant from the ultracentrifugation was saved to use as a control. The microvesicles from both sample types were conjugated to MagPlex beads (Luminex Corp, Austin TX). Optionally, the isolated microvesicles are incubated with anti-HSA/IgG/Fibrinogen beads to remove these highly abundant blood proteins. However, the conjugation step can be optimized to favor conjugation of the microvesicles such that removal of highly abundant proteins may be less of an issue.

The aptamer library used consisted of a 2'F SUL1 RNA aptamer library. The sequence is 5'-GGGAGGACGAUCGCGG-N4 O-CAGACGCUCGCUGAGGAUCCGAGA -3' (SEQ ID NO. 8). The aptamer library consists of three sections: Forward primer - 15 nucleotides, variable region - 40 nucleotides; reverse primer - 25 nucleotides. All pyrimidines (C and U) were 2'Fluoro modified.

The aptamer library was incubated with either the cancer or control microvesicle-conjugated beads. Thirteen rounds of positive selection for aptamers that bind the microvesicles were performed in parallel for both types of samples. See **Example 11** below for detailed protocol of the positive selection steps. Negative selection was not performed.

A number of representative sequences obtained from these procedures are shown in Table 8. The sequences in the table were identified in the aptamer pools from selection against BrCa microvesicles but were not in the aptamer pools selected against non-cancer samples. In Table 8, the sequences are shown 5' to 3' from left to right, wherein each complete sequence consists of a 5' leader sequence 5'-GGGAGGACGAUCGCGG (SEQ ID NO. 9) followed by the indicated Variable Sequence followed by the 3' tail sequence 5'-CAGACGCUCGCUGAGGAUCCGAGA (SEQ ID NO. 10). Each sequence is derived from a library having a leader and tail (see description above) with a variable sequence between. It is understood that the nucleotide sequences that are disclosed in Table 8 can also be modified to the extent that resulting modifications result in an aptamer having about 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 96, 97, 98, and 99
percent homology to the disclosed sequence and retain the functionality of binding to microvesicle antigens or functional fragments thereof.

Table 8: BrCa microvesicle aptamer candidate sequences

<table>
<thead>
<tr>
<th>ID</th>
<th>SEQ ID NO.</th>
<th>Variable Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>BRCA_APT1 RNA</td>
<td>11</td>
<td>CGCGUCUCCCCGCAUUGCCGCAUUUAUACAUUAA</td>
</tr>
<tr>
<td>BRCA_APT2 RNA</td>
<td>13</td>
<td>GUCCGGAACGCCUCAGCAUCCGCAUAUUGAUGUCUG</td>
</tr>
<tr>
<td>BRCA_APT3 RNA</td>
<td>15</td>
<td>GUCAAGUGUAACGGCGAAUCCGCAUAGCAUGUAA</td>
</tr>
<tr>
<td>BRCA_APT4 RNA</td>
<td>17</td>
<td>CACUAUGCUGUUUGCCGCAUUGGGAUUGGCAUUGC</td>
</tr>
<tr>
<td>BRCA_APT5 RNA</td>
<td>19</td>
<td>UCUUCGCAUCUGGUCGCAUACAGAAUACGAUAAACAUA</td>
</tr>
<tr>
<td>BRCA_APT6 RNA</td>
<td>21</td>
<td>GACUCACGUGCCCUUUUGGUAAGGCUUUUUGAACAAGC</td>
</tr>
<tr>
<td>BRCA_APT7 RNA</td>
<td>23</td>
<td>UAAUCGCGAGUGACACUAACUAUGAUGACGCUGGAA</td>
</tr>
</tbody>
</table>

Each sequence in Table 8 is synthesized in two variants for further investigation: 5' biotinylated and 3' biotinylated. This provides aptamer variants that can be captured at the 5' end or the 3' end as desired. The aptamers are further synthesized with each pyrimidine (C and U) 2'Fluoro modified.

The DNA sequence corresponding to each RNA sequence in Table 8 is provided in the sequence listing, where the DNA sequence directly follows its corresponding RNA sequence. For example SEQ ID NO. 12 is the DNA sequence corresponding to RNA sequence SEQ ID NO. 11, etc. The DNA forms of the aptamers are synthesized for further characterization as well.

The aptamers above were identified using positive selection for aptamers which recognize BrCa and non-BrCa microvesicles conjugated to microspheres.

Example 11: Aptamer library selection protocol

This Example provides the protocol for SUL1 RNA library selection performed in the Example above. The protocol can be followed for other aptamer libraries and sample input as desired.

Preparation

The working space is cleaned with 80% EtOH before working.

Beads are MagPlex beads (Luminex Corp., Austin, TX). Other beads can be substituted as desired.

Buffers/Reagents to Prepare:

- MilliQ water
- 100mM MgCl2
- 5x Transcription Buffer (200mM Tris pH 7.9)
- 1x PBS
- 1x PBS with 3mM MgCl2
- 10x PBS
- Selection buffer (1x PBS with 0.1% BSA and 3mM MgCl2)
Before starting with selection, remove the bead storage buffer, and wash beads with 1x PBS with 3mM MgCl₂ in 200µL total in all 4 tubes. 200,000 beads per selection are used.

**Binding 2F SUL1 RNA pool to microvesicle coated magnetic beads**

**Abbreviations:** TK- Transcription; NTC- No template control.

**Steps:**

1. **1st Round:** Mix 1nmol purified 2F SUL1 RNA with 20 µL of resuspended beads (conjugated with microvesicle). 100µL of 10x PBS +1% BSA, 3µL 100 mM MgCl₂, and 47µL H₂O. This gives a final concentration of 1x PBS, 0.1% BSA, 3mM MgCl₂.

   1.1 The addition of MgCl₂ in this step gives a concentration of 3mM MgCl₂. This is the binding concentration for the entire process.

   1.2 **Following Rounds:** Mix 20 µL of transcription product (15 mM MgCl₂ inside) with 20 µL of washed microvesicle coated beads, plus 9µL 10x PBS with 1% BSA, 5µL H₂O. No additional MgCl₂ is needed because the MgCl₂ in the diluted transcription product (TK) provides a final concentration of 3mM MgCl₂.

2. Incubate for 30 min at 37 °C, shake at 1000rpm, and pipet mix every 10 minutes.

3. Wash the beads:

   3.1 **One washing cycle comprises:**

      3.1.1 Remove the beads from the magnet.
      3.1.2 Resuspend beads in 100µL 1x PBS +3mM MgCl₂ off the magnet.
      3.1.3 Incubate sample for 30 seconds off of the magnet.
      3.1.4 Place the sample back onto the magnet, and wait until the beads are on the side.
      3.1.5 Remove and discard the supernatant.
      3.1.6 Resuspend in 100µL 1x PBS +3 mM MgCl₂ +0.1% BSA off of the magnet.
      3.1.7 Incubate sample for 3 minutes off of the magnet.
      3.1.8 Place the sample back onto the magnet, and wait until the beads are on the side.
      3.1.9 Remove and discard the supernatant.

   3.2 **1st Round:** Place bead mixture on a magnet and remove the supernatant. Wash once with 100 µL 1x PBS +3 mM MgCl₂ +0.1% BSA (by pipette mixing the beads), and discard buffer.

   3.3 **Following Rounds:** Increase the washing steps every second round by one more washing step up to 3 washing steps.

4. Add 55 µL MilliQ water to the bead sample.

5. Elute the RNA by incubating the bead sample for 5 min at 80 °C

   5.1 Check if there is 50 µL, if not spin the sample down to spin down the condensed water off the top.

   5.2 Transfer the supernatant to a new vial. Work quickly to avoid the strands rebinding the beads.

   5.2.1 Use 50 µL eluate for the following RT-PCR and store the rest at -20°C

**RT-PCR of recovered aptamer candidates**

**Tips**

- The rest of the RT-PCR sample and the TK-PCR sample is stored at -20 °C
- RNA can be stored at 4°C for ~1h
- RT-PCR product can be stored overnight at 4°C
- Proceed to the next selection cycle for optimal RNA quality immediately after transcription.
- Avoid vortexing RNA
- Mix on ice
- Use 0.5ml PCR tubes
- Every RT-PCR should have a no-template control (NTC) with water instead of template
- Do not freeze-thaw DTT more than one time
6. Prepare a Master Mix before the first round per Table 9, check it with 0.5 pmol RNA and store aliquots of 48 µl at -20°C until usage.

Table 9: RT-PCR master mix

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Volume (µl)/reaction</th>
<th>Final concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>5x Colorless GoTaq Flexi Buffer Promega cat# M890A</td>
<td>20</td>
<td>1x</td>
</tr>
<tr>
<td>5 x first strand buffer (Invitrogen) lot# 1300427</td>
<td>4</td>
<td>0.2 x</td>
</tr>
<tr>
<td>100 mM DTT</td>
<td>2</td>
<td>2 mM</td>
</tr>
<tr>
<td>100 µM SUL1 F primer</td>
<td>1</td>
<td>1 µM</td>
</tr>
<tr>
<td>100 µM SUL1 R primer</td>
<td>1</td>
<td>1 µM</td>
</tr>
<tr>
<td>100 mM MgCl₂</td>
<td>1.5</td>
<td>1.5 mM</td>
</tr>
<tr>
<td>25 mM (each) dNTPs</td>
<td>1.2</td>
<td>300 µM</td>
</tr>
<tr>
<td>MilliQ water</td>
<td>17.3</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>48</td>
<td></td>
</tr>
</tbody>
</table>

7. Add 50 µl MilliQ water as negative control (NTC) (pipette this first) or 50 µl selection eluate. Pipet mix.

8. Incubate at 65°C for 5 min.

9. After cooling to 4°C, add:
   9.1 1 µl Superscript II Reverse Transcriptase (Invitrogen, cat# 18064) (200 U/µl)
   9.2 1 µl GoTaqFlexi DNA polymerase (5 units/µl) Promega cat# M8305.

PCR -Program (SARTPCR)

   a) 10 min 54°C
      (This step is only for reverse transcriptase, should more rounds be needed, do not repeat step A.)
   b) 1 min 95°C
   c) 1 min 60°C
   d) 1 min 72°C

10. Cycle steps b-d for
    10.1 1st round b-d 4 cycles. Run 5 µL PCR products on a 4% agarose gel.
    10.1.1 Subsequent rounds: The amount of RNA is decreased after the first round, leading to an increase in required PCR-cycles. To determine the number of cycles needed each time, check the band intensity from the agarose gel from the previous round of selection. Use that number of cycles to start the next round of RT-PCR. Note: Always check results on an agarose gel.
    10.1.1.1 Agarose gel results: product band should be seen at the target length. The band intensity should be about the same as the 50bp ladder band (if not a little less intense). If the band is not intense enough (barely visible), cycle an appropriate amount more and re-check on an agarose gel.

[00747] Transcription

[00748] All mixing performed on ice. Prepare transcription Master Mix per Table 10 and store aliquots of 85.7 µl at -20°C until use.

11. Verify pH of stock 200 mM Tris pH 7.9 before use. A change in pH over time may cause problems with the transcription.
Table 10: Transcription (TK) Master Mix for SUL1 library

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Volume (µl) for one reaction</th>
<th>Volume (µl) for 20 reactions</th>
<th>Final concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>5x Transcription buffer (200 mMTris, pH 7.9)</td>
<td>20</td>
<td>400</td>
<td>1x</td>
</tr>
<tr>
<td>100 mM DTT</td>
<td>5</td>
<td>100</td>
<td>5 mM</td>
</tr>
<tr>
<td>100mM ATP</td>
<td>1</td>
<td>20</td>
<td>1 mM</td>
</tr>
<tr>
<td>100mM GTP</td>
<td>1</td>
<td>20</td>
<td>1 mM</td>
</tr>
<tr>
<td>100mM 2’F-dUTP</td>
<td>3</td>
<td>60</td>
<td>3 mM</td>
</tr>
<tr>
<td>100mM 2’F-dCTP</td>
<td>3</td>
<td>60</td>
<td>3 mM</td>
</tr>
<tr>
<td>100 mM MgCl₂</td>
<td>15</td>
<td>300</td>
<td>15 mM</td>
</tr>
<tr>
<td>MilliQ water</td>
<td>37</td>
<td>740</td>
<td></td>
</tr>
<tr>
<td>Total volume</td>
<td>85</td>
<td>1700 µl</td>
<td></td>
</tr>
</tbody>
</table>

12. Add 10 µl RT-PCR product to the mastermix.
13. Add 1 µl RNasin (40 units/µl)
13.1 Promega Recombinant RNasin Ribonuclease Inhibitor cat# N25 15/N25 11
14. Add 4 µl T7 Y639F mutant polymerase (2511/µl use: 100U total per reaction)
15. Perform the reaction for 30 min at 37 °C
16. Use the transcription-product directly for the next selection round. If the next step is not feasible, freeze transcription product at -20°C.

[00749] Subsequent Rounds

[00750] Repeat the bead incubation, the RT-PCR and transcription as often as needed. Try to have similar band intensity of the RT-PCR product for the sample in all rounds as noted above.

[00751] Binding Assay

[00752] A binding assay is performed after desired rounds of selection to determine to assess non-specific binding of cancer selected aptamers to control beads (conjugated to supernatant from plasma ultracentrifugation, see above) and likewise for non-cancer control samples. Binding assays can also be performed to assess binding of selected aptamers against the intended target microve sides.


[00754] Final concentration of selection buffer: 1x PBS+ 3mM MgCl₂ + 0.01% BSA pH 7.4

[00755] Wash buffer: 1x PBS + 3mM MgCl₂ pH 7.4

1. Remove microvesicle samples from -80°C freezer and thaw.
2. Place beads on magnet (200,000 per sample experiment), remove bead storage buffer.
3. Wash 1x 200µL for 1 minute each with 1x PBS, 3mM MgCl₂ buffer. Pool beads to make 200,000 in one tube.
4. Resuspend beads in 70µL of the selection buffer. (100µL of 10x PBS, 1% BSA + 3µL, 100mM MgCl₂ + 57 µL H₂O per sample).
5. Add 30µL radioactively labeled RNA aptamer library to their respective sample.
6. Incubate shaking at 1000 rpm at 37°C for 30 min.
7. Place samples on a magnet.
8. Remove and save supernatant.
9. Wash beads with 200µL wash buffer 1x PBS 3mM MgCl₂ pH 7.4, incubating off the magnet for 3 minute.
10. Place samples on the magnet, remove and save wash solution.
11. Repeat steps 9, 10.
12. Add 100µL water to the sample, pipette mix.
13. Heat at 80°C for 5 minutes.
14. Place samples on a magnet, remove supernatant, and save.
15. Resuspend beads in 100µL water.
16. Measure radioactivity of every fraction using scintillation counter.
17. Analyze amount of background binding present.

[00756] Negative Selection

[00757] As desired, a negative selection step is added prior to incubating the aptamer library with the beads conjugated to the target microvesicles (i.e., procedure "Binding 2F SUL1 RNA pool to microvesicle coated magnetic beads" above). The negative selection can be performed using beads conjugated to the supernatant or the input samples (e.g., plasma) after microvesicles are filtered or sedimented from the sample (referred to as "no microvesicle coated beads," "microvesicle depleted samples," or similar). The steps are:

[00758] 1) Start with aptamer library product from the desired round after transcription as described above. Wash the beads before start: remove storage buffer, wash beads with 200µL wash buffer, then replace buffer as stated below:

[00759] 2) Negative selection step: Add and pipet 20 µL of transcription product (15 mM MgCl₂) with freshly washed 'no microvesicle' coated beads with 10 µL 10x PBS with 1% BSA, 10 µL H₂O. No additional MgCl₂ is needed because the MgCl₂ in the diluted transcription product (TK) provides a final concentration of 3mM MgCl₂.

[00760] 3) Incubate for 30 min at 37 °C, shake at 1000rpm.

[00761] 4) Remove supernatant and add it to the positive selection beads (directly), which are washed microvesicle coated beads.

[00762] Continue with positive selection incubation. See Binding 2F SUL1 RNA pool to microvesicle coated magnetic beads above, starting at step 2. Additional steps through transcription are as detailed above.

Example 12: Additional Aptamers to Breast Cancer (BrCa) derived microvesicles

[00763] In this Example, an aptamer library is screened to identify aptamers that distinguish between microvesicles circulating in the blood of breast cancer patients and microvesicles circulating in the blood of healthy, control individuals (i.e., without breast cancer). The procedure used the same samples and aptamer library as in Example 10 above. The procedure in this Example differs in that negative selection was performed before each positive selection starting after the third round of positive selection.

[00764] Negative selection serves to remove aptamers that bind soluble/abundant/non-informative and common proteins for cancer and non-cancer proteins. Negative selection include performing negative selection on the aptamer candidates selected against BrCa+ microvesicles as follows: (i) using microbeads conjugated to the supernatant from the BrCa+ plasma ultracentrifugation step (which should not contain microvesicles); (ii) using microbeads conjugated to the supernatant from the BrCa- plasma ultracentrifugation step (which should not contain microvesicles); (iii) using microbeads conjugated to BrCa- microvesicles. Negative selection can also be performed on the aptamer candidates selected against BrCa- microvesicles as follows: (ii) using microbeads conjugated to the supernatant from the BrCa+ plasma ultracentrifugation step (which should not contain microvesicles); (ii) using microbeads
conjugated to the supernatant from the BrCa- plasma ultracentrifugation step (which should not contain microvesicles); (iii) using microbeads conjugated to BrCa+ microvesicles. Negative selection rounds are performed between rounds of positive selection as described herein.

[00765] Microvesicles were isolated from plasma of a pool of 60 breast cancer patients (BrCa+). Microvesicles were also isolated from pool of 60 non-cancer samples (BrCa-). Microvesicles were isolated from the plasma using ultracentrifugation (120,000 x g). Microvesicles were in the pellet from the ultracentrifugation. The supernatant from the ultracentrifugation was saved to use as a control. The microvesicles from both sample types were conjugated to MagPlex beads (Luminex Corp, Austin TX).

[00766] The aptamer library used consisted of a 2'F SUL1 RNA aptamer library. The sequence is 5'-GGGAGGACGAUGCGG -N40-CAGACGACUCCGUAGGAUCCGAGA-3' (SEQ ID NO. 8). The aptamer library consists of three sections: Forward primer - 15 nucleotides, variable region - 40 nucleotides; reverse primer - 25 nucleotides. All pyrimidines (C and U) were 2'Fluoro modified.

[00767] The aptamer library was incubated with either the cancer or control micro side -conjugated beads. Nine rounds of positive selection for aptamers that bind the microvesicles were performed in parallel for both types of samples. Negative selection against beads conjugated to the input plasma supernatant after ultracentrifugation before positive selection in rounds 4-9. See **Example 11** above for detailed protocol of the positive selection and negative selection steps.

[00768] The aptamers that were retained from the above positive selection were sequenced using Next Generation sequencing technology consisting of Ion Torrent NGS (Life Technologies, Inc., Carlsbad, CA). The MiSeq system may be used also (Illumina, Inc., San Diego, CA). The sequences are compared to identify aptamers that are found in the cancer samples and not the control samples, and vice versa. Such aptamers provide candidates that can be used to distinguish between BrCa and non-BrCa samples.

[00769] The sequencing data was analyzed according to the following procedure:

**Step 1:** Sequences were ranked according to frequencies in entire aptamer pool recovered in round 9 after negative selection against beads conjugated to micro side -depleted cancer plasma followed by positive selection against beads conjugated to cancer microvesicles.

**Step 2:** Fold changes were calculated between noted sample in Step 1 and: (i) same sample after additional negative selection against microvesicle depleted cancer plasma; (ii) same sample after additional negative selection against non-cancer microvesicles; (iii) same sample after additional negative selection against microvesicles depleted non-cancer plasma.

**Step 3:** Sequences were ranked based on fold changes calculated in Step 2 to identify sequences which are abundant or deficient in aptamer pool selected for breast cancer derived microvesicles.

**Step 4:** Possible mutant sequences (e.g., due to PCR or other errors) were removed based on results of consolidation analysis.

**Step 5:** Sequences were identified with fold changes greater than 3 and minimum frequency 50 in all three variants (i, ii and iii in step 2).

**Step 6:** The same selection schemes as in steps 1-5 were performed for aptamers selected against beads conjugated to non-cancer microvesicles.
A number of representative sequences obtained from these procedures are shown in Table 11. The sequences in the table were identified in the aptamer pools from selection against BrCa microvesicles but were not in the aptamer pools selected against non-cancer samples. In Table 11, the sequences are shown 5' to 3' from left to right, wherein each complete sequence consists of a 5' leader sequence 5'-GAGGAGGACGAUGCGG (SEQ ID NO. 9) followed by the indicated Variable Sequence followed by the 3' tail sequence 5'-CAGACGACUCGUGAGGAUGAGCAGA (SEQ ID NO. 10). Each sequence is derived from a library having a leader and tail (see description above) with a variable sequence between. It is understood that the nucleotide sequences that are disclosed in Table 11 can also be modified to the extent that resulting modifications result in an aptamer having about 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 96, 97, 98, and 99 percent homology to the disclosed sequence and retain the functionality of binding to microvesicle antigens or functional fragments thereof.

<table>
<thead>
<tr>
<th>ID (Figure)</th>
<th>SEQ ID NO.</th>
<th>Variable Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>BCE8 (FIG. 10Ai)</td>
<td>25</td>
<td>UACCGCCCUAAUCCGGACGACGACGUGUACGUUGCCUG</td>
</tr>
<tr>
<td>BCE9 (FIG. 10Ai)</td>
<td>26</td>
<td>GUCUGCCUCUGAUCGACGACGUGUACGUUGCCUG</td>
</tr>
<tr>
<td>BCE10 (FIG. 10Aii)</td>
<td>27</td>
<td>CGCCGUGUCGACGACGUGUACGUUGCCUG</td>
</tr>
<tr>
<td>BCE11 (FIG. 10Aiv)</td>
<td>28</td>
<td>CGCCGUGUCGACGACGUGUACGUUGCCUG</td>
</tr>
<tr>
<td>BCE12 (FIG. 10Av)</td>
<td>29</td>
<td>AUCCAAGUACGUACGACGUGUACGUUGCCUG</td>
</tr>
<tr>
<td>BCE13 (FIG. 10Aii)</td>
<td>30</td>
<td>CUGUGUGAAGUAGAAGCUAAAGCUACGACGUGUACGUUGCCUG</td>
</tr>
<tr>
<td>BCE14 (FIG. 10Avii)</td>
<td>31</td>
<td>AGUAUGCUAAGUAGAAGCUAAAGCUACGACGUGUACGUUGCCUG</td>
</tr>
<tr>
<td>BCE15 (FIG. 10Aivii)</td>
<td>32</td>
<td>AGUAUGCUAAGUAGAAGCUAAAGCUACGACGUGUACGUUGCCUG</td>
</tr>
<tr>
<td>BCE16 (FIG. 10Aix)</td>
<td>33</td>
<td>AGUAUGCUAAGUAGAAGCUAAAGCUACGACGUGUACGUUGCCUG</td>
</tr>
<tr>
<td>BCE17 (FIG. 10Bi)</td>
<td>34</td>
<td>AGUAUGCUAAGUAGAAGCUAAAGCUACGACGUGUACGUUGCCUG</td>
</tr>
<tr>
<td>BCE18 (FIG. 10Bii)</td>
<td>35</td>
<td>AGUAUGCUAAGUAGAAGCUAAAGCUACGACGUGUACGUUGCCUG</td>
</tr>
<tr>
<td>BCE19 (FIG. 10Bii)</td>
<td>36</td>
<td>AGUAUGCUAAGUAGAAGCUAAAGCUACGACGUGUACGUUGCCUG</td>
</tr>
<tr>
<td>BCE20 (FIG. 10Biv)</td>
<td>37</td>
<td>AGUAUGCUAAGUAGAAGCUAAAGCUACGACGUGUACGUUGCCUG</td>
</tr>
<tr>
<td>BCE21 (FIG. 10Bv)</td>
<td>38</td>
<td>AGUAUGCUAAGUAGAAGCUAAAGCUACGACGUGUACGUUGCCUG</td>
</tr>
<tr>
<td>BCE22 (FIG. 10Bvi)</td>
<td>39</td>
<td>AGUAUGCUAAGUAGAAGCUAAAGCUACGACGUGUACGUUGCCUG</td>
</tr>
<tr>
<td>BCE23 (FIG. 10Bvii)</td>
<td>40</td>
<td>AGUAUGCUAAGUAGAAGCUAAAGCUACGACGUGUACGUUGCCUG</td>
</tr>
<tr>
<td>BCE24 (FIG. 10Bviii)</td>
<td>41</td>
<td>AGUAUGCUAAGUAGAAGCUAAAGCUACGACGUGUACGUUGCCUG</td>
</tr>
<tr>
<td>BCE25 (FIG. 10Bxi)</td>
<td>42</td>
<td>AGUAUGCUAAGUAGAAGCUAAAGCUACGACGUGUACGUUGCCUG</td>
</tr>
</tbody>
</table>
[00777] Each sequence in Table 11 is synthesized in two variants for further investigation: 5’ biotinylated and 3’ biotinylated. This provides aptamer variants that can be captured at the 5’ end or the 3’ end as desired. The aptamers are further synthesized with each pyrimidine (C and U) 2’Fluoro modified. The aptamers may also be synthesized as the DNA sequence corresponding to each RNA sequence in Table 11. The aptamer libraries can also be filtered based on predicted secondary sequence, free energy, and other parameters as described herein.

[00778] FIGs. 10A-C show binding of the aptamers in Table 11 against microbeads conjugated to various input samples. The aptamer is indicated above each plot and the plot for each aptamer is indicated in the ID column in Table 11. The input sample is indicated on the X axis from left to right as follows: 1) Cancer Exosome: aptamer binding to microbeads conjugated to microvesicles isolated from plasma samples from breast cancer patients; 2) Cancer Non-exosome: aptamer binding to microbeads conjugated to plasma samples from breast cancer patients after removal of microvesicles by ultracentrifugation; 3) Non-Cancer Exosome: aptamer binding to microbeads conjugated to microvesicles isolated from plasma samples from normal (i.e., non-breast cancer) patients; 4) Non-Cancer Non-Exosome: aptamer binding to microbeads conjugated to plasma samples from breast cancer patients after removal of microvesicles by ultracentrifugation. As shown in FIGs. 10A-C, the aptamers were each able to distinguish between the cancer microvesicle samples versus the supernatant control samples and the non-cancer microvesicles. Further, all sequences in Table 11 were observed as binding more abundantly to cancer derived microvesicles as compared to non-cancer derived microvesicles with the exception of BCE10 and BCE14, which were observed as binding more abundantly to non-cancer derived microvesicles as compared to cancer derived microvesicles.

[00779] Based on the comparisons performed in this Example, aptamers that bind different starting input are obtained, including: 1) aptamers that preferentially bind cancer-derived microvesicles over non-cancer derived microvesicles; 2) aptamers that preferentially bind non-cancer-derived microvesicles over cancer derived microvesicles; 3) aptamers that bind both non-cancer-derived microvesicles and cancer derived microvesicles (e.g., "universal" binders); and 4) aptamers that bind plasma components that have been depleted of microvesicles.

[00780] The aptamer libraries in this Example are further subjected to four rounds of additional negative and positive selection. The positive selection is performed as described in this Example. The negative selection rounds are performed using the beads conjugated to non-cancer microvesicles as negative selection for aptamers obtained by positive selection against beads conjugated to cancer microvesicles. Similarly, the negative selection rounds are performed using the beads conjugated to cancer microvesicles as negative selection for aptamers obtained by positive selection against beads conjugated to non-cancer microvesicles.

| BCE26 (FIG. 10Ci) | 43 | GGUGAUCCUCGCCCGCCUCGAAGAUUGUG2CAUAUC |
| BCE27 (FIG. 10Cii) | 44 | GCUCACCAUCGUCUAGUGAGUCACCUCUAGUUCAUCAGGC |
Example 13: Disease Diagnosis

[00781] This example illustrates the use of oligonucleotide probes of the present invention to diagnose a proliferative disease.

[00782] A suitable quantity of an oligonucleotide or pool of oligonucleotides that bind a BrCa-derived population of microvesicles, such as identified in Example 10 or Example 12 or various Examples below, is synthesized via chemical means known in the art. The oligonucleotides are conjugated to a diagnostic agent suitable for detection, such as a fluorescent moiety, using a conjugation method known in the art.

[00783] The composition is applied to microvesicles isolated from blood samples taken from a test cohort of patients suffering from a proliferative disease associated with the overexpression of microvesicles, e.g. breast cancer. The composition is likewise applied to microvesicles isolated from blood samples taken from a negative control cohort, not suffering from a proliferative disease.

[00784] The use of appropriate detection techniques (e.g., microbead assay or flow cytometry) on the test cohort samples indicates the presence of disease, while the same techniques applied to the control cohort samples indicate the absence of disease.

[00785] The results show that the oligonucleotides of the present invention are useful in diagnosing proliferative diseases.

Example 14: Theranostics

[00786] This example illustrates the use of oligonucleotide probes of the present invention to provide a theranosis for a drug for treating a proliferative disease.

[00787] A suitable quantity of an oligonucleotide or pool of oligonucleotides that bind a BrCa-derived population of microvesicles, such as identified in Example 10 or Example 12 or various Examples below, is synthesized via chemical means known in the art. The probes are conjugated to an agent suitable for detection, such as a fluorescent moiety, using a method known in the art such as conjugation. The oligonucleotide probe or panel of oligonucleotide probes are within a suitable composition, such as a buffered solution.

[00788] Treatment selection. The composition is applied to microvesicles isolated from blood samples taken from a test cohort of patients suffering from a proliferative disease, e.g. breast cancer, that responded to a certain treatment, e.g., traotuzamab. The composition is likewise applied to microvesicles isolated from blood samples taken from a control cohort consisting of patients suffering from the same proliferative disease that did not respond to the treatment. The use of appropriate detection techniques (e.g., microbead assay or flow cytometry) on the test cohort samples indicates that probes which bind the samples are useful for identifying patients that will respond to the treatment, while the same techniques applied to the control cohort samples identifies probes useful for identifying patients that will not respond to the treatment.

[00789] Treatment monitoring. In another setting, the composition is applied to microvesicles isolated from blood samples taken from a test cohort of patients suffering from a proliferative disease, e.g. breast cancer, prior to or during a course of treatment, such as surgery, radiotherapy and/or chemotherapy. The
composition is then applied to microvesicles isolated from blood samples taken from the patients over a
time course. The use of appropriate detection techniques (e.g., microbead assay or flow cytometry) on the
test cohort samples indicates whether the detected population of disease-related microvesicles increases,
decreases, or remains steady in concentration over time during the course of treatment. An increase in the
population of disease-related microvesicles post-treatment may indicate that the treatment is ineffective
whereas a decrease in the population of disease-related microvesicles post-treatment may indicate that the
treatment has a beneficial effect.

[00790] The results show that the oligonucleotide probes of the present invention are useful in theranosing
proliferative diseases.

Example 15: Therapeutic Oligonucleotide Probes

[00791] This example illustrates the use of oligonucleotide probes of the present invention to treat a
proliferative disease.

[00792] A suitable quantity of an oligonucleotide or pool of oligonucleotides that bind a BrCa-derived
population of microvesicles, such as identified in Example 10 or Example 12 or various Examples
below, is synthesized via chemical means known in the art. The oligonucleotides are conjugated to a
chemotherapeutic agent, such as Doxil, using a conjugation method known in the art. The conjugate is
formulated in an aqueous composition.

[00793] The composition is administered intravenously, in one or more doses, to a test cohort of mice
suffering from a proliferative disease associated with the overexpression of the microvesicles, e.g. a breast
cancer model. A control cohort, not suffering from a proliferative disease is administered the identical
composition intravenously, according to a corresponding dosage regimen.

[00794] Pathological analysis of tumor samples and/or mouse survival indicates that mortality and/or
morbidity are improved in the test cohort over the control cohort.

[00795] The results show that the oligonucleotides of the present invention are useful in treating
proliferative diseases.

[00796] Useful oligonucleotides can be used to treat proliferative diseases in other organisms, e.g., a
human.

Example 16: Oligonucleotide - Sequencing Detection Method

[00797] This example illustrates the use of an oligonucleotide pool to detect microvesicles that are
indicative of a phenotype of interest. The method makes use of a pool of oligonucleotides that have been
enriched against a target of interest that is indicative of a phenotype of interest. The method in this
Example allows efficient use of a library of oligonucleotides to preferentially recognize a target entity.

[00798] For purposes of illustration, the method is described in the Example with a microvesicle target
from a bodily fluid sample. One of skill will appreciate that the method can be extended to other types of
target entity (e.g., cells, proteins, various other biological complexes), sample (e.g., tissue, cell culture,
biopsy, other bodily fluids) and other phenotypes (other cancers, other diseases, etc) by enriching an
aptamer library against the desired input samples.

[00799] General workflow:
[00800] 1) Obtain sample (plasma, serum, urine or any other biological sample) of patients with unknown medical etymology and pre-treating them accordingly to ensure availability of the target of interest (see below). Where the target of interest is a microvesicle population, the microvesicles can be isolated and optionally tethered to a solid support such as a microbead.

[00801] 2) Expose pre-treated sample to an oligonucleotide pool carrying certain specificity against target of interest. As described herein, an oligonucleotide pool carrying certain specificity against the target of interest can be enriched using various selection schemes, e.g., using non-cancer microvesicles for negative selection and cancer microvesicles for positive selection as described above. DNA or RNA oligonucleotides can be used as desired.

[00802] 3) Contact oligonucleotide library with the sample.

[00803] 4) Elute any oligonucleotides bound to the target.

[00804] 5) Sequence the eluted oligonucleotides. Next generation sequencing methods can be used.

[00805] 6) Analyze oligonucleotide profile from the sequencing. A profile of oligonucleotides known to bind the target of interest indicates the presence of the target within the input sample. The profile can be used to characterize the sample, e.g., as cancer or non-cancer.

[00806] Protocol variations:

[00807] Various configurations of the assay can be performed. Four exemplary protocols are presented for the purposes of the oligonucleotide-sequencing assay. Samples can be any appropriate biological sample. The protocols can be modified as desired. For example, the microvesicles can be isolated using alternate techniques instead or in addition to ultracentrifugation. Such techniques can be disclosed herein, e.g., polymer precipitation (e.g., PEG), column chromatography, and/or affinity isolation.

[00808] Protocol 1:

[00809] Ultracentrifugation of 1-5 ml bodily fluid samples (e.g., plasma/serum/urine) (120K x g, no sucrose) with two washes of the precipitate to isolate microvesicles.

[00810] Measure total protein concentration of recovered sample containing the isolated microvesicles.

[00811] Conjugate the isolated microvesicles to magnetic beads (for example MagPlex beads (Luminex Corp. Austin TX)).

[00812] Incubate conjugated microvesicles with oligonucleotide pool of interest.

[00813] Wash unbound oligonucleotides by retaining beads using magnet.

[00814] Elute oligonucleotides bound to the microvesicles.

[00815] Amplify and purify the eluted oligonucleotides.

[00816] Oligonucleotide sequencing (for example, Next generation methods; Ion Torrent: fusion PCR, emulsion PCR, sequencing).


[00818] Protocol 2:

[00819] This alternate protocol does not include a microvesicle isolation step, microvesicles conjugation to the beads, or separate partitioning step. This may present non-specific binding of the oligonucleotides against the input sample.
WO 2017/004243

Remove cells/debris from bodily fluid sample and dilute sample with PBS containing MgC^ (2mM).

Pre-mix sample prepared above with oligonucleotide library.

Ultracentrifugation of oligonucleotide/sample mixture (120K x g, no sucrose). Wash precipitated microvesicles.

Recover precipitate and elute oligonucleotides bound to microvesicles.

Amplify and purify the eluted oligonucleotides.

Oligonucleotide sequencing (for example, Next generation methods; Ion Torrent: fusion PCR, emulsion PCR, sequencing).

Assess oligonucleotide profile.

Protocol 3:

This protocol uses filtration instead of ultracentrifugation and should require less time and sample volume.

Remove cells/debris from bodily fluid sample and dilute it with PBS containing MgC^ (2mM).

Pre-mix sample prepared above with oligonucleotide library.

Load sample into filter (i.e., 150K or 300K MWCO filter or any other that can eliminate unbound or unwanted oligonucleotides). Centrifuge sample to concentrate. Concentrated sample should contain microvesicles.

Wash concentrate. Variant 1: Dilute concentrate with buffer specified above to the original volume and repeat centrifugation. Variant 2: Dilute concentrate with buffer specified above to the original volume and transfer concentrate to new filter unit and centrifuge. Repeat twice.

 Recover concentrate and elute oligonucleotides bound to microvesicles.

Amplify and purify the eluted oligonucleotides.

Oligonucleotide sequencing (for example, Next generation methods; Ion Torrent: fusion PCR, emulsion PCR, sequencing).

Assess oligonucleotide profile.

Protocol 4:

Ultracentrifugation of 1-5 ml bodily fluid sample (120K x g, no sucrose) with 2 washes of the precipitate to isolate microvesicles.

Pre-mix microvesicles with oligonucleotide pool.

Load sample into 300K MWCO filter unite and centrifuge (2000xg). Concentration rate is ~3x.

Wash concentrate. Variant 1: Dilute concentrate with buffer specified above to the original volume and centrifuge. Repeat twice. Variant 2: Dilute concentrate with buffer specified above to the original volume and transfer concentrate to new filter unit and centrifuge. Repeat twice

 Recover concentrate and elute oligonucleotides bound to microvesicles.

Amplify and purify the eluted oligonucleotides.

Oligonucleotide sequencing (for example, Next generation methods; Ion Torrent: fusion PCR, emulsion PCR, sequencing).
Assess oligonucleotide profile.

In alterations of the above protocols, polymer precipitation is used to isolate microvesicles from the patient samples. For example, the oligonucleotides are added to the sample and then PEG4000 or PEG8000 at 4% or 8% concentration is used to precipitate and thereby isolate microvesicles. Elution, recovery and sequence analysis continues as above.

Example 17: Plasma/Serum probing with an Oligonucleotide Probe Library

The following protocol is used to probe a plasma or serum sample using an oligonucleotide probe library.

Input oligonucleotide library:

Use 2 ng input of oligonucleotide library per sample.

Input oligonucleotide library is a mixture of two libraries, cancer and non-cancer enriched, concentration is 16.3 ng/ul.

Dilute to 0.2 ng/ul working stock using Aptamer Buffer (3 mM MgCl₂ in IX PBS)

Add 10 ul from working stock (equal to 2 ng library) to each optisalve tube

Materials:

PBS, Hyclone SH30256.0 1. LN: AYG165629, bottle# 8237, exp. 7/2015
Round Bottom Centrifuge Tubes, Beckman 326820, LN:P9 1207
OptiSeal Centrifuge tubes and plugs, polyallomer Konical, Beckman 361621, lot# Z10804SCA
Ultracentrifuge rotor: 50.4 TI
Ultracentrifuge rotor: 50.4 TI, Beckman Caris ID# 0478
Protocol:

1 Pre-chill tabletop centrifuge, ultracentrifuge, buckets, and rotor at 4°C.
2 Thaw plasma or serum samples
3 Dilute 1 ml of samples with 1:2 with Aptamer Buffer (3 mM MgCl₂ in IX PBS)
4 Spin at 2000 x g, 30 min, 4°C to remove debris (tabletop centrifuge)
5 Transfer supernatants for all samples to a round bottom conical
6 Spin at 12,000 x g, 45 min, 4°C in ultracentrifuge to remove additional debris.
7 Transfer supernatant about 1.8 ml for all samples into new OptiSeal bell top tubes (uniquely marked).
8 Add 2 ng (in 10 ul) of DNA Probing library to each optisalve tube
9 QS to 4.5 ml with Aptamer Buffer
10 Fix caps onto the OptiSeal bell top tubes
11 Apply Parafilm around caps to prevent leakage
12 Incubate plasma and oligonucleotide probe library for 1 hour at room temperature with rotation
13 Remove parafilm (but not caps)
14 Place correct spacer on top of each plugged tube
15 Mark pellet area on the tubes, insure this marking is facing outwards from center.
16 Spin tubes at 120,000 x g, 2 hr, 4°C (inner row, 33,400 rpm) to pellet microvesicles.
17 Check marking is still pointed away from center.

18 Completely remove supernatant from pellet, by collecting liquid from opposite side of pellet marker and using a 10 ml syringe barrel and 21G2 needle

19 Discard supernatant in appropriate biohazard waste container

20 Add 1 mL of 3 mM MgC\textsubscript{12} diluted with 1X PBS

21 Gentle vortex, 1600rpm for 5 sec and incubate 5 min at RT.

22 QS to -4.5 mL with 3 mM MgC\textsubscript{12} diluted with 1X PBS

23 Fix caps onto the OptiSeal bell top tubes.

24 Place correct spacer on top of each plugged tube.

25 Mark pellet area on the tubes, insure this marking is facing outwards from center.

26 Spin tubes at 120,000 x g, 70 min, 4°C (inner row 33,400 rpm) to pellet microvesicles

27 Check marking is still pointed away from center.

28 Completely remove supernatant from pellet, by collecting liquid from opposite side of pellet marker and using a 10 ml syringe barrel and 21G2 needle

29 Discard supernatant in appropriate biohazard waste container

30 Add 1 mL of 3 mM MgC\textsubscript{12} diluted with 1X PBS

31 Gentle vortex, 1600rpm for 5 sec and incubate 5 min at RT.

32 QS to -4.5 mL with 3 mM MgC\textsubscript{12} diluted with 1X PBS

33 Fix caps onto the OptiSeal bell top tubes.

34 Place correct spacer on top of each plugged tube.

35 Mark pellet area on the tubes, insure this marking is facing outwards from center.

36 Spin tubes at 120,000 x g, 70 min, 4°C (inner row 33,400 rpm) to pellet microvesicles

37 Check marking is still pointed away from center.

38 Save an aliquot of the supernatant (100ul into a 1.5ml tube)

39 Completely remove supernatant from pellet, by collecting liquid from opposite side of pellet marker and using a 10 ml syringe barrel and 21G2 needle

40 Add 50 ul of Rnase-free water to the side of the pellet

41 Leave for 15 min incubation on bench top

42 Cut top off tubes using clean scissors.

43 Resuspend pellet, pipette up and down on the pellet side

44 Measure the volume, make a note on the volume in order to normalize all samples

45 Transfer the measured resuspended eluted microvesicles with bound oligonucleotides to a Rnase free 1.5ml Eppendorf tube

46 Normalize all samples to 100ul to keep it even across samples and between experiments.

Next Generation Sequencing Sample Preparation:

1) Use 50 ul of sample from above, resuspended in 100 ul H2O and containing microvesicle/oligo complexes, as template in Transposon PCR, 14 cycles.

2) AMPure transposon PCR product, use entire recovery for indexing PCR, 10 cycles.
[00909] III) Check indexing PCR product on gel, proceed with AMPure if band is visible. Add 3 cycles if band is invisible, check on gel. After purification quantify product with QuBit and proceed with denaturing and diling for loading on HiSeq flow cell (Illumina Inc., San Diego, CA).

[00910] IV) 5 samples will be multiplexed per one flow cell. 10 samples per HiSeq.

Example 18: Enrichment of Oligonucleotides to Breast Cancer (BrCa) and non-BrCa derived microvesicles

[00911] In this Example, a naive oligonucleotide probe library is screened to enrich oligonucleotides that identify microvesicles circulating in the blood of breast cancer patients and microvesicles circulating in the blood of healthy, control individuals (i.e., without breast cancer). The procedure in this Example used the 9-round selection scheme shown in FIG. 11A.

[00912] In FIGS. 11A-11B, the following abreviations are used: CE- Cancer exosome samples; NCE-Non-cancer exosome samples; CNE- Cancer no exosome samples; NCNE- Non-cancer no exosome samples; MFI- Median Fluorescence Intensity; GC- GC content group analysis; Heat- heat elution; NaOH- sodium hydroxide elution. The terms exosome and microvesicle are used interchangeably.

[00913] As indicated in FIG. 11A, 0.1 mg/mL salmon sperm DNA was added as a competitor in selection rounds 7-9. FIG. 11B shows the enrichment of aptamers that bind to microvesicles (CE and NCE). Binding was determined by fluorescently labeling aptamer candidates and detecting binding to bead-capture microvesicles using a microbead assay format. At the end of round 9, only aptamers having GC content between -30% and 70% were selected to generate panels to differentiate cancer and non-cancer samples. In addition, several groups of aptamers were selected using the following criteria: 1) 12 most commonly observed; 2) 12 most commonly observed having > 2-fold change in binding observed between cancer and normal samples; 3) 30 most commonly observed having > 2-fold change in binding observed between cancer and normal samples, >10 read counts in the library determined by high throughput sequencing, heat elution after selection round 9; and 4) 30 most commonly observed having > 2-fold change in binding observed between cancer and normal samples, >10 read counts in the library determined by high throughput sequencing, NaOH elution after selection round 9.

[00914] Retained sequences are shown in Table 12 with selection criteria 1-4 indicated. In Table 12, the sequences are shown 5’ to 3’ from left to right, wherein each complete sequence consists of a 5’ leader sequence 5’-GCTGTCTACTCCGAA (SEQ ID NO. 45) followed by the indicated Variable Sequence followed by the 3’ tail sequence 5’-GCAGCTGTATCTTTGCTCC (SEQ ID NO. 46).

<table>
<thead>
<tr>
<th>SEQ ID NO.</th>
<th>Selection Criteria</th>
<th>Variable Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>47</td>
<td>3</td>
<td>GTCCGCTTGGGGTGTTACTCGAAAAATTCGCTGTCTGGGCGG</td>
</tr>
<tr>
<td>48</td>
<td>3</td>
<td>CCTGCGCTTCTTCAATGTCTCGTTCTGGGAAGGTCTTTAG</td>
</tr>
<tr>
<td>49</td>
<td>3</td>
<td>TAGGAGGATATATGTATCCGGCTGGGATCGGGCTACTTTCCCC</td>
</tr>
<tr>
<td>50</td>
<td>3</td>
<td>TAGGAGGATATATGTATCCGGCTGGGATCGGGCTACTTTCCCC</td>
</tr>
<tr>
<td>51</td>
<td>3</td>
<td>TAGGAGGATATATGTATCCGGCTGGGATCGGGCTACTTTCCCC</td>
</tr>
<tr>
<td>52</td>
<td>3</td>
<td>TGGCCTTCTCTTGGGTTAGGATGGGGTGAGGAGG</td>
</tr>
<tr>
<td>53</td>
<td>3</td>
<td>GGGGCGGCTGGGGCTGGTAAAGGTATCGGGGCTATTGTT</td>
</tr>
<tr>
<td>54</td>
<td>3</td>
<td>TAGGAGGATATATGTATCCGGCTGGGATCGGGCTACTTTCCCC</td>
</tr>
<tr>
<td>55</td>
<td>3</td>
<td>CCAATCTCGGAAGGTTTAAATAAGGTGGTCTTTAGCTTTAAGT</td>
</tr>
<tr>
<td>56</td>
<td>3</td>
<td>TACGGTTAAGTGGGTCGAGGGGAAAGGTGTTTGTTG</td>
</tr>
<tr>
<td>57</td>
<td>3</td>
<td>GTGAGGTCGTTGGCGGGGGGATGTTGAGGTTTATCATT</td>
</tr>
<tr>
<td>58</td>
<td>3</td>
<td>CGATATTGGGTTGGGTTGGGTGTATATTCCGGGTG</td>
</tr>
<tr>
<td>59</td>
<td>3</td>
<td>TCGGTTCCTGTATTGCCGCTTGTTAACATAGATTTAG</td>
</tr>
<tr>
<td>60</td>
<td>3</td>
<td>GCAGGGAGGTGGGGGTGTAGGTTGAGGTTTACCTTAC</td>
</tr>
<tr>
<td>61</td>
<td>3</td>
<td>CGGAGAATGGGGGATTTGAGGGGATGGGTTCGACTTC</td>
</tr>
<tr>
<td>62</td>
<td>3</td>
<td>TAGAGAATGATGATGATGATCAGGGGGGCTGGTTC</td>
</tr>
<tr>
<td>63</td>
<td>3</td>
<td>TATAGTTGGGGCTAGGGGGGGGATGTTGAGGTTAATC</td>
</tr>
<tr>
<td>64</td>
<td>3</td>
<td>TAGAGAATGATGATGATGATCAGGGGGGCTGGTTC</td>
</tr>
<tr>
<td>65</td>
<td>3</td>
<td>TCGATCTGGTCAAGATCTCTGCTTCAAATCTATAGAT</td>
</tr>
<tr>
<td>66</td>
<td>3</td>
<td>TGGCTGTCTGGCGGGGCGTCTTGCTTGCTTGTTATAGC</td>
</tr>
<tr>
<td>67</td>
<td>3</td>
<td>GCGAGAATGGGGGATTTGAGGGGATGGGTTCGACTTC</td>
</tr>
<tr>
<td>68</td>
<td>3</td>
<td>TAGAGAATGATGATGATCAGGGGGGCTGGTTC</td>
</tr>
<tr>
<td>69</td>
<td>3</td>
<td>CGGCTGTCTGGCGGGGCGTCTTGCTTGCTTGTTATAGC</td>
</tr>
<tr>
<td>70</td>
<td>3</td>
<td>CGGAGAATGGGGGATTTGAGGGGATGGGTTCGACTTC</td>
</tr>
<tr>
<td>71</td>
<td>3</td>
<td>TAGAGAATGATGATGATCAGGGGGGCTGGTTC</td>
</tr>
<tr>
<td>72</td>
<td>3</td>
<td>TAGAGAATGATGATGATGATCAGGGGGGCTGGTTC</td>
</tr>
<tr>
<td>73</td>
<td>3</td>
<td>CTTCGCTGTTCTTTCACTTTGCTTGCTTGTTATAGC</td>
</tr>
<tr>
<td>74</td>
<td>3</td>
<td>GCGAGAATGGGGGATTTGAGGGGATGGGTTCGACTTC</td>
</tr>
<tr>
<td>75</td>
<td>3</td>
<td>TAGAGAATGATGATGATCAGGGGGGCTGGTTC</td>
</tr>
<tr>
<td>76</td>
<td>3</td>
<td>TAGAGAATGATGATGATGATCAGGGGGGCTGGTTC</td>
</tr>
<tr>
<td>77</td>
<td>4</td>
<td>TGGATGTATGGGTCTCGGGGTGGGATTTTATGTTAG</td>
</tr>
<tr>
<td>78</td>
<td>4</td>
<td>GTCCCCGTGTTGGGGGTCTGATCAGAAATCTCGGCTCTTGCGGGG</td>
</tr>
<tr>
<td>79</td>
<td>4</td>
<td>TCTTGTACAAATAGGAGGGAAGGGGGTTTTGGGAGGTGGGTG</td>
</tr>
<tr>
<td>80</td>
<td>4</td>
<td>CTTCGCTGTTCTTTCACTTTGCTTGCTTGTTATAGC</td>
</tr>
<tr>
<td>81</td>
<td>4</td>
<td>CTAGATTTAGATAGGCTGGGGTCTCGGCTTGGGGGGTGGGGG</td>
</tr>
<tr>
<td>82</td>
<td>4</td>
<td>TGCCCTGTGAGCTGTGGGCTGCGGCTGGGCTGGGCTGGG</td>
</tr>
<tr>
<td>83</td>
<td>4</td>
<td>TAGAGAATGATGATGATGATCAGGGGGGCTGGTTC</td>
</tr>
<tr>
<td>84</td>
<td>4</td>
<td>CGGAGAATGGGGGATTTGAGGGGATGGGTTCGACTTC</td>
</tr>
<tr>
<td>85</td>
<td>4</td>
<td>TGGATGTATGGGTCTCGGGGTGGGATTTTATGTTAG</td>
</tr>
<tr>
<td>86</td>
<td>4</td>
<td>TAGAGAATGATGATGATGATCAGGGGGGCTGGTTC</td>
</tr>
<tr>
<td>87</td>
<td>4</td>
<td>TCTTGTACAAATAGGAGGGAAGGGGGTTTTGGGAGGTGGGTG</td>
</tr>
<tr>
<td>88</td>
<td>4</td>
<td>CTTCGCTGTTCTTTCACTTTGCTTGCTTGTTATAGC</td>
</tr>
<tr>
<td>89</td>
<td>4</td>
<td>GTGGGGGCGCTTCCGTGCTTGGCGGTGGGCTGGGCTGGG</td>
</tr>
<tr>
<td>90</td>
<td>4</td>
<td>TGGATGTATGGGTCTCGGGGTGGGATTTTATGTTAG</td>
</tr>
<tr>
<td>91</td>
<td>4</td>
<td>TGGATGTATGGGTCTCGGGGTGGGATTTTATGTTAG</td>
</tr>
<tr>
<td>92</td>
<td>4</td>
<td>TGGGCTGTGAGCTGTGGGCTGCGGCTGGGCTGGGCTGGG</td>
</tr>
<tr>
<td>93</td>
<td>4</td>
<td>CTGAGCTGGTTCTTTCACTTTGCTTGCTTGTTATAGC</td>
</tr>
<tr>
<td>94</td>
<td>4</td>
<td>TGGCTGTGAGCTGTGGGCTGCGGCTGGGCTGGGCTGGG</td>
</tr>
<tr>
<td>95</td>
<td>4</td>
<td>GGAGGAGGGGGGTGGGCTGCTGCTGCTGCTGCTGCTG</td>
</tr>
<tr>
<td>96</td>
<td>4</td>
<td>TGGATGTATGGGTCTCGGGGTGGGATTTTATGTTAG</td>
</tr>
<tr>
<td>97</td>
<td>4</td>
<td>TGGATGTATGGGTCTCGGGGTGGGATTTTATGTTAG</td>
</tr>
<tr>
<td>98</td>
<td>4</td>
<td>TGGATGTATGGGTCTCGGGGTGGGATTTTATGTTAG</td>
</tr>
<tr>
<td>99</td>
<td>4</td>
<td>TGGATGTATGGGTCTCGGGGTGGGATTTTATGTTAG</td>
</tr>
<tr>
<td>100</td>
<td>4</td>
<td>TGGATGTATGGGTCTCGGGGTGGGATTTTATGTTAG</td>
</tr>
<tr>
<td>101</td>
<td>4</td>
<td>TGGATGTATGGGTCTCGGGGTGGGATTTTATGTTAG</td>
</tr>
<tr>
<td>102</td>
<td>4</td>
<td>TGGATGTATGGGTCTCGGGGTGGGATTTTATGTTAG</td>
</tr>
<tr>
<td>103</td>
<td>4</td>
<td>TGGATGTATGGGTCTCGGGGTGGGATTTTATGTTAG</td>
</tr>
<tr>
<td>104</td>
<td>4</td>
<td>TGGATGTATGGGTCTCGGGGTGGGATTTTATGTTAG</td>
</tr>
<tr>
<td>105</td>
<td>4</td>
<td>TGGATGTATGGGTCTCGGGGTGGGATTTTATGTTAG</td>
</tr>
<tr>
<td>106</td>
<td>4</td>
<td>TGGATGTATGGGTCTCGGGGTGGGATTTTATGTTAG</td>
</tr>
<tr>
<td>107</td>
<td>1</td>
<td>TGGATGTATGGGTCTCGGGGTGGGATTTTATGTTAG</td>
</tr>
<tr>
<td>108</td>
<td>1</td>
<td>TGGATGTATGGGTCTCGGGGTGGGATTTTATGTTAG</td>
</tr>
<tr>
<td>109</td>
<td>1</td>
<td>TGGATGTATGGGTCTCGGGGTGGGATTTTATGTTAG</td>
</tr>
</tbody>
</table>
Individual aptamer candidates in Table 12 are synthesized. The aptamers are labeled with a fluorescent label and used to detect binding to bead-capture microvesicles using a microbead assay format. The microbead results are used to identify aptamers and panels of aptamers that differentiate cancer and non-cancer plasma samples.

Example 19: Aptamer based precipitation of microvesicles in plasma

Aptamers can specifically recognize target proteins with nanomolar affinity and have some potential advantages compared to antibodies due to their chemical stability, ease of synthesis and overall reproducibility. Microvesicles were isolated from a rat glioma cell line model that expresses or does not express the human EGFR protein in order to optimize aptamer-microvesicle complex formation and isolation in plasma. Microvesicles spiked into human plasma were precipitated with the positive EGFR DNA or RNA aptamer but not with a reverse complement control sequence. There was no observed binding of the positive or negative aptamer to microvesicles that do not express EGFR. Binding of aptamers to microvesicles was confirmed with aptamer based ELISAs, EMSA assays, and flow cytometry. These results demonstrate aptamer-based precipitation of microvesicles from a complex biological sample.

Example 20: Quantitative proteomics of EGFR, EGFRviii, and EGFR negative microvesicles from a rat glioma cell line

In this Example, we performed quantitative proteomics on microvesicles isolated from a rat glioma cell line F98 (parental line), F98(EGFR) and F98(EGFRviii). Different lots of microvesicles purified from each cell line type were considered biological replicates and each biological replicate was analyzed in four technical replicates by tandem mass tag 6-plex (TMT) peptide labeling and LC-MS/MS. The general microvesicle markers CD8, CD9, CD63, TsglOl, and Alix were identified in all technical
replicates. Multiple proteins were found to be up or down-regulated including a 7.5-fold increase in EGFR and a 2.6-fold increase in the EGFR binding partner HMCN2 in the F98[EGFR] compared with the parental line. These results indicate that expression of a single biomarker may affect global proteomic changes in microvesicles.

**Example 21: Oligonucleotide probe library**

This Example presents further development of the oligonucleotide probe library to detect biological entities such as described in Example 18 above. In this Example, steps were taken to reduce the presence of double stranded oligonucleotides (dsDNA) when probing the patient samples. The data were also generated comparing the effects of 8% and 6% PEG used to precipitate microvesicles (and potentially other biological entities) from the patient samples.

**Protocol:**

1) Pre-chill tabletop centrifuge at 4°C.

2) Protease inhibition: dissolve 2 tablets of "cOmplete ULTRA MINI EDTA-free EASYpack" protease inhibitor in 1100 ul of H₂O (2× stock of protease inhibitor).

3) Add 50 ul of protease inhibitor to the sample (on top of frozen plasma) and start thawing: 1 ml total ea.

4) To remove cells/debris, spin samples at 10,000 x g, 20 min, 4°C. Collect 1 ml supernatant (SN).

5) Mix 1 ml supernatant from step 4 with 1 ml of 2xPBS 6 mM MgCl₂, collect 400 ul into 3 tubes (replicates A, B, C) and use it in step 6.

6) Add competitor per Table 13: make dilutions in lxPBS, 3 mM MgCl₂, mix well, pour into trough, pipet using multichannel.

**Table 13: Competitors**

<table>
<thead>
<tr>
<th>units</th>
<th>Type of Competitor</th>
<th>Stock Concentration</th>
<th>Intermediate stock concentration</th>
<th>Number of samples</th>
<th>Volume from stock to make intermediate stock, ul</th>
<th>Buffer to make intermediate stock</th>
<th>Final Volume, ul</th>
<th>Final Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>ng/ul</td>
<td>Salmon DNA</td>
<td>-</td>
<td>40</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>425.5</td>
<td>0.8</td>
</tr>
<tr>
<td>ng/ul</td>
<td>RNA</td>
<td>-</td>
<td>40</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>425.5</td>
<td>0.8</td>
</tr>
<tr>
<td>x</td>
<td>S1</td>
<td>20</td>
<td>0.5</td>
<td>280</td>
<td>65.5</td>
<td>2555.6</td>
<td>425.5</td>
<td>0.01</td>
</tr>
</tbody>
</table>

7) Incubate for 10 min, RT, end-over-end rotation

Pool of 6-3S and 8-3S oligonucleotide probing libraries is ready: 2.76 ng/ul (~185 ng). Save pool stock and dilutions. New pool can be made by mixing 171.2ul (500ng) of library 6-3S (2.92 ng/ul) with 190.8ul (500ng) of library 8-3S (2.62 ng/ul). Aliquot pooled library into 30 ul and store at -80°C.
Add ssDNA oligonucleotide probing library to the final concentration 2.5 pg/ul for binding. Make dilutions in 1xPBS, 3mM MgCl2.

### Table 14: Probe library calculations

<table>
<thead>
<tr>
<th>Original stock, ng/ul</th>
<th>Lib Name</th>
<th>Required working stock (ng/ul)</th>
<th>ul from original stock to make working</th>
<th>ul of buffer to make working</th>
<th>Final volume, ul</th>
<th>Number of samples</th>
<th>Volume per sample from working stock</th>
<th>Final concentration (pg/ul)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.76</td>
<td>Pooled library 6-3S/8-3S</td>
<td>0.1</td>
<td>26.1</td>
<td>694.1</td>
<td>720.2</td>
<td>60</td>
<td>10.9</td>
<td>2.5</td>
</tr>
</tbody>
</table>

8) Binding: Incubate for 1h at RT with rotation.

9) Prepare polymer solution: 20% PEG8000 in 1x PBS 3mM MgCl2 (dilute 40% PEG8000 with 2xPBS with 6mM MgCl2). Add 20% PEG8000 to sample to the final concentration 6%. Invert few times to mix, incubate for 15 min at 4C

### Table 15: PEG calculations

<table>
<thead>
<tr>
<th>PEG MW</th>
<th>PEG stock, %</th>
<th>Final conc., %</th>
<th>Final volume, ul</th>
<th>Volume 20% PEG to add, ul</th>
<th>Volume of buffer to adjust final volume, ul</th>
<th>Sample volume before adding PEG</th>
<th>Total samples</th>
<th>Total 20% PEG needed, ml</th>
</tr>
</thead>
<tbody>
<tr>
<td>8000</td>
<td>20</td>
<td>6</td>
<td>622.8</td>
<td>186.9</td>
<td>-0.4</td>
<td>436.4</td>
<td>60</td>
<td>11.2</td>
</tr>
</tbody>
</table>

10) Spin at 10,000 x g for 5 min, RT.

11) Remove SN, add 1ml 1xPBS, 3 mM MgCl2 and wash pellet by gentle inversion with 1ml aptamer buffer.

12) Remove buffer, Re-suspend pellets in 100 ul H2O: incubate at RT for 10 min on mixmate 900rpm to re-suspend.

13) Make sure each sample is re-suspended by pipeting after step 13. Make notes on hardly re-suspendable samples.

14) 50 ul of re-suspended sample to indexing PCR -> next generation sequencing (NGS).

15) Keep leftover at 4C

Technical Validation:

The current protocol was tested versus a protocol using 8% PEG8000 to precipitate microvesicles. The current protocol further comprises steps to reduce dsDNA in the oligonucleotide probing libraries.

FIG. 12A shows the within sample variance (black) between binding replicates and the between sample variance (grey). Black is on top of grey, thus any observable grey oligo is informative about
differences in the biology of two patient samples. This evaluation of Sources of Variance shows that the technical variances is significantly smaller than the biological variance.

[00940] FIG. 12B shows the impact of using a higher proportion of single stranded DNA and PEG 6% isolation (white bars) compared to when there is a higher amount of double stranded DNA and 8% PEG (grey). This data indicates that the protocol in this Example improves biological separation between patients.

[00941] The plots in FIG. 12C show the difference between an earlier protocol (PEG 8% with increased dsDNA) and a modified protocol of the Example (PEG 6% no dsDNA). The black is the scatter between replicates (independent binding events) and the grey is the difference between patients. This data shows that the signal to noise increased significantly using the newer protocol.

[00942] Patient testing:

[00943] The protocol above was used to test patient samples having the following characteristics:

<table>
<thead>
<tr>
<th>Sample Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cancer</td>
<td>Mixed type carcinoma; Malignant;</td>
</tr>
<tr>
<td>Cancer</td>
<td>Invasive, predominant intraductal component (8500/3)</td>
</tr>
<tr>
<td>Cancer</td>
<td>Fibrocystic Changes; Invasive lobular carcinoma - 8520/3; Lobular carcinoma in situ - 8520/2; Benign; In situ and grade 3 intraepith; Malignant; Fat necrosis, periductal inflammation, malignant cells; Fat necrosis; Inflammation; Benign;</td>
</tr>
<tr>
<td>Cancer</td>
<td>Invasive, predominant intraductal component (8500/3)</td>
</tr>
<tr>
<td>Cancer</td>
<td>Mucinous (colloid) adenocarcinoma (8480/3)</td>
</tr>
<tr>
<td>Cancer</td>
<td>Invasive lobular carcinoma - 8520/3; Microcalcifications; Benign; Malignant;</td>
</tr>
<tr>
<td>Cancer</td>
<td>Otherfibrocystic change; Invasive, NOS (8500/3)</td>
</tr>
<tr>
<td>Cancer</td>
<td>Invasive ductal carcinoma, not otherwise specified (NOS) - 8500/3; Malignant;</td>
</tr>
<tr>
<td>Cancer</td>
<td>Invasive ductal carcinoma, not otherwise specified (NOS) - 8500/3; Malignant;</td>
</tr>
<tr>
<td>Cancer</td>
<td>Intraductal carcinoma, non-infiltrating, NOS (in situ) (8500/2)</td>
</tr>
<tr>
<td>Cancer</td>
<td>Atypical lobular hyperplasia</td>
</tr>
<tr>
<td>Cancer</td>
<td>Otherfibrocystic changes, inter and intralobular fibrosis, apocrine metaplasia, columnar cell change, microcalcifications; Invasive, NOS (8500/3)</td>
</tr>
<tr>
<td>Cancer</td>
<td>Fibroadenomal, NOS (8500/3)</td>
</tr>
<tr>
<td>Cancer</td>
<td>Ductal carcinoma in situ - 8520/2; Invasive ductal carcinoma, not otherwise specified (NOS) - 8500/3; Microcalcifications; Benign; In situ and grade 3 intraepith; Malignant;</td>
</tr>
<tr>
<td>Cancer</td>
<td>Ductal carcinoma in situ - 8520/2; Invasive lobular carcinoma - 8520/3; Lobular carcinoma in situ - 8520/2; In situ and grade 3 intraepith; Malignant;</td>
</tr>
<tr>
<td>Cancer</td>
<td>Ductal carcinoma in situ - 8520/2; Invasive ductal carcinoma, not otherwise specified (NOS) - 8500/3; Microcalcifications; Benign; In situ and grade 3 intraepith; Malignant; Focal Micropapillary Features, invasive ductal carcinoma with micropapillary features, invasive ductal carcinoma with mucinous and micropapillary features; Invasive micropapillary carcinoma - 8507/3; Malignant;</td>
</tr>
<tr>
<td>Cancer</td>
<td>Invasive, predominant intraductal component (8500/3)</td>
</tr>
<tr>
<td>Cancer</td>
<td>Invasive ductal carcinoma, not otherwise specified (NOS) - 8500/3; Malignant;</td>
</tr>
<tr>
<td>Cancer</td>
<td>Invasive, NOS (8500/3)</td>
</tr>
<tr>
<td>Cancer</td>
<td>Infiltrating duct and lobular carcinoma (8522/3)</td>
</tr>
<tr>
<td>Cancer</td>
<td>Invasive, predominant in situ component (8520/3)</td>
</tr>
<tr>
<td>Non-Cancer</td>
<td>Otherusual ductal hyperplasia, apocrine metaplasia, microcysts, elastosis</td>
</tr>
</tbody>
</table>
[00944] Microvesicles (and potentially other biological entities) were precipitated in blood (plasma) samples from the above patients using polymer precipitation with PEG as indicated above. The protocol was used to probe the samples with the oligonucleotide probe libraries. Sequences that bound the PEG precipitated samples were identified using next generation sequencing (NGS).

[00945] FIG. 12D shows scatter plots of a selection of results from testing the 40 patients listed previously. The spread in the data indicates that large numbers of oligos were detected that differed between samples. The number of significant oligos found is much greater than would be expected randomly as shown in Table 17. The table shows the number of oligonucleotides sorted by copy number detected and p-value. The d-# indicates the number copies of a sequence observed for the data in the rows.

Table 17: Expected versus observed sequences

<table>
<thead>
<tr>
<th></th>
<th>Total Number</th>
<th>P-0.1</th>
<th>P-0.05</th>
<th>P-0.01</th>
<th>P-0.005</th>
</tr>
</thead>
<tbody>
<tr>
<td>d-50</td>
<td>83,632</td>
<td>47,020</td>
<td>30,843</td>
<td>5,934</td>
<td>2,471</td>
</tr>
<tr>
<td>d-100</td>
<td>52,647</td>
<td>29,106</td>
<td>19,446</td>
<td>3,893</td>
<td>1,615</td>
</tr>
<tr>
<td>d-200</td>
<td>28,753</td>
<td>14,681</td>
<td>9,880</td>
<td>2,189</td>
<td>914</td>
</tr>
<tr>
<td>d-500</td>
<td>10,155</td>
<td>4,342</td>
<td>2,927</td>
<td>725</td>
<td>315</td>
</tr>
<tr>
<td>d-50</td>
<td>100.0%</td>
<td>56.2%</td>
<td>36.9%</td>
<td>7.1%</td>
<td>3.0%</td>
</tr>
<tr>
<td>d-100</td>
<td>100.0%</td>
<td>55.3%</td>
<td>36.9%</td>
<td>7.4%</td>
<td>3.1%</td>
</tr>
<tr>
<td>d-200</td>
<td>100.0%</td>
<td>51.1%</td>
<td>34.4%</td>
<td>7.6%</td>
<td>3.2%</td>
</tr>
<tr>
<td>d-500</td>
<td>100.0%</td>
<td>42.8%</td>
<td>28.8%</td>
<td>7.1%</td>
<td>3.1%</td>
</tr>
<tr>
<td>Maximum expected</td>
<td>10.0%</td>
<td>5.0%</td>
<td>1.0%</td>
<td>0.5%</td>
<td></td>
</tr>
</tbody>
</table>

[00946] As a control, the cancer and non-cancer samples were randomly divided into two groups. Such randomization of the samples significantly reduced the number of oligos found that differentiate between sample groups. Indeed, there was a 50-fold increase in informative oligos between the cancer/non-cancer
grouping versus random grouping. FIG. 12E shows data as in Table 17 and indicates the number of observed informative oligos between the indicated sample groups.

[00947] FIG. 12F shows distinct groups of oligos that differentiate between cancer and non-cancer samples. The figure shows a heatmap of the 40 samples tested with oligos selected that had more than 500 copies and p-value less than 0.005. There are clear subpopulations emerging with a distinct non-cancer cohort at the top. The non-cancer samples have boxes around them on the left axis. FIG. 12G is similar and shows results with an additional 20 cancer and 20 non-cancer samples. As shown, analysis with the 80 samples provides the emergence of more distinct and larger clusters.

[00948] The data for the additional 80 samples was also used to compare the consistency of informative oligos identified in different screening experiments. Of the 315 informative oligos identified using the first set of 40 patients, 86% of them showed fold-change in a consistent manner when tested on the independent set of 40 patients.

**Example 22: Enrichment of Oligonucleotide probes to Breast Cancer (BrCa) and non-BrCa derived microvesicles using a balanced library design**

[00949] In this Example, a naive ADAPT oligonucleotide library was screened to enrich oligonucleotides that identify microvesicles circulating in the blood of breast cancer patients and microvesicles circulating in the blood of healthy, control individuals (i.e., without breast cancer). The input library was the naive F-TRin-35n-B 8-3s library, which comprises a 5′ region (5′ CTAGCATGACTGCAGTACGT(SEQ ID NO. 131)) followed by the random naive aptamer sequences of 35 nucleotides and a 3′ region (5′ CTGTCTTTATACACATC≥GACGTGGCCGACGA (SEQ ID NO. 132)). The "balanced" design is described in Example 23 of Int'l Patent Publication WO/2015/031694 (Appl. No. PCT/US20 14/053306, filed August 28, 2014). The working library comprised approximately 2 x 10^13 synthetic oligonucleotide sequences. The naive library may be referred to as the "L0 Library" herein.

[00950] The L0 Library was enriched against fractionated plasma samples from breast cancer patients and from healthy (non-breast cancer) controls using the protocol shown in FIG. 21A. In Step 1, an aliquot of approximately 10^{11} sequences of PCR-amplified L0 was incubated with pooled blood-plasma from 59 breast cancer patients with positive biopsy (represented by "Source A" in FIG. 21A). In parallel, another aliquot of 10^{11} sequences was incubated with pooled blood-plasma from 30 patients with suspected breast cancer who proved negative on biopsy and 30 self-declared healthy women (represented by "Source B" in FIG. 21A). In Step 2, microvesicles (extracellular vesicles, "EV") were precipitated using ultracentrifugation (UC) from both LO-samples. The EV-associated oligodeoxynucleotides (ODNs) were recovered from the respective pellets. In Step 3, a counter-selection step (Step 3) was carried out by incubation of each enriched library with plasma from the different cohorts to drive the selection pressure towards enrichment of ODNs specifically associated with each sample cohort. In this step, sequences contained in the EV pellets were discarded. In Step 4, a second positive selection was performed. In this step, the sequences contained in the respective supernatants (sn) from Step 3 were mixed with plasma from another aliquot of each positive control sample-population, and EVs were again isolated. EV-associated ODNs were recovered, representing two single-round libraries called library L1 for positive
enrichment of cancer (positive biopsy) patients, and library L2 for the positive enrichment against control patients. In a final step, L1 and L2 were amplified by PCR, reverted to single stranded DNA (ssDNA), and mixed to yield library L3.

[00951] This enrichment scheme was iterated two times more using L3 as the input to further reduce the complexity of the profiling library to approximately $10^6$ different sequences. In Step 2, UC was used for partitioning of microvesicles, which may increase the specificity for the EV fraction. In Steps 3 and 4, partitioning was performed using PEG-precipitation. This procedure enriches for ODNs specific for each biological source. Library L3 contains those ODNs that are associated with targets characteristic for EV-populations from both sources, i.e. ODNs acting as aptamers that bind to molecules preferentially expressed in each source. A total of biopsy-positive ($n = 59$), biopsy-negative ($n = 30$), and self-declared normal ($n = 30$) were used in the first round of L3 enrichment, while only the cancer and non-cancer samples were used in the subsequent rounds.

[00952] The enriched libraries were characterized using next-generation-sequencing (NGS) to measure copy numbers of sequences contained in each profiling library. NGS of L0 shows that the vast majority of sequences existed in low copy numbers, whereas libraries L1 and L2 showed significantly higher average counts per sequence (FIG. 21B) and a reduced amount of different sequences, with unaltered total valid reads, (FIG. 21C) consistent with an enrichment process.

Example 23: Analysis of ADAPT-identified biomarkers

[00953] As described herein, e.g., in the section entitled "Aptamer Target Identification," an unknown target recognized by an aptamer can be identified. In this Example, an oligonucleotide probe library (also referred to as Adaptive Dynamic Artificial Poly-ligand Targeting (ADAPT) libraries or Topographical Oligonucleotide Probe "TOP" libraries) was developed as described here and targets of the screened oligonucleotides were determined. This Example used a ADAPT library generated by enriching microvesicles collected from the blood of breast cancer patients and normal controls (i.e., non-cancer individuals). The enrichment protocols are described herein in Example 22.

[00954] Materials & Methods

[00955] SBED library conjugation

[00956] A naive F-TRin-35n-B 8-3s library was enriched against microvesicles from normal female plasma. The naive unenriched library comprised a 5’ region (5’ CTAGCATGACTGCAGTACGT (SEQ ID NO. 131)) followed by the random naive aptamer sequences of 35 nucleotides and a 3’ region (5’ CTGTCTCTTATACACATCT GACGCTTGGCGACGA (SEQ ID NO. 132)). The naive library may be referred to as the "L0 Library" herein and the enriched library referred to as the "L2 library." See Example 22. The screened library was PCR amplified with a C6-amine sense primer (C6 Amine-5’ CTAGCATGACTGCAGTACGT 3’ (SEQ ID NO. 131)) and a 5’ phosphorylated anti-sense primer (5’ Phos TCGTCGGCAGCGTCA (SEQ ID NO. 133)), the purified product was strand separated and conjugated with sulfo-SBED (Thermo Scientific) according to Vinkenborg et al. (Angew Chem Int Ed Engl. 2012, 51:9176-80) with the following modifications: The reaction was scaled down to 5µg C6-amine DNA library (8.6 µM) in 25mM HEPES-KOH, 0.1M NaCl, pH 8.3 and incubated with either 100-fold molar
excess of sulfo-SBED or DMSO in a 2 lµL volume for 30 min at room temp in the dark. The SBED-
conjugated library was immediately separated from the unconjugated library and free sulfo-SBED by
injection onto a Waters X-BridgeTM OST C-18 column (4.6 mm x 50 mm) and fractionated by HPLC
(Agilent 1260 Infinity) with a linear gradient Buffer A: 100 mM TEAA, pH7.0, 0% ACN to 100 mM
TEAA, pH7.0, 25% ACN at 0.2ml/min, 65 °C. There SBED-conjugated fractions were desalted into water
with Glen Gel-PakTM Cartridges and concentrated by speed-vac. SBED conjugation was confirmed by
LC-MS and/or a dot blot with streptavidin-HRP detection.

[00957] Binding reaction and cross-linking
[00958] SBED library functionalization was tested by performing the ADP1T assay with SBED vs
DMSO mock conjugated control C6-amine library and sequenced on a HiSeq 2500TM (Illumina Corp.).
The aptamer precipitation was performed with forty-eight ADP1T reactions incubated for 1hr with end-
over-end rotation at room temp with a 5ng input of SBED conjugated library per 200 µL of plasma (pre-
spun to remove cellular debris at 10,000 xg for 20min, 4 °C) in IX PBS, 3mM MgCl2, 0.01mM dextran
sulfate, 40 ng/µg salmon sperm DNA and 40 ng/µg yeast transfer RNA, and cOmplete ULTRA Mini
EDTA-free TM protease inhibitors (Roche) equivalent to ~240ng library and 9.6 mis plasma. A duplicate
set of 48 reactions was prepared with the DMSO control C6-amine library. Aptamer library-protein
complexes were precipitated with incubation in 6% PEG8000 for 15 min at 4 °C then centrifuged at
10,000 xg for 5 min. Pellets were washed with 1ml 1x PBS, 3mM MgCl2 by gentle inversion to remove
unbound aptamers. The washed pellets were resuspended in 100µL of water and subjected to photo-cross-
linking at 365nm with a hand-held 3UV (254NM/302NM/365NM) lamp, 115 volts (Thermo Scientific)
for 10 min on ice with 1-2 cm between the 96-well plate and lamp.

[00959] Oligonucleotide precipitation
[00960] Cross-linked reactions were subsequently pooled (~4.8ml) per library or 4.8 ml of IX PBS (AP
bead only control) and incubated with 10 µL of Prepared Dynabeads® MyOne™ Streptavidin C1
(10mg/ml) (Life Technologies) (pre-washed with IX PBS, 0.01% Triton X-100) shaking for 1 hr at room
temp. Beads were transferred to an eppendorf tube and lysed for 20 min with lysis buffer (50 mM Tris-
HCl, 10mM MgCl2, 200mM NaCl, 0.5% Triton X-100, 5% glycerol, pH 7.5) on ice, washed 3 times with
wash buffer 1 (100mM Tris-HCl, 1mM EDTA, 2M NaCl, 1% Triton X-100), followed by 2 times with
wash buffer 2 (100mM Tris-HCl, 1mM EDTA, 2M NaCl, 0.01% Triton X-100) as described by
Vinkenborg et al. (Angew Chem Int Ed Engl. 2012, 51:9176-80). Cross-linked proteins were eluted by
boiling 15 min in IX LDS sample buffer with reducing agent added (Life Technologies) and loaded on a
4-12% SDS-PAGE gradient gel (Life Technology). Proteins and DNA were detected with double staining
with Imperial Blue Protein Stain (Thermo Scientific) followed by Prot-SIL2 TM silver stain kit (Sigma)
used according to manufacturer’s instructions in order to enhance sensitivity and reduce background.

[00961] Protein identification
[00962] Protein bands that appeared to differ between the cancer and normal were excised from the
gradient gels and subjected to liquid chromatography-tandem mass spectrometry (LC-MS/MS).

[00963] Results
ADAPT protein targets were identified from bands cut from a silver stained SDS-PAGE gel (FIG. 13). Apatam-SBED complexes (lane 3) or Apatam-DMSO complexes (control-lane 4) were precipitated with 6% PEG8000, subjected to UV photo-cross-linking, and pulled-down with Streptavidin coated beads. Eluate was analyzed under reducing conditions by SDS-PAGE and silver staining. Apatam library alone (50ng) (lane 1) was loaded as a control for migration of the library (second to bottom arrows) and an equal volume of eluate from a bead only sample (lane 4) was loaded as a streptavidin control to control for potential leaching of the streptavidin monomer (bottom arrow) under the harsh elution conditions. Upper arrows ("Targets") indicate specific or more predominant bands identified with the SBED-conjugated library vs. the mock DMSO treated control C6-amine library. Indicated target protein bands were cut out and sent for LC-MS/MS protein identification or indicated DNA library bands were eluted, reamplified and sequenced. The identified proteins are those that appeared as upregulated in the normal samples.

Tables 18-25 list human proteins that were identified in 8 bands excised from the silver stained gel. In all tables the proteins are those identified in the oligo-SBED complexes with proteins identified in the corresponding control lanes removed. The band numbers in the tables indicate different bands cut from the gel (FIG. 13). Accession numbers in the table are from the UniProt database (www.uniprot.org). "GN=" is followed by the gene name. Various protein classifications indicated in the Tables 18-25 include Nucleic Acid Binding Proteins (NAB), Tumor suppressors (TS), cell adhesion/cytoskeletal (CA/CK) and abundant plasma proteins (ABP). In Table 18, the proteins listed below the row "SBED associated" were identified by peptide fragments linked to SBED, indicating that the peptides are from near the cross-linking sites. Class is not reported for these proteins.

**Table 18: Band 3**

<table>
<thead>
<tr>
<th>Accession number</th>
<th>Class</th>
<th>Protein name</th>
</tr>
</thead>
<tbody>
<tr>
<td>P02538</td>
<td>CA/CK</td>
<td>Keratin, type II cytoskeletal 6A GN=KRT6A</td>
</tr>
<tr>
<td>P15924</td>
<td>CA/CK</td>
<td>Desmoplakin GN=DSP</td>
</tr>
<tr>
<td>P04259</td>
<td>CA/CK</td>
<td>Keratin, type II cytoskeletal 6B GN=KRT6B</td>
</tr>
<tr>
<td>P60709</td>
<td>CA/CK</td>
<td>Actin, cytoplasmic 1 GN=ACTB</td>
</tr>
<tr>
<td>P20930</td>
<td>CA/CK</td>
<td>Filaggrin GN=FLG</td>
</tr>
<tr>
<td>P07476</td>
<td>CA/CK</td>
<td>Involucrin GN=IVL</td>
</tr>
<tr>
<td>P31947</td>
<td>TS</td>
<td>14-3-3 protein sigma GN=SFN</td>
</tr>
<tr>
<td>Q7Z794</td>
<td>CA/CK</td>
<td>Keratin, type II cytoskeletal 1b GN=KRT77</td>
</tr>
<tr>
<td>P02545</td>
<td>NAB</td>
<td>Prelamin-A/C GN=LMNA</td>
</tr>
<tr>
<td>P19012</td>
<td>CA/CK</td>
<td>Keratin, type I cytoskeletal 15 GN=KRT15</td>
</tr>
<tr>
<td>P47929</td>
<td>CA/CK &amp; TS</td>
<td>Galectin-7 GN=LGALS7</td>
</tr>
<tr>
<td>P11142</td>
<td></td>
<td>Heat shock cognate 71 kDa protein GN=HSPA8</td>
</tr>
<tr>
<td>P58107</td>
<td>NAB</td>
<td>Epilplakin GN=EPPK1</td>
</tr>
<tr>
<td>P08107</td>
<td></td>
<td>Heat shock 70 kDa protein 1A/1B GN=HSPA1A</td>
</tr>
<tr>
<td>Q02413</td>
<td>CA/CK</td>
<td>Desmoglein-1 GN=DSG1</td>
</tr>
<tr>
<td>P06396</td>
<td>CA/CK</td>
<td>Gelsolin GN=GSN</td>
</tr>
<tr>
<td>0608 14</td>
<td>NAB</td>
<td>Histone H2B type 1-K GN=HIST1H2BK</td>
</tr>
<tr>
<td>P68104</td>
<td>NAB</td>
<td>Elongation factor 1-alpha 1 GN=EEF 1A1</td>
</tr>
<tr>
<td>P05387</td>
<td>NAB</td>
<td>60S acidic ribosomal protein P2 GN=RPLP2</td>
</tr>
<tr>
<td>Q7RTS7</td>
<td>CA/CK</td>
<td>Keratin, type II cytoskeletal 74 GN=KRT74</td>
</tr>
<tr>
<td>P31946</td>
<td>TS</td>
<td>14-3-3 protein beta/alpha GN=YWHAB</td>
</tr>
<tr>
<td>Accession</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>Q13835</td>
<td>CA/CK Plakophilin-1 GN=PKP1</td>
<td></td>
</tr>
<tr>
<td>P14923</td>
<td>CA/CK Junction plakoglobin GN=JUP</td>
<td></td>
</tr>
<tr>
<td>P09651</td>
<td>NAB Heterogeneous nuclear ribonucleoprotein A1 GN=HNRNPA1</td>
<td></td>
</tr>
<tr>
<td>P07900</td>
<td>NAB Heat shock protein HSP 90-alpha GN=HSP90AA1</td>
<td></td>
</tr>
<tr>
<td>Q96KK5</td>
<td>NAB Histone H2A type 1-H GN=HIST1H2AH</td>
<td></td>
</tr>
<tr>
<td>P04406</td>
<td>CA/CK Glyceraldehyde-3-phosphate dehydrogenase GN=GapDH</td>
<td></td>
</tr>
<tr>
<td>P10412</td>
<td>NAB Histone HI 4 GN=HIST1H1E</td>
<td></td>
</tr>
<tr>
<td>P04792</td>
<td>NAB Heat shock protein beta-1 GN=HSPB1</td>
<td></td>
</tr>
<tr>
<td>Q9NKT1</td>
<td>CA/CK Calmodulin-like protein 5 GN=CALML5</td>
<td></td>
</tr>
<tr>
<td>P41505</td>
<td>NAB Permaminogen activator inhibitor 1 RNA-binding protein GN=SERBP1</td>
<td></td>
</tr>
<tr>
<td>P07237</td>
<td>NAB Protein disulfide-isomerase GN=P4HB</td>
<td></td>
</tr>
<tr>
<td>060437</td>
<td>CA/CK Periplakin GN=PPL</td>
<td></td>
</tr>
<tr>
<td>P01771</td>
<td>ABP Ig lambda chain V-IV region Hil</td>
<td></td>
</tr>
<tr>
<td>P55884</td>
<td>NAB Eukaryotic translation initiation factor 3 subunit B GN=EIF3B</td>
<td></td>
</tr>
<tr>
<td>P11021</td>
<td>NAB 78 kDa glucose-regulated protein GN=HSPA5</td>
<td></td>
</tr>
<tr>
<td>P01024</td>
<td>NAB Complement C3 GN=C3</td>
<td></td>
</tr>
<tr>
<td>P04350</td>
<td>CA/CK Tubulin beta-4A chain GN=TUBB4A</td>
<td></td>
</tr>
<tr>
<td>P01857</td>
<td>ABP Ig gamma-1 chain C region GN=IGHG1</td>
<td></td>
</tr>
<tr>
<td>P61247</td>
<td>NAB 40S ribosomal protein S3a GN=RPS3A</td>
<td></td>
</tr>
<tr>
<td>P62937</td>
<td>Peptidyl-prolyl cis-trans isomerase A GN=PPIA</td>
<td></td>
</tr>
<tr>
<td>015020</td>
<td>CA/CK Spectrin beta chain, non-erythrocytic 2 GN=SPTBN2</td>
<td></td>
</tr>
<tr>
<td>P30401</td>
<td>NAB Protein disulfide-isomerase A3 GN=PDIA3</td>
<td></td>
</tr>
<tr>
<td>Q6KB66</td>
<td>CA/CK Keratin, type II cytoskeletal 80 GN=KRT80</td>
<td></td>
</tr>
<tr>
<td>Q9UJ6</td>
<td>CA/CK Drebrin-like protein GN=DBNL</td>
<td></td>
</tr>
<tr>
<td>P47914</td>
<td>NAB 60S ribosomal protein L29 GN=RPL29</td>
<td></td>
</tr>
<tr>
<td>P39023</td>
<td>NAB 60S ribosomal protein L3 GN=RPL3</td>
<td></td>
</tr>
<tr>
<td>A6NY6</td>
<td>CA/CK Putative annexin A2-like protein GN=ANXA2P2</td>
<td></td>
</tr>
<tr>
<td>P60174</td>
<td>CA/CK Triosephosphate isomerase GN=TP1I</td>
<td></td>
</tr>
<tr>
<td>P35421</td>
<td>CA/CK Radixin GN=RDX</td>
<td></td>
</tr>
<tr>
<td>P07305</td>
<td>NAB Histone HI 0 GN=H1F0</td>
<td></td>
</tr>
<tr>
<td>P15259</td>
<td>CA/CK Phosphoglycerate mutase 2 GN=PGAM2</td>
<td></td>
</tr>
<tr>
<td>P0CG05</td>
<td>ABP Ig lambda-2 chain C regions GN=IGLC2</td>
<td></td>
</tr>
<tr>
<td>Q92817</td>
<td>CA/CK Envoplakin GN=EVPL</td>
<td></td>
</tr>
<tr>
<td>P06733</td>
<td>NAB MBP-1 of Alpha-enolase GN=EN01</td>
<td></td>
</tr>
<tr>
<td>P22626</td>
<td>NAB Heterogeneous nuclear ribonucleoproteins A2/B 1 GN=HNRNPA2B1</td>
<td></td>
</tr>
<tr>
<td>P62424</td>
<td>NAB 60S ribosomal protein L7a GN=RPL7A</td>
<td></td>
</tr>
<tr>
<td>P60660</td>
<td>NAB 60S ribosomal protein L7a GN=RPL7A</td>
<td></td>
</tr>
<tr>
<td>P04083</td>
<td>NAB Annexin A1 GN=ANXA1</td>
<td></td>
</tr>
<tr>
<td>Q14134</td>
<td>NAB Tripartite motif-containing protein 29 GN=TRIM29</td>
<td></td>
</tr>
<tr>
<td>P39019</td>
<td>NAB 40S ribosomal protein S19 GN=RPS19</td>
<td></td>
</tr>
<tr>
<td>Q8WVV4</td>
<td>CA/CK Protein POF1B GN=POF1B</td>
<td></td>
</tr>
<tr>
<td>Q02887</td>
<td>NAB 60S ribosomal protein L6 GN=RPL6</td>
<td></td>
</tr>
<tr>
<td>Q9YX9</td>
<td>NAB MORC family CW-type zinc finger protein 2 GN=MORC2</td>
<td></td>
</tr>
<tr>
<td>Q9NQC3</td>
<td>NAB Reticulin-4 GN=RTN4</td>
<td></td>
</tr>
<tr>
<td>Q5T753</td>
<td>CA/CK Late cornified envelope protein 1E GN=CA/CK E</td>
<td></td>
</tr>
</tbody>
</table>

**SBED associated**

<table>
<thead>
<tr>
<th>Accession</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>P56202</td>
<td>Neutrophil gelatinase-associated lipocalin precursor</td>
</tr>
<tr>
<td>P80188</td>
<td>Rho GTPase-activating protein 5</td>
</tr>
</tbody>
</table>

-238-
### Table 19: Band 9

<table>
<thead>
<tr>
<th>Accession number</th>
<th>Class</th>
<th>Protein name</th>
</tr>
</thead>
<tbody>
<tr>
<td>P61626</td>
<td></td>
<td>Lysozyme C GN=LYZ</td>
</tr>
<tr>
<td>Q9HCK1</td>
<td>NAB</td>
<td>DBF4-type zinc finger-containing protein 2 GN=ZDBF2</td>
</tr>
</tbody>
</table>

### Table 20: Band 1

<table>
<thead>
<tr>
<th>Accession number</th>
<th>Class</th>
<th>Protein name</th>
</tr>
</thead>
<tbody>
<tr>
<td>P01834</td>
<td>ABP</td>
<td>Ig kappa chain C region GN=IGKC</td>
</tr>
<tr>
<td>P01765</td>
<td>ABP</td>
<td>Ig heavy chain V-III region TIL</td>
</tr>
<tr>
<td>P04003</td>
<td>NAB</td>
<td>C4b-binding protein alpha chain GN=C4BPA</td>
</tr>
<tr>
<td>P60709</td>
<td>CA/CK</td>
<td>Actin, cytoplasmic 1 GN=ACTB</td>
</tr>
<tr>
<td>Q5T751</td>
<td>CA/CK</td>
<td>Late cornified envelope protein 1C GN=LCE1C</td>
</tr>
</tbody>
</table>

### Table 21: Band 5

<table>
<thead>
<tr>
<th>Accession number</th>
<th>Class</th>
<th>Protein name</th>
</tr>
</thead>
<tbody>
<tr>
<td>P01860</td>
<td>ABP</td>
<td>Ig gamma-3 chain C region GN=IGHG3</td>
</tr>
<tr>
<td>O60902</td>
<td>NAB</td>
<td>Short stature homeobox protein 2 GN=SHOX2</td>
</tr>
</tbody>
</table>

### Table 22: Band 7

<table>
<thead>
<tr>
<th>Accession number</th>
<th>Class</th>
<th>Protein name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q04695</td>
<td>CA/CK</td>
<td>Keratin, type I cytoskeletal 17 GN=KRT17</td>
</tr>
<tr>
<td>Q7Z794</td>
<td>CA/CK</td>
<td>Keratin, type II cytoskeletal 1b GN=KRT77</td>
</tr>
<tr>
<td>Q6KB66</td>
<td>CA/CK</td>
<td>Keratin, type II cytoskeletal 80 GN=KRT80</td>
</tr>
<tr>
<td>P01833</td>
<td></td>
<td>Polymeric immunoglobulin receptor GN=PIGR</td>
</tr>
<tr>
<td>P01042</td>
<td></td>
<td>Kininogen-1 GN=KNG1</td>
</tr>
<tr>
<td>Q02413</td>
<td>CA/CK</td>
<td>Desmoglein-1 GN=DSG1</td>
</tr>
<tr>
<td>P15924</td>
<td>CA/CK</td>
<td>Desmoplakin GN=DSP</td>
</tr>
<tr>
<td>Q8TF72</td>
<td></td>
<td>Protein Shroom3 GN=SHROOM3</td>
</tr>
<tr>
<td>P02671</td>
<td>ABP</td>
<td>Fibrinogen alpha chain GN=FGA</td>
</tr>
<tr>
<td>Q5T749</td>
<td>CA/CK</td>
<td>Keratinocyte proline-rich protein GN=KPRP</td>
</tr>
<tr>
<td>Q5VZP5</td>
<td></td>
<td>Inactive dual specificity phosphatase 27 GN=DUSP27</td>
</tr>
<tr>
<td>Q5T751</td>
<td>CA/CK</td>
<td>Late cornified envelope protein 1C GN=LCE1C</td>
</tr>
<tr>
<td>Q9UL12</td>
<td></td>
<td>Sarcosine dehydrogenase, mitochondrial GN=SAR DH</td>
</tr>
<tr>
<td>P00698</td>
<td></td>
<td>Lysozyme C OS=Gallus gallus GN=LYZ</td>
</tr>
<tr>
<td>Q8N114</td>
<td></td>
<td>Protein shisa-5 GN=SHISA5</td>
</tr>
</tbody>
</table>
Certain proteins were identified in multiple bands. For example, IGLC2 was identified in bands 3, 15 and 19 and SHROOM3 was identified in bands 7, 15, 19. This may be due to degradation products, isoforms or the like. These experiments identified 108 proteins (plus 2 lysozyme controls), comprising 108 proteins (plus 2 lysozyme controls), comprising...

### Table 23: Band 15

<table>
<thead>
<tr>
<th>Accession number</th>
<th>Class</th>
<th>Protein name</th>
</tr>
</thead>
<tbody>
<tr>
<td>P08238</td>
<td>Heat shock protein HSP 90-beta GN=HSP90AB1</td>
<td></td>
</tr>
<tr>
<td>P68104</td>
<td>NAB</td>
<td>Elongation factor 1-alpha 1 GN=EEF1A1</td>
</tr>
<tr>
<td>P02675</td>
<td>ABP</td>
<td>Fibrinogen beta chain GN=FGB</td>
</tr>
<tr>
<td>Q8TF72</td>
<td>Protein Shroom3 GN=SHROOM3</td>
<td></td>
</tr>
<tr>
<td>P0CG05</td>
<td>ABP</td>
<td>Ig lambda-2 chain C regions GN=IGLC2</td>
</tr>
<tr>
<td>P78386</td>
<td>CA/CK</td>
<td>Keratin, type II cuticular Hb5 GN=KRT85</td>
</tr>
<tr>
<td>Q7Z5Y6</td>
<td>Bone morphogenetic protein 8A GN=BMP8A</td>
<td></td>
</tr>
<tr>
<td>O14633</td>
<td>CA/CK</td>
<td>Late cornified envelope protein 2B GN=LCE2B</td>
</tr>
</tbody>
</table>

### Table 24: Band 17

<table>
<thead>
<tr>
<th>Accession number</th>
<th>Class</th>
<th>Protein name</th>
</tr>
</thead>
<tbody>
<tr>
<td>P02538</td>
<td>CA/CK</td>
<td>Keratin, type II cytoskeletal 6A GN=KRT6A</td>
</tr>
<tr>
<td>P01834</td>
<td>ABP</td>
<td>Ig kappa chain C region GN=IGKC</td>
</tr>
<tr>
<td>P06702</td>
<td>Protein S100-A9 GN=S100A9</td>
<td></td>
</tr>
<tr>
<td>P68104</td>
<td>NAB</td>
<td>Elongation factor 1-alpha 1 GN=EEF1A1</td>
</tr>
<tr>
<td>P01024</td>
<td>Complement C3 GN=C3</td>
<td></td>
</tr>
<tr>
<td>P81605</td>
<td>Dermcidin GN=DCD</td>
<td></td>
</tr>
<tr>
<td>P05109</td>
<td>Protein S100-A8 GN=S100A8</td>
<td></td>
</tr>
<tr>
<td>Q5T751</td>
<td>CA/CK</td>
<td>Late cornified envelope protein 1C GN=LCE1C</td>
</tr>
</tbody>
</table>

### Table 25: Band 19

<table>
<thead>
<tr>
<th>Accession number</th>
<th>Class</th>
<th>Protein name</th>
</tr>
</thead>
<tbody>
<tr>
<td>P02768</td>
<td>NAB</td>
<td>Serum albumin GN=ALB</td>
</tr>
<tr>
<td>P0CG05</td>
<td>ABP</td>
<td>Ig lambda-2 chain C regions GN=IGLC2</td>
</tr>
<tr>
<td>P06702</td>
<td>Protein S100-A9 GN=S100A9</td>
<td></td>
</tr>
<tr>
<td>P08238</td>
<td>Heat shock protein HSP 90-beta GN=HSP90AB1</td>
<td></td>
</tr>
<tr>
<td>P60709</td>
<td>CA/CK</td>
<td>Actin, cytoplasmic 1 GN=ACTB</td>
</tr>
<tr>
<td>P13647</td>
<td>CA/CK</td>
<td>Keratin, type II cytoskeletal 5 GN=KRT5</td>
</tr>
<tr>
<td>P01616</td>
<td>ABP</td>
<td>Ig kappa chain V-II region MIL</td>
</tr>
<tr>
<td>Q86Y23</td>
<td>CA/CK</td>
<td>Hornerin GN=HRNR</td>
</tr>
<tr>
<td>P01857</td>
<td>ABP</td>
<td>Ig gamma-1 chain C region GN=IGHG1</td>
</tr>
<tr>
<td>P62805</td>
<td>NAB</td>
<td>Histone H4 GN=HIST1H4A</td>
</tr>
<tr>
<td>P59665</td>
<td>Neutrophil defensin 1 GN=DEFA1</td>
<td></td>
</tr>
<tr>
<td>P61626</td>
<td>NAB</td>
<td>Lysozyme C GN=LYZ</td>
</tr>
<tr>
<td>P01024</td>
<td>ABP</td>
<td>Complement C3 GN=C3</td>
</tr>
<tr>
<td>Q8TF72</td>
<td>Protein Shroom3 GN=SHROOM3</td>
<td></td>
</tr>
<tr>
<td>P83593</td>
<td>ABP</td>
<td>Ig kappa chain V-IV region STH (Fragment)</td>
</tr>
<tr>
<td>P01700</td>
<td>ABP</td>
<td>Ig lambda chain V-I region HA</td>
</tr>
<tr>
<td>P01877</td>
<td>ABP</td>
<td>Ig alpha-2 chain C region GN=IGHA2</td>
</tr>
<tr>
<td>Q9UL12</td>
<td>Sarcosine dehydrogenase, mitochondrial GN=SARDH</td>
<td></td>
</tr>
<tr>
<td>Q6NXT2</td>
<td>NAB</td>
<td>Histone H3.3C GN=H3F3C</td>
</tr>
<tr>
<td>P02788</td>
<td>NAB</td>
<td>Lactotransferrin GN=LTF</td>
</tr>
<tr>
<td>P02787</td>
<td>ABP</td>
<td>Serotransferrin GN=TF</td>
</tr>
</tbody>
</table>
among others 34 Nucleic Acid Binding Proteins (NAB) where 7 of the 34 are putative tumor suppressors/repressors; 37 cell adhesion/cytoskeletal (CA/CK); and 14 abundant plasma proteins (ABP).

All of the tumor suppressors/repressors are DNA/RNA binding proteins. Other proteins comprise chaperones, signaling molecules etc.

[00967] For comparison to the above pull-down results with the enriched L2 library, mass spectrometry was used to identify proteins in whole plasma samples ("neat plasma") (Table 26, row "Neat"), in PEG enriched plasma fractions (comprising a microvesicle-containing fraction) (Table 26, row "PEG"), and in unenriched L0 library pull-down (Table 26, row "Unenriched") using the same healthy (non-breast cancer) plasma samples as described above (column "Plasma," rows "Non-cancer" in the table).

Descriptions in the table comprise the gene name followed by accession numbers from the UniProt database (www.uniprot.org) in parentheses. 81 proteins were pulled down by L2 (Tables 18-25) that were not detectable in the PEG-precipitated and neat plasma, indicating that these proteins became enriched due to their interaction with L2 oligonucleotide probes. Further details are found in Example 28 of International Patent Application PCT/US 15/62184, filed November 23, 2015, which application is incorporated by reference herein in its entirety. The methods above were further repeated to identify proteins that were identified in cancer samples using pull-down with the L1 library (column "Plasma," rows "Cancer" in the table). With the proteins identified in the cancer plasma samples, the row "Non-SBED" and "SBED refer to proteins identified by peptide fragments that were unassociated with or associated with SBED, respectively.

Table 26: Proteins identified in plasma

<table>
<thead>
<tr>
<th>Plasma</th>
<th>Sample</th>
<th>Description (Gene Name (Accession))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-cancer</td>
<td>Neat</td>
<td>ALB (P02768), C3 (P01024), APOB (P04114), TF (P02787), A2M (P01023), C4B (P00450), C4A (P00450), FGA (P20671), FGB (P20675), CP (P00450), IGKC (P01834), IGHG1 (P01857), HP (P00738), FGG (P02679), IGHG3 (P01860), CFH (P08603), GC (P02774), IGHG2 (P01859), SERPINA1 (P01009), ITIH4 (P14624), SERPINC1 (P01008), IGHM (P01871), HPR (P00739), CFB (P00751), APOA1 (P02647), FN1 (P02751), IGL5 (B9A064), PLG (P00747), A1BG (P04217), ITIH2 (P19823), SERPINA3 (P01011), KNG1 (P01042), IGHG4 (P01861), APOA4 (P06727), F2 (P00734), HPX (P02790), C4BPA (P04003), ITIH1 (P19827), IGHA1 (P01876), IGLC2 (P02649), SERPINA1 (P01009), APOH (P02749), HRG (P04196), APCS (P02743), AFM (P43652), IGHA2 (P01877), APOE (P02649), AHSG (P02765), SERPIND1 (P05546), C5 (P01031), AMBP (P02760), ORM1 (P02763), AZGP1 (P25311), KRT1 (P04264), CFI (P05156), HBB (P06871), CD5L (P10310), C9 (P02748), RBP4 (P02753), VTN (P04004), SERPINA6 (P08185), SERPINF2 (P08697), CLU (P01090), KRT2 (P13590), C1Q (P02747), TTR (P02766), Ig kappa chain V-II region TEW (P01617), Ig heavy chain V-III region BUT (P01767), Ig heavy chain V-III region TUR (P01779), APOA2 (P02652), ORM2 (P01652), CFHR1 (Q03591), PGLYRP2 (Q96PD5), LYZ (P00698), Ig kappa chain V-IV region Len (P01625), Ig heavy chain V-III region GAL (P01781), IGHD (P01880), C1QB (P02746), C8A (P07357), C8G (P07360), KRT10 (P13645), APOL1 (014791), IGI (P01591), KLKB1 (P03952), Ig lambda chain V region 4A (P04211), Ig kappa chain V-III region VG (Fragment) (P04433), HBA1 (P069905), APOM (095445), C1R (P00736), F12 (P00748), Ig lambda chain V-III region SH (P01714), Ig lambda chain V-IV region Hil (P01717), Ig heavy chain V-III region WEA</td>
</tr>
<tr>
<td>Non-cancer</td>
<td>Unenriched</td>
<td>MOL (P06889), Ig heavy chain V-III region GAR (P80419), FCGBP (Q9Y6R7), APOM (Q02413-2), F13A1 (P00488), Ig heavy chain V-I region HG3 (P01743), CIQA (P02745), Ig lambda chain V-VI region WLT (P06318), C2 (P06681), C4BPB (P20851), CFP (P27918), SERPINA4 (P29622), SAA4 (P35542), SERPINF1 (P36955), LGALS3BP (Q80380), HABP2 (Q14520), RCBTB1 (Q8NDN9), APOL1 (Q01791), KCNQ2 (043526), F9 (P00740), Ig heavy chain V-III region TRO (P01762), Ig kappa chain V-III region HIL (P01771), Ig heavy chain V-II region OU (P01814), APOA2 (P02652), F11 (P03951), Ig lambda chain V-I region WAH (P04208), Ig lambda chain V region 4A (P04211), Ig kappa chain V-II region RPMI 6410 (P06310), Ig kappa chain V-III region IARC/BL41 (P06311), KRT5 (P13647), IGLL1 (P15814), Ig heavy chain V-I region V35 (P23083), HBA1 (P69905), ADIPOQ (Q15488), PGLYRP2 (Q96PD5), UPF3A (Q9H1J1), BCO1 (Q9HAY6), ARFGAP3 (Q9NP61), SARDH (Q9UL12), SERPINA1 (P01009), KNG1 (P01042), Ig kappa chain V-I region Kue (P01604), Ig kappa chain V-I region Lay (P01605), Ig kappa chain V-I region OU (P01606), Ig kappa chain V-II region MIL (P01616), Ig heavy chain V-III region VH26 (P01764), Ig heavy chain V-III region GA (P01769), FN1 (P027510), TTR (P02766), SERPING1 (P05155), APOA4 (P06727), PRSS1 (P07477), ANX46 (P08133), CFTR (P13569), LBP (P18428), FBLN1 (P23142), SPAG17 (Q6Q759), PDLIM2 (Q96LY6), ARHGEF17 (Q96PE2), IGLC7 (A0M8Q6), AGRN (Q00468), AGT (P01019), RBP4 (P02753), AHSG (P02765), Ig kappa chain V-III region GOL (P04206), SERPINA5 (P05154), GSN (P06396), Ig kappa chain V-III region HAH (P18135), CFHR2 (P36980), GIT2 (Q141611), INCAPC (Q9NQ57)</td>
</tr>
<tr>
<td>Cancer</td>
<td>Non-SBED</td>
<td>MUC5B (Q9HC84), FABP5 (Q01469), HPX (P02790), CP (P00450), SPRR2E (P22531), SPRR2D (P22532), PDE4D (Q08499-7), GC (P02774-2), CPD (Q07596), CD14 (P08571), LAP3 (P28838-2), AFM (P43652), FCN2 (Q15485-2), DMBT1 (Q9UGM3-9), LIFR (P42702), SNX27 (Q96L92-3), LCN1 (P31025), ARFIP1 (P53367-2), APOH (P02749), KLLB1 (P03952), XP32 (Q57750), H2AFV (Q7IU19-5), KRT75 (Q05678), KRT6C (P48668), KRT83 (P78385), KRT76 (Q01546), KRT33B (Q14525), KRT72 (Q14CN4-3), KRT31 (Q15323), KRT73 (Q86Y46-2), DSG1 (Q02413-2), LCE1C (Q5T751), LCE1A (Q5T7P2), CFB (P00751), CFH (P08603), SERPINA1 (P01009-2), Ig kappa chain V-I region EU (P01598), Ig kappa chain V-II region MIL (P01616), Ig lambda chain V-I region Bau (P01715), Ig heavy chain V-III region GAL (P01781), IGLC6 (P0CF74) and ACTG2 (P63267-2)</td>
</tr>
<tr>
<td>Cancer</td>
<td>SBED</td>
<td>Probable ubiquitin carboxyl-terminal hydrolase FAF-Y (Q00507-1; Q00507-2); Homeobox protein Hex-C12 (P31275); putative uncharacterized protein C18orf65 (Q6ZTR6); Piwi-like protein 2 (Q8TC59-2; Q8TC59-1); Probable ubiquitin carboxyl-terminal hydrolase FAF-X (Q93008-1; Q93008-3); Histone-lysine N-methyltransferase ASH1L (Q9NR48; Q9NR48-2); PNMA-like protein 2 (Q9ULN7-5)</td>
</tr>
</tbody>
</table>

[00968] The biomarkers in this Example can be used to detect microvesicles that are indicative of cancer or non-cancer samples.
Example 24: Identification of biomarkers through affinity enrichment with an enriched oligonucleotide library and mass spectrometry

[00969] This Example continues upon the Example above. Identification of protein-protein and nucleic acid-protein complexes by affinity purification mass spectrometry (AP-MS) can be hampered in samples comprising complex mixtures of biological components (e.g., bodily fluids including without limitation blood and derivatives thereof). For example, it may be desirable to detect low abundance protein and nucleic acid-protein complexes in a complex milieu comprising various components that may interact promiscuously with specific binding sites such as high abundance proteins that interact non-specifically with the affinity resin. AP-MS has been used previously to enrich for pre-identified targets of interest using individual DNA or RNA aptamers or specific nucleic acid binding domains. In this Example, an enriched oligonucleotide probing library was used as the affinity reagent. This approach combined with mass spectrometry enables the identification of differentially expressed biomarker from different disease states or cellular perturbations without relying on a priori knowledge of the targets of interest. Such biomarker may comprise proteins, nucleic acids, miRNA, mRNA, carbohydrates, lipid targets, combinations thereof, or other components in a biological system.

[00970] The method comprises identification of an enriched oligonucleotide probe library according to the methods of the invention followed by target identification with affinity purification of the bound probing library and mass spectrometry. The members of the enriched oligonucleotide probing library comprise an affinity tag. A biological sample is probed with the oligonucleotide probe library, affinity purification of the oligonucleotide probe library via the affinity tag is performed which will accordingly purify biological entities in complex with various members of the probe library, and read-out of targets that purified with the members of the probe library is performed using liquid chromatography-tandem mass spectrometry (LC-MS/MS) for proteins or oligonucleotide targets (e.g., miRNA or mRNA) with next generation sequencing (NGS). Confirmation of protein targets is performed using quantitative mass spectrometry (MS), e.g., using MRM/SRM or SWATH based methods.

[00971] The method of the Example lends itself to various options. For example, any appropriate affinity tags can be used for affinity pull-down, including without limitation anti-sense oligonucleotides, biotin, polyhistidine, FLAG octapeptide (i.e., N-DYKDDDDK-C (SEQ ID NO. 134), where N stands for Aminoterminus and C stands for Carboxy terminus), 3X FLAG, Human influenza hemagglutinin (HA)-tag (i.e., N-YPYDVPDYA-C (SEQ ID NO. 135)), myc-tag (N-EQKLISEEDL-C (SEQ ID NO. 136)), other such as known in the art, and combinations thereof. Similarly, any appropriate enrichment support can be used in addition to the magnetic streptavidin beads exemplified herein, including without limitation other bead systems, agarose beads, planar arrays or column chromatography supports. It follows that the various supports can be coupled with the various affinity reagents appropriate for the oligonucleotide library, including without limitation streptavidin, avidin, anti-His tag antibodies, nickel, and the like. The different affinity tags and supports can be combined as desired. This Example used cross-linking but in certain cases such cross-linking is not necessary and may even be undesirable, e.g., to favor identification of high affinity complex formation. When cross-linking is desired, any appropriate cross-linkers can be used to
carry out the invention, including BS2G, DSS, formaldehyde, and the like. Other appropriate cross-linkers and methods are described herein. See, e.g., Section "Aptamer Target Identification." Lysis buffers and wash stringent lysis/wash conditions may produce a wider array of potential protein complexes of interest whereas more stringent lysis/wash conditions may favor higher affinity oligo-target complexes and/or targets comprising specific proteins (e.g., by disassociating larger complexes bound to the oligos). One of skill will further appreciate that qualitative and/or quantitative LC-MS/MS may be used for target detection and verification. Similarly, metabolic labeling and label-free approaches may be used for quantitative MS, including without limitation spectral counting, SILAC, dimethyl labeling, TMT labeling. Targeted MS with SRM/MRM or SWATH, and the like.

Example 25: Protocol for Affinity capture using oligonucleotide probing library

This Example presents a detailed protocol for the method of affinity capture using an oligonucleotide probing library presented in the Example above.

Protocol:

The oligonucleotide probe library comprises F-TRi-n-35n-B-8-3s described herein either desthiobiotin labeled or unlabeled library and binding to normal (i.e., non-cancer) female plasma. The oligonucleotide probe library is enriched against the plasma samples as described elsewhere (e.g., in Example 21). The plasma samples are processed separately against the desthiobiotin labeled or unlabeled oligonucleotide libraries. General parameters included the following:

48 normal plasma samples are pooled for enrichment of each oligonucleotide library (Desthiobiotin or Unlabeled)

200 μl input plasma per sample

Ultracentrifugation (UC) is used to pre-clear the samples

5 ng of each aptamer library is added to each sample

Binding competitors for all library samples include 0.01X SI (dextran sulfate), 340ng fortRNA and 340 ng Salmon sperm DNA as described elsewhere herein

6% PEG 8000 is used for precipitation of microvesicles within the samples

Affinity purification is performed with Cl Streptavidin beads (MyOne Strptavidin Beads Cl-65001, lot 2ml (10mg/ml))
Buffers:

Plasma dilution: 6 mM MgCl₂ in 2X PBS

Pellet Wash Buffer: 1X PBS, 3mM MgCl₂

PEG Ppt Buffer: 20% Peg8000 in 1X PBS, 3mM MgCl₂

Bead Prep Buffer: 1X PBS containing 0.01% Triton X-100

Lysis Buffer: prepare a 2X stock solution consisting of 100mM Tris-HCl, 20mM MgCl₂, 400mM NaCl, 1% Triton X-100, 10% glycerol, pH 7.5. Diluted to 1X with water 1:1 prior to using.

AP Wash buffer 1: 10mM Tris-HCl, 1mM EDTA, 2M NaCl, 1% Triton X-100, pH 7.5

AP wash buffer 2: 10mM Tris-HCl, 1mM EDTA, 2M NaCl, 0.01% Triton X-100, pH 7.5

Biotin Elution buffer 1: 5mM Biotin, 20mM Tris, 50mM NaCl, pH 7.5

IX LDS, IX Reducing buffer 2

Reagent/Instrument Prep:

Pre-chill Ultracentrifuge to 4 °C.

Protease inhibition: dissolve 2 tablets of "cOmplete ULTRA MINI EDTA-free EASYpack" protease inhibitor in 1100 μl of H2O (20x stock of protease inhibitor).

Plasma Preparation (for each of Desthiobiotin or Unlabeled oligonucleotide libraries):

1. Add 50 μl of protease inhibitor to each ml of sample (on top of frozen plasma) in a room temperature (RT) water bath. Will use 22 ml of pooled plasma, so 1100 μl inhibitor.

2. To remove cell/debris, spin samples at 7500 xg 20min, 4 °C in the Ultracentrifuge.

3. Collect the supernatant, pool and measure volume & record.

4. Add an equal volume of 2X PBS, 6mM MgCl₂ to the plasma.

5. Label low-retention eppendorf tubes 1-96.

6. Transfer 400μl of each sample to eppendorf tubes based on appropriate tube map

7. Using an electronic P200, add competitors: 8.6 μl of 40%/μl Salmon sperm DNA; 8.6 μl of 40%/μl tRNA; 8.6μl of 0.5X S1.

8. Incubate at RT with end over end rotation for 10 min.

9. Add 10μL of appropriate oligo library, mix well. Save any leftover diluted library for gel control (see below).

10. Incubate 1 hr at RT with end over end rotation.

11. Using an electronic repeat P100, add 187μl of 20% PEG 8000 to sample for a final 6% concentration to the 435.5 μl of sample/oligo library. Invert a few times to mix and incubate for 15 min at 4 °C

12. Spin each sample in table top centrifuge at 10,000 xg for 5 min.

13. Remove supernatant and discard, add 1ml 1x PBS, 3mM MgCl₂ to pellet.

14. Wash pellet by gentle inversion
1.5. Remove buffer, re-suspend pellets in 100µl IX PBS, 3mM MgCl₂; incubate at RT for 10 min on mixmate @ 900rpm to re-suspend. Make sure each sample is well re-suspended by pipetting.

16. Pool all desthiobiotin library samples into one 50ml falcon tube, and the unlabeled library into another, total volume for each should be 4800µl.

17. Take 10µL aliquot for the input into AP sample for gel (add 10 µL of 2X LDS buffer w/ 2X reducing agent.

18. Affinity Purification:

19. Prepare 10µL of MyOne Strep-coated Magnetic beads per each condition into a 1.5 ml eppendorf tube and place on a magnetic bead rack. Have a Bead only control as well (n=3)

20. Remove supernatant and wash IX 500µl with Bead buffer.

21. Resuspend beads in an equal volume of IX PBS, 3mM MgCl₂ (equal vol to what was taken out originally = 10µl)

22. Add the 10µl of beads directly to the 4780µl from step 19. To Bead only control add PBS.

23. Incubate samples with streptavidin beads 1hr RT on plate shaker (taped).

24. Place on the large magnetic stand for 1 min and remove supernatant

25. Add 1.5 mL of IX lysis buffer to the samples (do 3 X 500µl with a good rinse of the 50mL falcon tube for each to collect all the beads) and transfer to a new set of eppendorf tubes.

26. Incubate for 20 min on ice.

27. Place tubes in magnetic bead rack, let equilibrate 1 min and remove the supernatant.

28. Wash the beads with wash buffer #1 via vortexing. Resuspend well.

29. Place tubes on magnetic bead rack, let equilibrate 1 min and remove the supernatant

30. Wash 2 additional times as with wash buffer #1 steps 27-29 (total 3 washes with wash buffer #1)

31. Repeat steps 27-29 (2) additional times with wash buffer #2

32. During the last wash transfer beads to a new eppendorf tube. (to reduce non-specific binding)

33. Do one dry spin to make sure all residual wash buffer is removed.

34. Add 10µl of Biotin Elution buffer 1 to beads

35. Incubate for 15 minutes at 37°C.

36. Place on magnetic stand for 1 min, collect sup and transfer to a new tube, add 10µL of 2X LDS, 2X Reducing agent to eluted sample. Save as Elution #1.

37. Add 10µl of IX LDS Sample Buffer, IX Reducing buffer to magnetic beads.
38. Boil the samples for 15 min at 90 °C. The boiling time is 15 minutes to ensure the streptavidin on the beads unfolds and releases the biotinylated aptamer-protein complex.

39. Place samples on magnetic stand on ice and collect the eluted sample. This is Elution #2. Discard the beads.

40. Gel 1 layout:
Lane 1: 5ng Desthiobiotin library
Lane 2: IX LDS
Lane 3: Marker
Lane 4: Desthiobiotin Elution #1
Lane 5: Unlabeled Elution #1
Lane 6: Bead only Elution #1
Lane 7: Desthiobiotin Elution #2
Lane 8: Unlabeled Elution #2
Lane 9: Bead only Elution #2
Lane 10: Input for AP (saved from step 17)

41. Running Reducing SDS gel:
Prepare IX MOPS SDS Running Buffer from 20X MOPS SDS Buffer
Use 10 or 12 well 4 -12 % Bis Tris gel
Peel off tape seal and place in the gel box. Insert spacer for second gel cassette if needed

42. Fill the inside/upper chamber with running buffer MOPS (IX) and 500ul Antioxidant
Remove the comb carefully, not disturbing the wells
Rinse the wells with the running buffer to remove the storage buffer which can interfere with sample running

43. Slowly load samples to each well carefully using L-20 tip
Fill the outer/lower chamber with approximately 600ml of running buffer MOPS (IX)
Place top portion of unit and secure correct electrodes
Run the gel to migrate proteins
100 V constant for samples to move through stack (until all samples line up) for 15 min
Increase to 150 V constant for running (until visible sample buffer comes to bottom) for ~ 1 hr
At the end of the run, stop the power supply and remove the gel cassettes from cell
Disassemble the gel cassette by with gel knife.
Remove one side of cassette case. Trim off the gel foot and wells (avoid drying gel).
Transfer gel into container filled with Mili Q water and perform a quick wash.
Silver staining:
Materials:
ProteoSilver TM Silver Stain Kit, Sigma Catalog No. PROT-SIL1, Lot No. SLBJ0252V
Fixing solution. Add 50 ml of ethanol and 10 ml of acetic acid to 40 ml of ultrapure water.

30% Ethanol solution. Add 30 ml of ethanol to 70 ml of ultrapure water.

Sensitizer solution. Add 1 ml of ProteoSilver Sensitizer to 99 ml of ultrapure water. The prepared solution should be used within 2 hours. A precipitate may form in the ProteoSilver Sensitizer. This precipitate will not affect the performance of the solution. Simply allow the precipitate to settle and remove 1 ml of the supernatant.

Silver solution. Add 1 ml of ProteoSilver Silver Solution to 99 ml of ultrapure water. The prepared solution should be used within 2 hours.

Developer solution. Add 5 ml ProteoSilver Developer 1 and 0.1 ml ProteoSilver Developer 2 to 95 ml of ultrapure water. The developer solution should be prepared immediately (<20 minutes) before use.

All steps should be carried out in the hood and waste needs to be collected in toxic designated container.

Procedure

A. Direct Silver Staining

All steps are carried out at room temperature on an orbital shaker at 60 to 70 rpm.

Fixing - After electrophoresis of the proteins in the mini polyacrylamide gel, place the gel into a clean tray with 100 ml of the Fixing solution overnight in the hood. Cover tightly.

Ethanol wash - Decant the Fixing solution and wash the gel for 10 minutes with 100 ml of the 30% Ethanol solution.

Water wash - Decant the 30% Ethanol solution and wash the gel for 10 minutes with 200 ml of ultrapure water.

Sensitization - Decant the water and incubate the gel for 10 minutes with 100 ml of the Sensitizer solution.

Water wash - Decant the Sensitizer solution and wash the gel twice, each time for 10 minutes with 200 ml of ultrapure water.

Silver equilibration - Decant the water and equilibrate the gel for 10 minutes with 100 ml of the Silver solution.

Water wash - Decant the Silver solution and wash the gel for 1 to 1.5 minutes with 200 ml of ultrapure water.

Gel development - Decant the water and develop the gel with 100 ml of the Developer solution. Development times of 3 to 7 minutes are sufficient to produce the desired staining.
intensity for most gels. Development times as long as 10 to 12 minutes may be required to detect bands or spots with very low protein concentrations (0.1 ng/mm2).

[001093] 10. Stop - Add 5 ml of the ProteoSilver Stop Solution to the developer solution to stop the developing reaction and incubate for 5 minutes. Bubbles of CO₂ gas will form in the mixture.

[001094] 11. Storage - Decant the Developer/Stop solution and wash the gel for 15 minutes with 200 ml of ultrapure water. Store the gel in fresh, ultrapure water and take picture for documentation.

[001095] Protein identification

[001096] Protein bands of interest were excised from the gradient gels and subjected to liquid chromatography-tandem mass spectrometry (LC-MS/MS) as above.

**Example 26: Oligonucleotide probes: breast cancer versus non-cancer**

[001097] This Example presents breast cancer oligonucleotide probes identified in a library enriched against balanced fractions pool of plasma-derived microvesicles from patients with 50% aggressive cancer. General methodology is as presented in Example 21 above. The samples comprised pools of 30 each of breast cancer patient plasma and healthy plasma (i.e., non-cancer controls). A set of cancer specific aptamers was identified where each aptamer has a fold change exceeding two when compared to either the healthy plasma pool (normal and non-cancer) or process control (no negative selection enriched library).

[001098] Methodology

[001099] The F-TRIn-35n-B 8-3s library as described herein was enriched against microvesicles from the plasma samples. See Example 21 with the modifications noted below. The screened library comprised a 5' region (5' CTAGCATGACTGCAGTACGT3' (SEQ ID NO. 131)) followed by the random naive aptamer sequences and a 3' region (5' CTGTCTTTATAACATCTGACGCTGCGACGASEQ ID NO. 132)). In the previous enrichment protocol in Example 21, positive, negative and positive selections were performed before each cycle of PCR to re-amplify the library. But in the current enrichment protocol, the aptamer library was purified with streptavidin beads after each selection and the beads were directly used for PCR amplification. Also as compared to prior experiments, the new sample pool for aptamer enrichment was balanced for different plasma fractions and collection vial for each of cancer, non-cancer and normal patients. Specifically, plasma was initially collected in 4 tubes from each patient, each of those 4 tubes were split into 3 aliquots, resulting in 12 aliquots from each patient. For example, a 1st tube out of the 4 results in aliquots 1A, 1B and 1C, the same split is repeated for the other three aliquots, which results in 12 aliquots from each patient (1A-C, 2A-C, 3A-C, 4A-C). Correspondingly, a pool of 60 patients consists of each variant repeated 5 times (5x12). Enrichment was done according to the selection methodology outlined above, where the library was PCR amplified after each binding round for 7 rounds, which include 3 positive selections against cancer-derived samples, 3 negative selections against controls, and a final positive selection. The enriched library was subjected to the probing test on cancer and healthy pools of plasma samples. There was a subset of 296 cancer specific aptamers, which have a relatively higher read count and fold change > 2 as compared to healthy pool and process control. The detailed protocol is as follows:
[001100] Equipment & Supplies

[001101] IxPBS (HiClone): SH30256.0 1, Lot #: AZC186921, bottle# 1476, exp. 05/20 16.

Supplemented with 3mM MgCl₂ (4227844, USB) in steps 6, 7, 11.

[001102] Table Top centrifuge: 0363

[001103] 20X SI: Aptamer Science TT 0702 14, LN: 14F-01-SI, exp. 2015-06

[001104] PEG8000- lot #SLBJ9928V cat # 91458, protease inhibitor Ref - 05892791, Water - ref 10977-0 15, lot # 1606173

[001105] 2X PBS+ 6mM MgCl₂ - PBS (Sigma) -SLBK2636V, Water - RNBD29 18, MgCl₂ - 4227844 (USB). Used in steps: 5, 9.

[001106] Stock yeast tRNA (Ambion) - lot #1406019. Salmon DNA (Invitrogen) - lot # 1617974

[001107] Starting solution comprises 5 ng Non-Enriched F-Trin-35n-B aptamer library, 300 ul of plasma, 0.01x SI + 0.8ng/ul Salmon DNA/tRNA (competitor DNAs), 6% PEG8000 (to precipitate microvesicles); final volume 600 ul.

[001108] Round 1 (1st positive enrichment)

[001109] Step 1: Pre-chill tabletop centrifuge at 4°C.


[001111] Step 3: Add 50 ul of protease inhibitor to the sample (on top of frozen plasma) and start thawing: 1 ml total ea.

[001112] Step 4: Cell spin: To remove cells/debris, spin samples at 10,000 x g, 20 min, 4°C. Collect the entire volume of supernatant (SN) without disturbing the pellet.

[001113] Step 5: Mix SN from step 4 with equal volume of 2xPBS 6 mM MgCl₂, collect 600 ul ea into 2ml Fisher Low binding tubes for use in step 6. Store remaining sample at 4°C for the following rounds.

[001114] Step 6: Blocking: Add competitors in order: 1) Salmon DNA Stock; 2) tRNA; 3) SI. Make dilutions to desired concentration (see starting solution above) in IxPBS, 3mM MgCl2, mix well. Incubate for 10 min at room temperature (RT), end-over-end rotation.

[001115] Step 7: Binding: Add ssDNA Probing library to the final concentration 12.5 pg/ul for binding. Make dilutions in IxPBS, 3mM MgCl₂.

[001116] Step 8: Precipitation: Add buffer (20% PEG8000 in 1x PBS with 3mM MgCl₂) to sample to the final PEG concentration 6%.

[001117] Step 9: Spin at 10,000 x g for 5 min, RT.

[001118] Step 10: Wash: Remove SN, add 1ml IxPBS, 3 mM MgCl2 and wash pellet by gentle inverion with 1ml aptamer buffer.

[001119] Step 11: Resuspension: Remove buffer, Re-suspend pellets in 200 ul H2O: incubate at RT for 10 min on mixmate 900rpm. Ensure each sample is re-suspended by pipeting after step 11.
Step 12: Purification: Aptamers elution from PEG/Protein pellet with Streptavidin beads (Dynabeads #65001: Dynabeads® MyOne™ Streptavidin C1):

- Beads stock: 10 mg/ml; Capacity: >2,500 pmoles/mg;
- 10 ul beads should bind 250 pmoles Biotin
- 3 ul beads for each aptamer library (AL) sample can bind 75 pmoles Biotin

Library input in the protocol above is ~5ng, which is 0.17 pmol

12.1) Pre-washing Streptavidin Magnetic Beads:

12.1.1 Add 10 uL of Streptavidin Magnetic Beads into 1.5mL microcentrifuge tube (3uL x 2 samples + overage)

12.1.2 Place the tube into a magnetic stand to collect the beads against the side of the tube. Remove and discard the supernatant.

12.1.3 Add 0.5mL of Wash Buffer (IxPBS with 0.1% Tween 20) to the tube. Invert the tube several times or vortex gently to mix. Collect the beads with a magnetic stand, then remove and discard the supernatant.

12.1.4 Wash once with 0.5ml IxPBS, collect the beads with a magnetic stand, then remove and discard the supernatant.

12.1.5 Add 35 ul IxPBS to tubes. Aliquot into 10ul per well (using repeater pipet) in 1.5 ml Fisher low-binding tubes. Add 40 ul of IxPBS.

12.2 Incubate 100 ul of sample, recovered in step 12, at 50°C for 10 min (mixmate, 500 rpm) (to denature/remove protein/PEG from aptamer library)

12.3 Sample Binding:

12.3.6 Add heat denatured samples to beads (aliquoted in step 5).

12.3.7 Incubate for 30min, 37°C, mixmate at 800 rpm. Spin down at 2000rpm for 20 sec.

12.3.8 Collect the beads (with bound aptamers) using magnet and remove the supernatant by multichannel pipet.

12.3.9 Take tubes off the magnet, add 300ul of IxPBS, pipet well. Collect the beads with magnetic stand, then remove supernatant.

12.3.10 Add 100ul H2O to resuspend beads by multichannel pipet.

12.3.11 Block the beads with Biotin: add 7.5 ul from 10uM stock, incubate 15 min, RT, 800rpm.

12.3.12 Collect the beads (with bound aptamers) using magnet and remove the supernatant by multichannel pipet. Take tubes off the magnet, add 200ul of IxPBS, pipet well. Collect the beads with magnetic stand, then remove supernatant.

12.3.13 Add 100ul H2O

Step 13: Use beads with bound aptamer library directly in re-AMP PCR (3x33ul)

Step 14: Agarose gel
SYBR Gold gel lot #: H204044-0 1

2ul of sample + 8ul of loading buffer. Run for 10 cycles

If no dsDNA bands appear, run additional 3 cycles

Optional: if PCR product has non-specific bands, perform gel cut.

Step 15: dsDNA purification with Nucleospin column (NTI binding buffer):

- combine 3 wells per column

- 5 min elution in 30 ul NE buffer, RT. Add 20 ul of NE after elution.

Step 16: Optional: Gel, SybrGold, Agarose: to verify product with correct size is observed

Quantify dsDNA (QuBit, Life Technologies) (keep 5 ul of dsDNA as control for gel later)

Lambda digestion at 37C for 2 h; heat inactivation at 80C for 10 min.

ssDNA purification with Nucleospin column (NTC binding buffer).

- combine 3 samples per column

- elution: 5 min in 30 ul NE buffer, RT

Quantify dsDNA (QuBit, Life Technologies) ssDNA (5 ul)

Gel ssDNA: load between 2ng and 10 ng

Round 2 (2d positive enrichment)

Repeat protocol starting from step 6 above, using diluted plasma samples stored in step 5.

All steps are the same except for step 7. Input of library is 5 times less

Step 7: Binding: Add ssDNA Probing library to the final concentration 2.5 pg/ul for binding. Make dilutions in 1xPBS, 3mM MgCl2.

Steps 6-13: As above. Use entire 100 ul of re-suspended sample to re-amplify via PCR

Steps 14-20: ssDNA preparation as above.

Round 3 (3rd positive enrichment)

Repeat steps as shown for round 2

Round 4 (Negative enrichment)

Make sure to use correct negative samples (e.g., non-cancer sample in the case of positive cancer selection above, or vice versa)

Steps 1-9: Repeat as in Round 2 above (ssDNA input is 2.5pg/ul).

Collect SN (~900ul) and discard pellet after PEG precipitation.

Resuspenstion: Not needed in this step

Purification: Aptamers elution from PEG/Protein pellet with Streptavidin beads (Dynabeads #6500 1: Dynabeads® MyOne™ Streptavidin Cl):

Beads stock: 10 mg/ml; Capacity: >2,500 pmoles/mg;

10 ul beads should bind 250 pmoles Biotin
3 uL beads for each AL sample can bind 75 pmoles Biotin.

Library input in the protocol above is ~5ng, which is 0.17 pmol.

CHANGE FOR SN -> For each sample, split 900uL SN collected in step 10 above into two aliquots 450uL ea, consider 3uL of beads per each aliquot.

12.1) Pre-washing Streptavidin Magnetic Beads:
12.1.1 Add 30 uL of Streptavidin Magnetic Beads into 1.5mL microcentrifuge tube. Four samples make 8 aliquots (for beads treatment) x 3uL beads + overage.
12.1.2 Place the tube into a magnetic stand to collect the beads against the side of the tube. Remove and discard the supernatant.
12.1.3 Add 0.5mL of Wash Buffer (1xPBS with 0.1% Tween 20) to the tube. Invert the tube several times or vortex gently to mix. Collect the beads with a magnetic stand, then remove and discard the supernatant.
12.1.4 Wash 1 time with 0.5ml 1xPBS, collect the beads with a magnetic stand, then remove and discard the supernatant.
12.1.5 Add 105 ul 1xPBS to tubes. Aliquot into 10ul per well (using repeater pipet) in 1.5 ml Fisher low-binding tubes. Add 40 ul of 1xPBS.
12.2) Incubate 900 ul of sample, recovered in step 10, at 50°C for 10 min (mixmate, 500 rpm) (to denature/remove protein/PEG from aptamer library)
12.3) Sample Binding:
12.3.6 Add 1/2 (450ul) heat denatured samples to beads ( aliquoted in step 5).
12.3.7 Incubate for 30min, 37°C, end-over-end rotation. Spin down at 2000rpm for 20 sec.
12.3.8 Collect the beads (with bound aptamers) using magnet and remove the supernatant by multichannel pipet.
12.3.9 Take tubes off the magnet, add 300ul of 1xPBS, pipet well. Collect the beads with magnetic stand, then remove supernatant.
12.3.10 Add 100ul H2O to resuspend beads by multichannel pipet.
12.3.11 Block the beads with Biotin: add 7.5 ul from 10uM stock, incubate 15 min, RT, 800rpm.
12.3.12 Collect the beads (with bound aptamers) using magnet and remove the supernatant by multichannel pipet. Take tubes off the magnet, add 200ul of 1xPBS, pipet well. Collect the beads with magnetic stand, then remove supernatant.
12.3.13 Add 100ul H2O per purification sample (total 200ul per negative enrichment sample)
12.3.14 Split each sample into 3 aliquots (33ul each), this results into 6 wells per enrichment sample.

Step 13: Perform PCR

Rounds 5, 6: same as round 4
Round 7: repeats steps as shown for round 2

Results

FIGs. 14A-B are scatter plots showing high correlation of frequencies between replicates of cancer pool (FIG. 14A) and appearance of cancer specific subset of aptamers when cancer and healthy pools are compared (circled region in FIG. 14B). FIGs. 14C-F are scatter plots showing the appearance of cancer specific subset of aptamers when the cancer pool profile of the enriched library (i.e., C-RN7) was compared to the cancer pool of the process control library (CP; FIG. 14C) and healthy pool profile of process control library (HP; FIG. 14D). We did not observe a similar subset when the healthy pool profile (C-RN7 v HP) of the enriched library was compared to both variants of profile of process control (FIGs. 14E-F, see points along x-axis).

Selections of aptamer sequences identified in the enrichment screening are shown in Tables 27-29. In these tables, each complete aptamer sequence is assembled from 5’ to 3’ as a 5’ region, the variable region, and 3’ region. The sequences are shown 5’ to 3’ from left to right, wherein each complete sequence consists of a 5’ leader sequence 5’- CTAGCATGACTGCAGTACGT (SEQ ID NO. 131) followed by the indicated Variable Region sequence followed by the 3’ tail sequence 5’- CTGTCTTTTATAACATCT GCAGCTGCCGACGA (SEQ ID NO. 132). The top 50 cancer specific sequences are shown in Table 27, each having a relatively high copy number in these experiments and a cancer/normal copy number ratio of at least 2-fold higher in the cancers. Table 28 comprises 25 sequences with mid-level copy number and a cancer/normal copy number difference of 0.8-1.2-fold. Table 29 shows a selection of 25 more sequences with low copy numbers and cancer/normal copy number ratio of -1.0 (i.e., no significant difference).

### Table 27: Cancer-Specific Sequences

<table>
<thead>
<tr>
<th>SEQ ID NO.</th>
<th>Variable Region</th>
</tr>
</thead>
<tbody>
<tr>
<td>137</td>
<td>AACCAGGACTGTATTGGGTCATGCCCCATGCGACTTTT</td>
</tr>
<tr>
<td>138</td>
<td>AACCTGAGCCCAATTAACCTCCACTGTCATATCT</td>
</tr>
<tr>
<td>139</td>
<td>AGATGGCCGCTGTATCCTACTGTCCTGCCATTATAGA</td>
</tr>
<tr>
<td>140</td>
<td>AGACATGGCCACGCGCCTTTACATCTCCAGCTGGCA</td>
</tr>
<tr>
<td>141</td>
<td>AGCCACCGAGACCTACTTCATTCTACTGTCGCAAATGTT</td>
</tr>
<tr>
<td>142</td>
<td>AGGTACGTAAGGCTGCCAGTACTCCTACTGTCGCGAAG</td>
</tr>
<tr>
<td>143</td>
<td>AGTCTAGGATCGACTCTTTTCTACTAGCTGCTCTTC</td>
</tr>
<tr>
<td>144</td>
<td>AGTGTCGCACTGTAAATCCATCTCCACTGTGCTCTGA</td>
</tr>
<tr>
<td>145</td>
<td>AGTTGAGAGCCCTTGGTATTCTACTGCGCTCTATAGA</td>
</tr>
<tr>
<td>146</td>
<td>CAGTTGTTAAGCTATTGTGGAACCTCGGCTG</td>
</tr>
<tr>
<td>147</td>
<td>CATGCCACAGCTCCATTTAACACATGGCATATTATAGA</td>
</tr>
<tr>
<td>148</td>
<td>CATGGACAGCCATGCTCCTACTACTGCTGCAATCATA</td>
</tr>
<tr>
<td>149</td>
<td>CTCTGATAGCTCTCTACTACTGTTGCAATTTATG</td>
</tr>
<tr>
<td>150</td>
<td>CTTTGAAAATGCGCACAACCCTCCACTACCTGGTG</td>
</tr>
<tr>
<td>151</td>
<td>CGGAGGACGAGGTACTGCGCTGGTCTAAACAAATGAA</td>
</tr>
<tr>
<td>152</td>
<td>CGTGAGCTATGTTTCCTACTGTCGCAATACAGATA</td>
</tr>
<tr>
<td>153</td>
<td>TAGAGGGCTATTCCCTCTACTGCTGCTGTTAT</td>
</tr>
<tr>
<td>154</td>
<td>GCAAAACAAATTAGTGACTCTACAGCTGGCCCAATGAA</td>
</tr>
<tr>
<td>155</td>
<td>GCAACTCTCTCTCTACGCGCCTGAGATTCTCTCCATAGA</td>
</tr>
<tr>
<td>156</td>
<td>GAGCCACATGTGAACTCTACGCTGCTGACTCATA</td>
</tr>
<tr>
<td>157</td>
<td>GAGGAGACAAAGATTGCTCCCCTACTAGCTGGCCCTTA</td>
</tr>
<tr>
<td>158</td>
<td>GAGTTTGGGCTGGATACCTACTGCTGCTGACCTTA</td>
</tr>
</tbody>
</table>
Table 28: Moderate Control Sequences

<table>
<thead>
<tr>
<th>SEQ ID NO.</th>
<th>Variable Region</th>
</tr>
</thead>
<tbody>
<tr>
<td>187</td>
<td>ACATGCATACCCCTACGGAATCCTCCATCACCATAC</td>
</tr>
<tr>
<td>188</td>
<td>AGTCAGTGCGCCGCTATTTACGATCTCACTGTTC</td>
</tr>
<tr>
<td>189</td>
<td>AGTGCAGTCGCGATGGGAACTTCTTCTTTGCTTTA</td>
</tr>
<tr>
<td>190</td>
<td>ATGGCAGTCCAGGCTAATCTAATCTGCTGTGCT</td>
</tr>
<tr>
<td>191</td>
<td>ATTGCGATCCGCAGATTTTTTACTCTACTGTGCT</td>
</tr>
<tr>
<td>192</td>
<td>TATGCGGATTTCCTACGGCTAATCGTTTTTT</td>
</tr>
<tr>
<td>193</td>
<td>GCAGGACACCCCCCTACGGAATCCTCCATCACCATAC</td>
</tr>
<tr>
<td>194</td>
<td>ACGTACCCTACGGAATCCTCCATCACCATAC</td>
</tr>
<tr>
<td>195</td>
<td>TTGCGATCCGCAGATTTTTTACTCTACTGTGCT</td>
</tr>
<tr>
<td>196</td>
<td>CATGCCCGTAAACGCCTAATCTACCACCTTTTGCT</td>
</tr>
<tr>
<td>197</td>
<td>CATGCCGCCTTTCGACTTCCCATATCCCAACACGCC</td>
</tr>
<tr>
<td>198</td>
<td>CATGCCGGTCATTCCATACTCAAACTCATTCACTC</td>
</tr>
<tr>
<td>199</td>
<td>CATGCGAGCTACCCTTATCATCTCATGTTTGCTTT</td>
</tr>
<tr>
<td>200</td>
<td>CATGCGATTGCGCCCCATTCCTTGCCTTATCTCAC</td>
</tr>
<tr>
<td>201</td>
<td>CATGCGCACATCGTACTCCGTAGCCCTAAATTATCA</td>
</tr>
<tr>
<td>202</td>
<td>CATGCGGAGCTACCTACGGAATCCTCCATCACCATAC</td>
</tr>
<tr>
<td>203</td>
<td>GCCATGCGAGGTTTATGCGTCTAATCCTCTTCTT</td>
</tr>
<tr>
<td>204</td>
<td>GCCATACAGGCGAGTACCTACGGAATCCTCCATCACCATAC</td>
</tr>
<tr>
<td>205</td>
<td>GCGCCACGACTACCCGACTTCCGATTACTTTCTT</td>
</tr>
<tr>
<td>206</td>
<td>GCGGACATCGGGCAGTTTCTATACACCCACCTTTTG</td>
</tr>
<tr>
<td>207</td>
<td>GCTGCAAGGCCTAAAATTCTGCCTCCACACATA</td>
</tr>
<tr>
<td>208</td>
<td>GCTGCGAGGCGGCTACCTACGGAATCCTCCATCACCATAC</td>
</tr>
<tr>
<td>209</td>
<td>TATGCGGACACCCACATATTACACGACCTTATCCT</td>
</tr>
</tbody>
</table>
Table 29: Negative Control Sequences

<table>
<thead>
<tr>
<th>SEQ ID NO.</th>
<th>Variable Region</th>
</tr>
</thead>
<tbody>
<tr>
<td>212</td>
<td>AAAACACGCAACTACCTGGAAAACCTTGACCTCTTT</td>
</tr>
<tr>
<td>213</td>
<td>AAAACACGCAACTACCTGGAAAACCTTGACCTCTTT</td>
</tr>
<tr>
<td>214</td>
<td>AAAACACGCAACTACCTGGAAAACCTTGACCTCTTT</td>
</tr>
<tr>
<td>215</td>
<td>AAAACACGCAACTACCTGGAAAACCTTGACCTCTTT</td>
</tr>
<tr>
<td>216</td>
<td>AAAACACGCAACTACCTGGAAAACCTTGACCTCTTT</td>
</tr>
<tr>
<td>217</td>
<td>AAAACACGCAACTACCTGGAAAACCTTGACCTCTTT</td>
</tr>
<tr>
<td>218</td>
<td>AAAACACGCAACTACCTGGAAAACCTTGACCTCTTT</td>
</tr>
<tr>
<td>219</td>
<td>AAAACACGCAACTACCTGGAAAACCTTGACCTCTTT</td>
</tr>
<tr>
<td>220</td>
<td>AAAACACGCAACTACCTGGAAAACCTTGACCTCTTT</td>
</tr>
<tr>
<td>221</td>
<td>AAAACACGCAACTACCTGGAAAACCTTGACCTCTTT</td>
</tr>
<tr>
<td>222</td>
<td>AAAACACGCAACTACCTGGAAAACCTTGACCTCTTT</td>
</tr>
<tr>
<td>223</td>
<td>AAAACACGCAACTACCTGGAAAACCTTGACCTCTTT</td>
</tr>
<tr>
<td>224</td>
<td>AAAACACGCAACTACCTGGAAAACCTTGACCTCTTT</td>
</tr>
<tr>
<td>225</td>
<td>AAAACACGCAACTACCTGGAAAACCTTGACCTCTTT</td>
</tr>
<tr>
<td>226</td>
<td>AAAACACGCAACTACCTGGAAAACCTTGACCTCTTT</td>
</tr>
<tr>
<td>227</td>
<td>AAAACACGCAACTACCTGGAAAACCTTGACCTCTTT</td>
</tr>
<tr>
<td>228</td>
<td>AAAACACGCAACTACCTGGAAAACCTTGACCTCTTT</td>
</tr>
<tr>
<td>229</td>
<td>AAAACACGCAACTACCTGGAAAACCTTGACCTCTTT</td>
</tr>
<tr>
<td>230</td>
<td>AAAACACGCAACTACCTGGAAAACCTTGACCTCTTT</td>
</tr>
<tr>
<td>231</td>
<td>AAAACACGCAACTACCTGGAAAACCTTGACCTCTTT</td>
</tr>
<tr>
<td>232</td>
<td>AAAACACGCAACTACCTGGAAAACCTTGACCTCTTT</td>
</tr>
<tr>
<td>233</td>
<td>AAAACACGCAACTACCTGGAAAACCTTGACCTCTTT</td>
</tr>
<tr>
<td>234</td>
<td>AAAACACGCAACTACCTGGAAAACCTTGACCTCTTT</td>
</tr>
<tr>
<td>235</td>
<td>AAAACACGCAACTACCTGGAAAACCTTGACCTCTTT</td>
</tr>
<tr>
<td>236</td>
<td>AAAACACGCAACTACCTGGAAAACCTTGACCTCTTT</td>
</tr>
</tbody>
</table>

The data described above (i.e., FIGs. 15A-F and Tables 27-29) were obtained with 1 ng input of the specified oligonucleotide libraries used to probe the various samples. In addition, probing experiments were performed with lower titers of the C-R7N oligonucleotide library, specifically 0.1, 0.01, 0.001 and 0.0001 ng. A total of 826 unique oligonucleotide sequences were detected between all titers. An eighth enrichment round (round 8) was performed with four titers of the library C-R7N-1 (i.e., 1 ng input in round 7): 1, 0.1, 0.01 and 0.001 ng input. From each titer, sets of oligonucleotides that passed the same filters described above were obtained. The composite set of oligonucleotides was compared to the sets of oligonucleotides obtained from the titrations of the C-R7N library noted above. The variable regions of 733 oligonucleotides observed in at least two of the five sets of oligonucleotides are listed in rank order by occurrence in SEQ ID NOs. 237-969. These oligonucleotides were identified as oligonucleotide probes that selectively bind to cancer samples. As in Table 27, the oligonucleotides were synthesized with a 5' region consisting of the sequence (5'-CTAGCATGACGTACATG) (SEQ ID NO. 131) and a 3' region consisting of the sequence (5'-CTGTCTTTCATACATGACGTACATG) (SEQ ID NO. 132) flanking the variable regions. The 733 cancer specific sequences which can be used to identify cancer samples along with those in Table 27 above.
As noted, next generation sequencing technologies can be used to identify and quantify members of the oligonucleotide library that bind to a sample. We also designed probes that can be used to identify and quantify members of the oligonucleotide library that bind to a sample without requiring a PCR step. This procedure relies on hybridization of the bound oligonucleotide with oligonucleotide-specific complementary probe. The complementary probe can directly or indirectly carry a tag that can be detected. For example, the oligonucleotide-specific complementary probe may carry a fluorescent tag. A general design is shown in FIG. 14G. In the figure, an individual member of the oligonucleotide library 1401 is bound by complementary probe 1402. Complementary probe 1402 consists of three sections from 5’ to 3’: 1) probe part B 1403; 2) probe complement 1404; and 3) probe part A 1405. The member of the oligonucleotide library 1401 is bound by complementary probe 1402 via base pair hybridization between the variable region of oligonucleotide library 1401 and the probe complement region 1404 of complementary probe 1402. Tag probe 1406 which is fluorescently labeled (indicated by the circles in the figure) and capture probe 1407 hybridize to the probe part A 1405 and probe part B 1403 of complementary probe 1402, respectively. These features allow for the capture of the entire complex (i.e., oligonucleotide library 1401, complementary probe 1402, tag probe 1406, and capture probe 1407) via the capture probe, combined with detection via the label of the tag probe 1406.

For these experiments, the complementary probes were designed to work with tag probe and capture probe supplied as part of the nCounter nucleic acid detection system from NanoString Technologies, Inc. (Seattle, WA). The complementary probes for nCounter quantification were designed as reverse complement to the top 50 oligonucleotide probe binders and 50 negative controls, described above. The complete sequences are shown in Table 30. In the table, the ID column is an identifier and the Type column indicates if the sequence is to a control oligonucleotide probe ("neutral") or to a cancer binding oligonucleotide probe ("binder"). Each complete sequence comprises from 5' to 3': 1) 5' Region (Probe Part B 1403); 2) Probe Complement Region 1404; and 3) 3' Region (Probe Part A 1405), where region identification numbers correspond to FIG. 14G. The probe complement region may be shorter than the variable sequence of the corresponding oligonucleotide probe in order to optimize binding parameters (e.g., specificity and melting point). The probe complement region may also be so optimized to bind against the variable sequence of the corresponding oligonucleotide probe and also to at least a portion of its 5’ and 3’ flanking sequences (e.g., as shown in Tables 27-29 above). The oligonucleotides each have a 5’ region 1403 consisting of the sequence (5’- CGAAGCCATGACCTGATCGACT (SEQ ID NO. 970)) 5’ of the Probe Complement Region 1404 and a 3’ region 1405 as shown in Table 30 located 3’ of the Probe Complement Region 1404. Probe Part A 1405 shown in FIG. 14G corresponds to SEQ ID NO. 971.

<table>
<thead>
<tr>
<th>SEQ ID NO</th>
<th>Type</th>
<th>Probe Complement Region 1504</th>
<th>3’ Region (Probe Part A 1505)</th>
</tr>
</thead>
<tbody>
<tr>
<td>972</td>
<td>neutral</td>
<td>CAAGTTTCAGATTGGTCTGTTTCTA</td>
<td>CTTTAGCCCGGACACGCTG</td>
</tr>
<tr>
<td>973</td>
<td>neutral</td>
<td>CAGGGTAGGTATCGTGAGTTTTTATTACT</td>
<td>CATCCGTCTTTTTGTGTTT</td>
</tr>
<tr>
<td>974</td>
<td>neutral</td>
<td>GAAAGTCAGGAAAATGCCTGCTTTT</td>
<td>CACAAATTGTGCGGTTAGCAGAA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----</td>
<td>----</td>
<td>----</td>
<td></td>
</tr>
<tr>
<td>975</td>
<td>neutral</td>
<td>GAGACAGAGTGGGCAATGCGATAATG</td>
<td>GTTTGAGGAGGAC</td>
</tr>
<tr>
<td>976</td>
<td>neutral</td>
<td>GAGACAGAGGAAAAAGGGATAGTTGGTCTCGT</td>
<td>CAAAGAGGCTATCTTCTTCGTTT</td>
</tr>
<tr>
<td>977</td>
<td>neutral</td>
<td>GAGCCTGTTGGGGTCTACGTTAC</td>
<td>CAACTCTACACCTCAGCTTAC</td>
</tr>
<tr>
<td>978</td>
<td>neutral</td>
<td>GCGGAGAACAGTTTACGTAATC</td>
<td>CATATGCGGAGCTTTAGT</td>
</tr>
<tr>
<td>979</td>
<td>binder</td>
<td>GAGACAGAAGATCAGCAGTATGGCAGTA</td>
<td>CTTTGTGATGTATCTATCATTTA</td>
</tr>
<tr>
<td>980</td>
<td>binder</td>
<td>AGTGGAGTTTAATGGGCCTTCAGTTCGCT</td>
<td>CAAACCCACTTTTTTTTCAAAT</td>
</tr>
<tr>
<td>981</td>
<td>neutral</td>
<td>GAACAGATGGTGGATGATCGGATCTAATG</td>
<td>CACGTTGGAGAAAAGGAGAGAGAGAC</td>
</tr>
<tr>
<td>982</td>
<td>binder</td>
<td>AGTATATCAAGCCGGGCTGTTATCAGTAC</td>
<td>CTTGGATTTATGGTGCTATGCAAG</td>
</tr>
<tr>
<td>983</td>
<td>binder</td>
<td>TAAAGGGGGGCCATGCCATCTACGTA</td>
<td>CAGATAAGGGTGTTATAGTGAGG</td>
</tr>
<tr>
<td>984</td>
<td>binder</td>
<td>GAGACAGAACATGTCGACGCTATGAGATA</td>
<td>CTTTCTCTCTGTGCTTCAGCTACA</td>
</tr>
<tr>
<td>985</td>
<td>binder</td>
<td>TAATGGCAGCCTTACGTACCTACTACGTA</td>
<td>CAATGCGGATAGAAGTCGAGAG</td>
</tr>
<tr>
<td>986</td>
<td>neutral</td>
<td>CGGCGCACTGACTACGTAC</td>
<td>CGGTTGAGACATTTGCGAGAG</td>
</tr>
<tr>
<td>987</td>
<td>binder</td>
<td>GAGACAGAAGGGGCCACGCTATA</td>
<td>CTTTTGTGGGGGCTTTGAGG</td>
</tr>
<tr>
<td>988</td>
<td>neutral</td>
<td>CCATCGGACTGCACCTACGTA</td>
<td>CAGCGCAGTCAATACATCAAGAG</td>
</tr>
<tr>
<td>989</td>
<td>binder</td>
<td>GAGACAGTCAAGGGCGACAGTTGAGATG</td>
<td>CTTGCGCAATGCACTCAGTCCTGTC</td>
</tr>
<tr>
<td>990</td>
<td>binder</td>
<td>GTGAGAATACCAAGGGTGCTGCAACTACGTA</td>
<td>CAAACTGCGGAGAAGAGAGGAGAGAG</td>
</tr>
<tr>
<td>991</td>
<td>neutral</td>
<td>CCGTCCGATGCGCAGTACGTA</td>
<td>CATTGCTGCTGATTAAGCTGTCAGGAA</td>
</tr>
<tr>
<td>992</td>
<td>neutral</td>
<td>GAGGGGATGACGGCAATACGTA</td>
<td>CTGGACAGTGTTAGGCAATAG</td>
</tr>
<tr>
<td>993</td>
<td>neutral</td>
<td>CAGTGACTGCACTGAGTACGTA</td>
<td>CTTAAGCGGCTGCTTCTCAGGAA</td>
</tr>
<tr>
<td>994</td>
<td>binder</td>
<td>TTGGATGCTCAGACATCGATGCTAGTG</td>
<td>CCACCGGTAGCGGCTTCAGCTACAGAG</td>
</tr>
<tr>
<td>995</td>
<td>neutral</td>
<td>CCTGCTCAGACATCTACGTA</td>
<td>CATTGGGATAGTGTAGCTGACGGA</td>
</tr>
<tr>
<td>996</td>
<td>neutral</td>
<td>CGGACTGCTGACCTACGTA</td>
<td>CACCGAGATTTCCGCTAGTACGAG</td>
</tr>
<tr>
<td>997</td>
<td>neutral</td>
<td>GAAAAATGACGACAGAGCCGGTTACTAGCTACGTA</td>
<td>CTTGGAGTGTAGTTGCTATCAGGTT</td>
</tr>
<tr>
<td>998</td>
<td>neutral</td>
<td>AGGTTGCCAGACTAGTTGACTGCTAGTA</td>
<td>CTTACAGATCTTCGCTAGTACGATTC</td>
</tr>
<tr>
<td>999</td>
<td>neutral</td>
<td>CGCACTGACCTCCTTACTGAGCAGTACT</td>
<td>CGGAGGACTTGTGATAGTGTGACGGA</td>
</tr>
<tr>
<td>1000</td>
<td>neutral</td>
<td>GAGACAGTCAAGATTTTACAGGCGG</td>
<td>CATTGGGAGGTTGATAGTGGTAAA</td>
</tr>
<tr>
<td>1001</td>
<td>neutral</td>
<td>GAGACAGTCAAGATTTTACAGGCGG</td>
<td>CTTTCTGATTATATCCAGAGGAGAGAGAG</td>
</tr>
<tr>
<td>1002</td>
<td>neutral</td>
<td>GGAAAGAATCAGGAGTTTGGTACTGAGTACGTA</td>
<td>CACCCCTCCAAAGCGATTCTATTG</td>
</tr>
</tbody>
</table>

-259-
<table>
<thead>
<tr>
<th></th>
<th>Description</th>
<th>Sequence 1</th>
<th>Sequence 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1003</td>
<td>neutral</td>
<td>GCCGTTTACGGCGTACGACT</td>
<td>CCGGAAACATCTACTGTCGTCATCT</td>
</tr>
<tr>
<td>1004</td>
<td>neutral</td>
<td>GTGAAAGGCGCGTACGACT</td>
<td>CCGGGAATGGCGATTTCCGATTTCT</td>
</tr>
<tr>
<td>1005</td>
<td>neutral</td>
<td>GAGAGCCGAGATGCTACGACT</td>
<td>CCGATCTCTCATACGGACAACTG</td>
</tr>
<tr>
<td>1006</td>
<td>binder</td>
<td>GATGAGAGGATGCTACGACT</td>
<td>CCGTATCGAGCTGCTCACGACGAG</td>
</tr>
<tr>
<td>1007</td>
<td>neutral</td>
<td>ATGATGCTCGTACGACT</td>
<td>CATTGCCAACCATGTGAGAATGAG</td>
</tr>
<tr>
<td>1008</td>
<td>neutral</td>
<td>GGCACCGACTGACGACACT</td>
<td>CACCAGTTAGCGTGCGTATACCA</td>
</tr>
<tr>
<td>1009</td>
<td>neutral</td>
<td>GAGACGTGATAATTATTGACGCTACGACT</td>
<td>CCGGTAATGATAGAAACATATAGC</td>
</tr>
<tr>
<td>1010</td>
<td>neutral</td>
<td>GAGACAGATGATTAGGAGATAGAAGACGACT</td>
<td>CCGGTTGATAATTAGCAGACGCGCT</td>
</tr>
<tr>
<td>1011</td>
<td>binder</td>
<td>GAGACAGTTGATACGCTACGACT</td>
<td>CCGTCTCAGATGGTGGTATAC</td>
</tr>
<tr>
<td>1012</td>
<td>binder</td>
<td>GAGAGATTGAGACGTACGACTGAGACGACT</td>
<td>CTCAGACATTTAGTACGCTACGACG</td>
</tr>
<tr>
<td>1013</td>
<td>neutral</td>
<td>TGGTTGCTCGCTACGACT</td>
<td>CGTGAACATGATTGCTAGAGACG</td>
</tr>
<tr>
<td>1014</td>
<td>binder</td>
<td>GGTTCGCATGAGACGTACGACT</td>
<td>CATTACGAAAATTTGAGCAAGCAATT</td>
</tr>
<tr>
<td>1015</td>
<td>binder</td>
<td>CTCCTGAGCTCGGACGTAC</td>
<td>CATACGACATACGCTAG</td>
</tr>
<tr>
<td>1016</td>
<td>binder</td>
<td>CAGTAGAGAACATTACGTACGACT</td>
<td>CATACGCTACGCTACGACT</td>
</tr>
<tr>
<td>1017</td>
<td>neutral</td>
<td>GAGACGGCAGATGAAAGGGAAGGCACT</td>
<td>CTTTAGTCTTAGGACGCCAACAGC</td>
</tr>
<tr>
<td>1018</td>
<td>neutral</td>
<td>TGCCCGCCTCAGCTAC</td>
<td>CTAGCCAGATCTACGAGATGAG</td>
</tr>
<tr>
<td>1019</td>
<td>neutral</td>
<td>GTGAGGAGGCTCCGCGATAC</td>
<td>CAAATCGCCTATATGGAGGAG</td>
</tr>
<tr>
<td>1020</td>
<td>neutral</td>
<td>ACTCCCTAGCAGCTCAGCT</td>
<td>CCCTGCTTAGGTATCTACATTG</td>
</tr>
<tr>
<td>1021</td>
<td>neutral</td>
<td>GCCATTAAGCTTACGCTACGACT</td>
<td>CTTAGCCTCGGATGCTATACGCT</td>
</tr>
<tr>
<td>1022</td>
<td>neutral</td>
<td>GACAGGAGTACATGCTACGACT</td>
<td>CAGGACGTCTAGTTGACCTCTTTC</td>
</tr>
<tr>
<td>1023</td>
<td>neutral</td>
<td>GAGACAGTGTACAAGAAATGTACGACT</td>
<td>CTGTGGTCCGTCAGGACATAG</td>
</tr>
<tr>
<td>1024</td>
<td>neutral</td>
<td>GAGACAGGAGGACAGAGGAAATTGAGGAAAT</td>
<td>CATATGGGAGAGAGAGACG</td>
</tr>
<tr>
<td>1025</td>
<td>neutral</td>
<td>GGAGGCCCCTACGACT</td>
<td>CGCTCAATTTTGACCATACGAT</td>
</tr>
<tr>
<td>1026</td>
<td>neutral</td>
<td>ATGAGCAGGCCCCTACGACT</td>
<td>CTTATGGCATATGCTACGCT</td>
</tr>
<tr>
<td>1027</td>
<td>binder</td>
<td>CAGAGGAGGAGATGCTACGACT</td>
<td>CCTAAATTGGAAAAAGGTAAAA</td>
</tr>
<tr>
<td>1028</td>
<td>binder</td>
<td>GAGACAGTGTGACGACGACGACT</td>
<td>CTGTGATGCAAGGCTATCTGTC</td>
</tr>
<tr>
<td>1029</td>
<td>binder</td>
<td>GACAGGGTAGAGGTAGTTCCGACGACT</td>
<td>CTTAAGCTATCAGGAAATGCA</td>
</tr>
<tr>
<td>1030</td>
<td>binder</td>
<td>GAGACAGGATGAGAATGATAGACGAGTACGAG</td>
<td>CCCGAATGTTAAATGTGCACGTT</td>
</tr>
<tr>
<td>1031</td>
<td>binder</td>
<td>GAGACAGTGGGGACGCGATAGAG</td>
<td>CCTATGGAAAGCTACCTCCTCC</td>
</tr>
</tbody>
</table>

-260-
<table>
<thead>
<tr>
<th>WO 2017/004243</th>
<th>PCT/US2016/040157</th>
</tr>
</thead>
<tbody>
<tr>
<td>1032 neutral</td>
<td>GGGGTGTCCCTGACGATAC</td>
</tr>
<tr>
<td>1033 binder</td>
<td>GTAGAGTATCCACGCAAATGACGTAC</td>
</tr>
<tr>
<td>1034 binder</td>
<td>GCAGGTTAGGTTGATGCTGACGATAC</td>
</tr>
<tr>
<td>1035 binder</td>
<td>GAGAACAGGCGACGAGTAAATATATAT</td>
</tr>
<tr>
<td>1036 neutral</td>
<td>AAAGTGCCCTAGTACGACGATAC</td>
</tr>
<tr>
<td>1037 binder</td>
<td>CTCGCAGAAGGATACGACGATAC</td>
</tr>
<tr>
<td>1038 neutral</td>
<td>GAGAACAGGCGACGACGATAC</td>
</tr>
<tr>
<td>1039 binder</td>
<td>GAGAACAGGCGACGACGATAC</td>
</tr>
<tr>
<td>1040 binder</td>
<td>GAGAACAGGCGACGACGATAC</td>
</tr>
<tr>
<td>1041 binder</td>
<td>GAGAACAGGCGACGACGATAC</td>
</tr>
<tr>
<td>1042 binder</td>
<td>GAGACAGGAGAAGTACGAGTACGAGTAC</td>
</tr>
<tr>
<td>1043 neutral</td>
<td>GAGAACAGGCGACGACGATAC</td>
</tr>
<tr>
<td>1044 binder</td>
<td>GAGAACAGGCGACGACGATAC</td>
</tr>
<tr>
<td>1045 binder</td>
<td>GAGAACAGGCGACGACGATAC</td>
</tr>
<tr>
<td>1046 binder</td>
<td>GAGAACAGGCGACGACGATAC</td>
</tr>
<tr>
<td>1047 binder</td>
<td>GAGAACAGGCGACGACGATAC</td>
</tr>
<tr>
<td>1048 binder</td>
<td>GAGAACAGGCGACGACGATAC</td>
</tr>
<tr>
<td>1049 neutral</td>
<td>GAGAACAGGCGACGACGATAC</td>
</tr>
<tr>
<td>1050 binder</td>
<td>GAGAACAGGCGACGACGATAC</td>
</tr>
<tr>
<td>1051 neutral</td>
<td>GAGAACAGGCGACGACGATAC</td>
</tr>
<tr>
<td>1052 binder</td>
<td>GAGAACAGGCGACGACGATAC</td>
</tr>
<tr>
<td>1053 binder</td>
<td>GAGAACAGGCGACGACGATAC</td>
</tr>
<tr>
<td>1054 binder</td>
<td>GAGAACAGGCGACGACGATAC</td>
</tr>
<tr>
<td>1055 binder</td>
<td>GAGAACAGGCGACGACGATAC</td>
</tr>
<tr>
<td>1056 binder</td>
<td>GAGAACAGGCGACGACGATAC</td>
</tr>
<tr>
<td>1057 binder</td>
<td>GAGAACAGGCGACGACGATAC</td>
</tr>
<tr>
<td>1058 binder</td>
<td>GAGAACAGGCGACGACGATAC</td>
</tr>
<tr>
<td>1059 neutral</td>
<td>GGGGCCTGACGACGACGATAC</td>
</tr>
</tbody>
</table>
Example 27: Oligonucleotide probes: aggressive breast cancer, non-aggressive breast cancer, and non-cancer

[001203] This Example presents oligonucleotide probes identified in an oligonucleotide probe library enriched against plasma-derived microvesicles from patients with aggressive breast cancer, non-aggressive breast cancer, and non-cancer (breast biopsy-negative samples). The general methodology follows that in Example 26 above. Oligonucleotides were screened using plasma samples from 60 individuals with breast cancer and 60 individuals without breast cancer.

[001204] A combination of ultracentrifugation and PEG based partitioning of microvesicles from each input plasma sample. The F-TRin-35n-B 8-3s library as described herein was enriched against microvesicles from pooled plasma samples (cancer or non-cancer pools) which were isolated using ultracentrifugation. Subsequent enrichments were performed by isolating the microvesicles using PEG precipitation. As in Example 26, the input oligonucleotide library was exposed to positive, negative and again positive targets sequentially and then PCR amplified. We performed two parallel tracks of initial enrichment against pooled samples, one track where pooled cancer samples were considered positive and non-cancer samples were negative, and another track where pooled non-cancer samples were considered positive and cancer samples were negative. The enriched libraries resulting from each track were pooled together before probing individual patient samples in the following experiments.

[001205] Oligonucleotides that distinguish aggressive and non-aggressive breast cancer

[001206] We first identified oligonucleotides that distinguished between aggressive and non-aggressive breast cancer. For these experiments, the aggressive breast cancer patient samples were considered as positive samples (31 samples) and non-aggressive breast cancer patient samples were...
considered as negative samples (29 samples). The enriched libraries from above were subjected to the oligonucleotide probing assay on the individual cancer plasma samples. A two-step normalization procedure was performed to normalize the sum of the copy numbers for all oligonucleotides associated with the PEG-precipitated pellet. The copy number of each recovered oligonucleotide was divided by the total number of sequence reads for each sample, then multiplied by the mean of totals of all samples. This data treatment accounts for the different number of sequence reads for each patient even though the same amount of DNA library was added for each patient-sample. The final value for each oligonucleotide for each patient is the average of normalized copy numbers from three probing replicates. We excluded oligonucleotides with an average read-count across all patients <100, and subsequently selected among the remaining oligonucleotides only those that show a >20% CV across the 60 patient samples. This process resulted in approximately 12000 remaining oligonucleotides.

We compared patient profiles using clustering analysis with one minus Pearson correlation coefficients (1 - P_r) as the distance measure. The resulting heat-map based on the number of oligonucleotide sequence counts allows for visualisation of similarities in oligonucleotide copy numbers among all patient samples. See FIG. 15. In the figure, the X-axis shows individual oligonucleotides and the Y-axis shows samples (aggressive or other (i.e., non-aggressive)). Oligonucleotides from the indicated Cluster 1 have relatively low counts for the majority of aggressive cancers. Oligonucleotides from the indicated Clusters 2 and 3 have relatively higher counts in a majority of aggressive cancer samples. It is noted that oligonucleotide probes that are observed at either higher or lower relative copy number in the advanced cancer samples versus non aggressive can be used as markers to distinguish these sample groupings. Oligonucleotides from the indicated Cluster 4 were more selective toward particular aggressive cancer samples, and were able to identify additional 12 advanced cancer samples that were missed by oligonucleotides in Clusters 1-3. Cluster 4 oligonucleotides may be useful for detecting particular subsets of aggressive cancers.

The distances between samples within each cluster are considerably smaller than the distance between the two clusters. As can be seen in FIG. 15, these two groups of samples have very different proportions of advanced cancer cases versus other; in cluster I, 15 of the 19 patients are advanced cancer cases (79%), while in cluster II, only 16 of 40 patients are advanced cases (40%). Fisher exact p-value equals 0.01 with an odds ratio of 4.9, indicating that the ratio of advanced to other cancer in these two clusters is statistically significant.

The SEQ ID NOs. of the variable regions from the oligonucleotides in Clusters 1-4 in FIG. 15 are shown in Table 31. As in Table 27, the oligonucleotides were each synthesized with a 5' region consisting of the sequence (5’ -CTAGCAGCTGTACGTAGT (SEQ ID NO. 131)) and a 3’ region consisting of the sequence (5’ -CTGTCCTCTACACATCTCGCATCGCTGCGACGA (SEQ ID NO. 132)) flanking the variable region. The sequences can be used in the identification of aggressive breast cancer samples.

<table>
<thead>
<tr>
<th>Cluster</th>
<th>SEQ ID NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster 4</td>
<td>1072-1121</td>
</tr>
</tbody>
</table>

Table 31: Clusters 1-4 Variable Regions SEQ ID NOs
A selection of 50 negative control oligonucleotides was made from the same experiments as above. As criteria, we chose those oligonucleotides with the lowest variability (% cv) between biological samples when taking all 120 samples from the noted probing into account. The variable regions of the 50 negative controls are listed in SEQ ID NOs 1422-1471. As in Table 27, the oligonucleotides were synthesized with a 5’ region consisting of the sequence (5’ -CTAGCATGACTGCAGTACGT (SEQ ID NO. 131)) and a 3’ region consisting of the sequence (5’ -CTGTCTCTTATACATCTGACGCTCGGACGA (SEQ ID NO. 132)) flanking the variable regions. These sequences can be used as negative controls when identifying aggressive breast cancer samples. 

Oligonucleotides specific to non-cancer samples versus cancer samples

We also identified a set of oligonucleotides that were specific to the non-cancer plasma samples. In addition to probing the cancer samples noted above, the enriched oligonucleotide probe library was used to perform the probing test on the 60 individual non-cancer plasma samples. Each sample was probed as described above. The oligonucleotides recovered with each individual non-cancer and cancer plasma sample were identified using next-generation sequencing. The averaged frequencies of individual oligonucleotide sequence reads between non-cancer and cancer samples were used to calculate a fold-change of non-cancer versus cancer. Oligonucleotides with both read counts above 100 and fold change (non-cancer/cancer) greater or equal to 2 were identified as non-cancer specific oligonucleotide probes (SEQ ID NOs. 1472-1486). In addition, a selection of negative control oligonucleotides were selected as those with fold change less than or equal to 1 and read count below 20 (SEQ ID NOs 1487-1501).

The variable regions of the non-cancer specific oligonucleotides and negative control oligonucleotides are listed by rank order in SEQ ID NOs. 1472-1486 and SEQ ID NOs 1487-1501, respectively. As in Table 27, the oligonucleotides were synthesized with a 5’ region consisting of the sequence (5’ -CTAGCATGACTGCAGTACGT (SEQ ID NO. 131)) and a 3’ region consisting of the sequence (5’ -CTGTCTCTTATACATCTGACGCTCGGACGA (SEQ ID NO. 132)) flanking the variable regions. The sequences can be used to identify non-cancer samples (SEQ ID NOs. 1472-1486) or used as negative controls for such experiments (SEQ ID NOs 1487-1501).

Oligonucleotides specific to cancer samples versus non-cancer

We then identified a set of oligonucleotides that were specific to the cancer plasma samples as compared to the non-cancer samples. Pooled cancer and non-cancer samples were probed as described above. The oligonucleotides recovered with each individual cancer and non-cancer plasma sample were identified using next-generation sequencing. The averaged frequencies of individual oligonucleotide sequence reads between cancer and non-cancer samples were used to calculate a fold-change of cancer versus non-cancer. Oligonucleotides with both read counts above 10 and fold change (non-cancer/cancer) greater or equal to 1.5 were identified as cancer specific oligonucleotide probes (SEQ
ID NOs. 1502-1539). In addition, a selection of negative control oligonucleotides were selected as those with fold change less than or equal to 1 and read count below 20. See SEQ ID NOs 1487-1501 as described above.

The variable regions of the cancer specific oligonucleotides are listed by rank order in SEQ ID NOs. 1502-1539. As in Table 27, the full length oligonucleotides were synthesized with a 5’ region consisting of the sequence (5’-CTAGCATGACTGCAGTACGT (SEQ ID NO. 131)) and a 3’ region consisting of the sequence (5’-CTGTCCTTATACATCTGACGCTGCGCAGA (SEQ ID NO. 132)) flanking the variable regions. These sequences which can be used to identify cancer samples.

Example 28: Oligonucleotide probes: Aggressive Breast Cancer, Non-Aggressive Cancer and Non-cancer/Healthy blood plasma targets with UC and PEG

This Example builds on the work presented in Example 27. Based on the results of the probing Aggressive cancer pool (50% patients with aggressive breast cancer), Non-Aggressive cancer pool and Healthy Pool (50% biopsy confirmed non-cancer and 50% healthy non-cancer), we obtained aptamers specific for:

Aggressive cancer compared to non-aggressive cancer and to non-cancer/healthy samples.

Non-aggressive cancer compared to aggressive and to non-cancer/healthy samples.

Non-Cancer/Healthy samples compared to aggressive and non-aggressive samples.

A combination of ultracentrifugation and PEG based partitioning of microvesicles from each input plasma sample. This approach potentially provides a cleaner patient sample and thus provide higher specificity for the intended target. As in Example 27, the input oligonucleotide library was exposed to positive, negative and again positive targets sequentially and then PCR amplified. Enriched libraries were subjected to the oligonucleotide probing assay on the plasma samples using sample pools as follows: ACP - aggressive cancer pool; NACP - non-aggressive cancer pool; HP - healthy pool (mix of non-cancer and normal samples). Each sample pool was probed two naive (i.e., untrained) oligonucleotide probe libraries (6-3S and 8-3S described hereing) with three replicates. Samples were sequenced in multiplex with 9 samples per flow cell. Indexing PCR was performed either with or without internal standard. The average number of sequence read counts between three replicates was considered for fold change and TTEST calculations. Sequences were identified that met the following criteria: fold change between pools >= 2; p-value =<0.05, % cv =< 20%, counts >=100.

The joint selection of oligonucleotides that passed the above criteria is shown in Tables 32 and 33. As in Table 27, the oligonucleotides each have a 5’ region consisting of the sequence (5’-CTAGCATGACTGCAGTACGT (SEQ ID NO. 131)) and a 3’ region consisting of the sequence (5’-CTGTCCTTATACATCTGACGCTGCGCAGA (SEQ ID NO. 132)). Tables 32 and 33 list in rank order the SEQ ID NOs of the variable region that are positioned between the 5’ and 3’ regions. The differences are that the starting libraries were derived from a first round of enrichment against a cancer pool (Table 32) or a first round of enrichment against a non-cancer pool (Table 33). However, after the first round the aliquots were pooled for subsequent enrichment. The combined oligonucleotide probe library can be used to distinguish between aggressive breast cancer, non-aggressive breast cancer, and
non-cancer plasma samples in a single assay according to the column "Comparison." In the Comparison column, sequences elevated in group 1 as compared to group 2 are shown as group 1/ group 2. The legend for this column is "ACP": Aggressive cancer pool; "HP": healthy & non-cancer pool; "NACP": Non-aggressive cancer pool. "IntSt" in the comparison row indicates that an internal standard was used during the enrichment. By way of illustration, the first two sets of comparisons in Table 32 are labeled ACP/HP, which indicates that these oligonucleotides were observed at high copy numbers after probing the aggressive breast cancer sample pool versus the healthy/non-cancer sample pool. Continuing with this illustration, higher numbers of such oligonucleotide probes binding to a sample can indicate the presence of an aggressive breast cancer.

**Table 32: Oligonucleotide Probe Comparison and Variable Regions**

<table>
<thead>
<tr>
<th>Comparison</th>
<th>SEQ ID NO.</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACP/HP (IntSt)</td>
<td>1540-1551</td>
</tr>
<tr>
<td>ACP/HP</td>
<td>1552-1585</td>
</tr>
<tr>
<td>HP/ACP (IntSt)</td>
<td>1586-1611</td>
</tr>
<tr>
<td>HP/ACP</td>
<td>1612-1650</td>
</tr>
<tr>
<td>NACP/HP (IntSt)</td>
<td>1651-1655</td>
</tr>
<tr>
<td>NACP/HP</td>
<td>1656-1733</td>
</tr>
<tr>
<td>HP/NACP (IntSt)</td>
<td>1734-1754</td>
</tr>
<tr>
<td>HP/NACP</td>
<td>1755-1789</td>
</tr>
<tr>
<td>NACP/ACP (IntSt)</td>
<td>1790-1793</td>
</tr>
<tr>
<td>ACP/NACP (IntSt)</td>
<td>1794-1805</td>
</tr>
<tr>
<td>ACP/NACP</td>
<td>1806-1815</td>
</tr>
</tbody>
</table>

**Table 33: Oligonucleotide Probe Comparison and Variable Regions**

<table>
<thead>
<tr>
<th>Comparison</th>
<th>SEQ ID NO.</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACP/HP (IntSt)</td>
<td>1816-1822</td>
</tr>
<tr>
<td>ACP/HP</td>
<td>1823-2004</td>
</tr>
<tr>
<td>HP/ACP (IntSt)</td>
<td>2005-2054</td>
</tr>
<tr>
<td>HP/ACP</td>
<td>2055-2234</td>
</tr>
<tr>
<td>NACP/HP (IntSt)</td>
<td>2235-2242</td>
</tr>
<tr>
<td>NACP/HP</td>
<td>2243-2318</td>
</tr>
<tr>
<td>HP/NACP (IntSt)</td>
<td>2319-2356</td>
</tr>
<tr>
<td>HP/NACP</td>
<td>2357-2491</td>
</tr>
<tr>
<td>NACP/ACP (IntSt)</td>
<td>2492-2507</td>
</tr>
<tr>
<td>NACP/ACP</td>
<td>2508-2583</td>
</tr>
<tr>
<td>ACP/NACP (IntSt)</td>
<td>2584-2593</td>
</tr>
<tr>
<td>ACP/NACP</td>
<td>2594-2728</td>
</tr>
</tbody>
</table>

**Example 29: Oligonucleotide probes: Non-Aggressive Cancer probes identified in libraries enriched with multiple protocols**

This Example builds on the work presented in Example 28. Here, we identified a set of oligonucleotide probes specific for non-aggressive cancer samples. The oligonucleotide probes were identified based on probing using an aggressive cancer pool (ACP), non-aggressive cancer pool (NACP) and healthy pool (HP) of plasma samples with naive oligonucleotide libraries. Compared to the work in Example 28, four different enrichment schemes were used, as further detailed below. For purposes of
these experiments, non-aggressive (or non-advanced) cancer is defined as sample with tumor size not exceeding stage T1, presence of nodes NO, Metastasis MO and positive estrogen/progesterone receptors. In each case, the enriched libraries were subjected to the probing test on pooled plasma samples. The library enrichment was performed as described herein, see e.g., Example 21. The libraries were exposed to positive and negative targets and PCR amplified either after each binding or after series positive-negative-positive binding as described below. The positive and negative pools are indicated below for the different enrichment schemes. After each round of enrichment, next generation sequencing was used to determine a copy number of the recovered oligonucleotide probes. Averaged frequencies of individual oligonucleotide between the indicate pools were used to calculate the fold change. Oligonucleotide with counts above 50, fold change >=2 and %cv between replicates below 20% were considered as candidates and shown in the tables below.

<table>
<thead>
<tr>
<th>Scheme 1</th>
</tr>
</thead>
</table>

The positive target was the NACP plasma and the negative target was the HP plasma. The enrichment started with the naive (non-enriched; NE) library. Microvesicles in the samples were isolated using ultracentrifugation in all stages. Each round of enrichment consisted of a complete set of positive-negative-positive binding events followed by PCR amplification of the enriched library. Two such rounds of binding were performed. The resulting library after the second enrichment (R2) is used for subsequent probing of the pooled samples (aggressive cancer, non-aggressive cancer and healthy). As in Example 28 above, comparisons performed included ACP/HP, HP/ACP, ACP/NACP, NACP/ACP, HP/NACP, NACP/HP.

Resulting oligonucleotides identified in these selections are shown in Table 34. As in Table 27, the oligonucleotides each have a 5' region consisting of the sequence (5' - CTAGCATGACTGCAGTACGT (SEQ ID NO. 131)) and a 3' region consisting of the sequence (5' - CTGTCTTTATACATCTAGC(TGCGGACGA (SEQ ID NO. 132)). Table 34 lists the SEQ ID NOs in rank order of the variable regions that are positioned between the 5' and 3' regions. The combined oligonucleotide probe library can be used to distinguish between aggressive breast cancer, non-aggressive breast cancer, and non-cancer plasma samples in a single assay according to the column "Comparison." In the Comparison column, sequences elevated in group 1 as compared to group 2 are shown as group 1/ group 2. The legend for this column is "ACP": Aggressive cancer pool; "HP": healthy & non-cancer pool; "NACP": Non-aggressive cancer pool. By way of illustration, the first two sets of comparisons in Table 34 are labeled ACP/HP, which indicates that these oligonucleotides were observed at high copy numbers after probing the aggressive breast cancer sample pool versus the healthy/non-cancer sample pool. Continuing with this illustration, higher numbers of such oligonucleotide probes binding to a sample can indicate the presence of an aggressive breast cancer.

<table>
<thead>
<tr>
<th>Comparison</th>
<th>SEQ ID NO.</th>
</tr>
</thead>
<tbody>
<tr>
<td>HP/ACP</td>
<td>2729-2737</td>
</tr>
<tr>
<td>ACP/HP</td>
<td>2738-2742</td>
</tr>
<tr>
<td>ACP/NACP</td>
<td>2743-2745</td>
</tr>
</tbody>
</table>
Scheme 2

In this enrichment, the positive target is NACP plasma and the negative target is the HP plasma. The starting input library is the C-R8N-1 (i.e., previously enriched for a cancer pool with 48% of aggressive cancer samples; see description of the C-RN7 library in Example 26 with one additional round of positive selection performed). The microvesicle partitioning was performed using PEG precipitation. PCR amplification is done after each binding. The resulting library after this ninth (Round 9) round of enrichment is used for probing. Three different input criteria for used for round 9: 1) R9-2 - library input 0.1 ng; 2) R9-lc - library input is 1 ng; 3) R9 - 5x input of competitor DNA. The R9-2 and R9-lc oligonucleotide libraries were then used for subsequent probing of the pooled samples (aggressive cancer, non-aggressive cancer and healthy). As above, comparisons performed included ACP/HP, HP/ACP, ACP/NACP, NACP/ACP, HP/NACP, NACP/HP.

Resulting oligonucleotides identified from the R9-lc and R9-2 oligonucleotide libraries in these selections are shown in Table 35 and Table 36, respectively. As in Table 27, the oligonucleotides each have a 5’ region consisting of the sequence (5’ -CTAGCATGACTGCAGTACGT (SEQ ID NO. 131)) and a 3’ region consisting of the sequence (5’ -CTGTCTCTTATACACATCTGACGCTGCCGACGA (SEQ ID NO. 132)). Tables 35 and 36 list the SEQ ID NOs in rank order of the variable regions that are positioned between the 5’ and 3’ regions. When combined, the oligonucleotide probe libraries can be used to distinguish between aggressive breast cancer, non-aggressive breast cancer, and non-cancer plasma samples in a single assay according to the column "Comparison." In the Comparison column, sequences elevated in group 1 as compared to group 2 are shown as group 1 / group 2. The legend for this column is "ACP": Aggressive cancer pool; "HP": healthy & non-cancer pool; "NACP": Non-aggressive cancer pool. By way of illustration, the first two sets of comparisons in Table 35 are labeled ACP/HP, which indicates that these oligonucleotides were observed at high copy numbers after probing the aggressive breast cancer sample pool versus the healthy/non-cancer sample pool. Continuing with this illustration, higher numbers of such oligonucleotide probes binding to a sample can indicate the presence of an aggressive breast cancer.

Table 35: Enrichment Scheme 2 R9-lc Oligonucleotide Probe Comparison and Variable Regions

<table>
<thead>
<tr>
<th>Comparison</th>
<th>SEQ ID NO.</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACP/HP</td>
<td>2764-2791</td>
</tr>
<tr>
<td>HP/ACP</td>
<td>2792-2793</td>
</tr>
<tr>
<td>ACP/NACP</td>
<td>2794-2815</td>
</tr>
<tr>
<td>NACP/ACP</td>
<td>2816-2830</td>
</tr>
<tr>
<td>HP vs NACP</td>
<td>2831-2845</td>
</tr>
<tr>
<td>NACP vs HP</td>
<td>2846-2858</td>
</tr>
</tbody>
</table>

Table 36: Enrichment Scheme 2 R9-2 Oligonucleotide Probe Comparison and Variable Regions

<table>
<thead>
<tr>
<th>Comparison</th>
<th>SEQ ID NO.</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACP/HP</td>
<td>2746-2753</td>
</tr>
<tr>
<td>NACP/HP</td>
<td>2754</td>
</tr>
<tr>
<td>HP/NACP</td>
<td>2755-2763</td>
</tr>
</tbody>
</table>
[001231] Scheme 3

[001232] The positive target was the NACP plasma and the negative target was the HP plasma. The enrichment started with the naive (non-enriched; NE) library. Microve sides in the samples were isolated using PEG precipitation. Six rounds of enrichment were performed with PCR amplification performed after each round. The round alternated between a single selection of positive or negative selection (i.e., positive selection in rounds 1, 3, 5 and negative selection in rounds 2, 4, and 6). The resulting library after round 6 is used for probing of the pooled samples (aggressive cancer, non-aggressive cancer and healthy) with library input 0.1 ng. As above, comparisons performed included ACP/HP, HP/ACP, ACP/NACP, NACP/ACP, HP/NACP, NACP/HP.

[001233] Resulting oligonucleotides identified in these selections are shown in Table 37. As in Table 27, the oligonucleotides each have a 5’ region consisting of the sequence (5’ -CTAGCATGACTGAGCTACGT (SEQ ID NO. 131)) and a 3’ region consisting of the sequence (5’ -CTGTCTTTATACACATCTCTACGCTGGCGAAGA (SEQ ID NO. 132)). Table 37 lists the SEQ ID NOs in rank order of the variable regions that are positioned between the 5’ and 3’ regions. The combined oligonucleotide probe library can be used to distinguish between aggressive breast cancer, non-aggressive breast cancer, and non-cancer plasma samples in a single assay according to the column “Comparison.” In the Comparison column, sequences elevated in group 1 as compared to group 2 are shown as group 1/ group 2. The legend for this column is "ACP": Aggressive cancer pool; "HP": healthy & non-cancer pool; "NACP": Non-aggressive cancer pool. By way of illustration, the first two sets of comparisons in Table 37 are labeled ACP/HP, which indicates that these oligonucleotides were observed at high copy numbers after probing the aggressive breast cancer sample pool versus the healthy/non-cancer sample pool. Continuing with this illustration, higher numbers of such oligonucleotide probes binding to a sample can indicate the presence of an aggressive breast cancer.

Table 37: Enrichment Scheme 3 Oligonucleotide Probe Comparison and Variable Regions

<table>
<thead>
<tr>
<th>Comparison</th>
<th>SEQ ID NO.</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACP/HP</td>
<td>2951-2960</td>
</tr>
<tr>
<td>HP/ACP</td>
<td>2961-2972</td>
</tr>
<tr>
<td>ACP/NACP</td>
<td>2973-2976</td>
</tr>
<tr>
<td>NACP/ACP</td>
<td>2977-2990</td>
</tr>
<tr>
<td>HP vs NACP</td>
<td>2991-3003</td>
</tr>
<tr>
<td>NACP/HP</td>
<td>3004-3017</td>
</tr>
</tbody>
</table>

[001234] Scheme 4

[001235] Starting aptamer library is C-R8N-3 (library previously enriched for Cancer pool with 48% of aggressive cancer samples and input O.OOlng). Microve sides in the samples were isolated using
PEG precipitation. PCR amplification is done after each binding. There was only negative target used (the NACP), which is unique for the selection scheme versus the three above. This enrichment was meant to further enrich C-R8N library toward aggressive cancer targets by removing oligonucleotides specific to non-aggressive targets. Four negative selection rounds were performed. Two different input criteria for used for probing after round 13 (R13): 1) R13-3 - library input 0.01 ng; 2) R13-4 - library input 0.001 ng. The R13-3 and R13-4 oligonucleotide libraries were then used for subsequent probing of the pooled samples (aggressive cancer, non-aggressive cancer and healthy). As above, comparisons performed included ACP/HP, HP/ACP, ACP/NACP, NACP/ACP, HP/NACP, NACP/HP.

The variable regions of the resulting oligonucleotides identified from the R13-3 and R13-4 oligonucleotide libraries in these selections are shown in Table 38 and Table 39, respectively. As in Table 27, the oligonucleotides each have a 5’ region consisting of the sequence (5’ - CTAGCATGACTGCAATAGT  (SEQ ID NO. 131)) and a 3’ region consisting of the sequence (5’ - CTTCTCTATACATCTT GACGTGCGACGA  (SEQ ID NO. 132)). Tables 38 and 39 list the SEQ ID NOs in rank order of the variable regions that are positioned between the 5’ and 3’ regions. When combined, the oligonucleotide probe libraries can be used to distinguish between aggressive breast cancer, non-aggressive breast cancer, and non-cancer plasma samples in a single assay according to the column "Comparison." In the Comparison column, sequences elevated in group 1 as compared to group 2 are shown as group 1/ group 2. The legend for this column is "ACP": Aggressive cancer pool; "HP": healthy & non-cancer pool; "NACP": Non-aggressive cancer pool. By way of illustration, the first two sets of comparisons in Table 38 are labeled ACP/HP, which indicates that these oligonucleotides were observed at high copy numbers after probing the aggressive breast cancer sample pool versus the healthy/non-cancer sample pool. Continuing with this illustration, higher numbers of such oligonucleotide probes binding to a sample can indicate the presence of an aggressive breast cancer.

Table 38: Enrichment Scheme 4 R13-3 Oligonucleotide Probe Comparison and Variable Regions

<table>
<thead>
<tr>
<th>Comparison</th>
<th>SEQ ID NO.</th>
<th>Variable Region Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACP/HP</td>
<td>3018-3030</td>
<td>ACCATCGGCTACGTAAAGCGTTTCTAA</td>
</tr>
<tr>
<td>HP/ACP</td>
<td>3031-3043</td>
<td>ACCGACTGCGCCGAATCCTGATCCCCCTAAGTTTC</td>
</tr>
<tr>
<td>ACP/NACP</td>
<td>3044-3052</td>
<td>ACCACTGACTCAGTGGTACTAATCACACACACATT</td>
</tr>
<tr>
<td>NACP/ACP</td>
<td>3053-3061</td>
<td>ACTGACTGACATACATGCGCATATATAAACCGCTATT</td>
</tr>
<tr>
<td>HP/ACP</td>
<td>3062-3083</td>
<td>ACCGCTGTACGATGTCTCGTGCTCAGTTATTTCCTTCC</td>
</tr>
<tr>
<td>ACP/HP</td>
<td>3084-3096</td>
<td>ACTGCGGTCATCGAGGTTACTGCTATGCGCTCAGC</td>
</tr>
</tbody>
</table>

Table 39: Enrichment Scheme 4 R13-4 Oligonucleotide Probe Comparison and Variable Regions

<table>
<thead>
<tr>
<th>Comparison</th>
<th>SEQ ID NO.</th>
<th>Variable Region Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACP/HP</td>
<td>3097-3099</td>
<td>ATGCATTTGCGTATTCTCAATGGAAGGCTCCTAAA</td>
</tr>
<tr>
<td>HP/ACP</td>
<td>3100-3107</td>
<td>CATACGACAGTGTCCCTCCTACTTATAATTGTCCACCA</td>
</tr>
<tr>
<td>ACP/NACP</td>
<td>3108-3117</td>
<td>AGCCCGTTGCTACGTTGCAATGAGTTAATCTAATCTAG</td>
</tr>
<tr>
<td>NACP/ACP</td>
<td>3118-3130</td>
<td>AGCGTACCGGCTCAGTTATTCTATGTCATGTCAGT</td>
</tr>
<tr>
<td>HP/NACP</td>
<td>3131-3141</td>
<td>ACCGCGTGTACGTTGCAATGAGTTAATCTAATCTAG</td>
</tr>
<tr>
<td>NACP/HP</td>
<td>3142-3150</td>
<td>ACCACGTTGATTGCTACTGCTCCTGCTTGAAGTCCCTTACG</td>
</tr>
</tbody>
</table>
Example 30: Use of an oligonucleotide probe library to characterize Breast Cancer samples

[001237] An oligonucleotide probe library comprising approximately 2000 different probe sequences was constructed and used to probe approximately 500 individual breast cancer and non-cancer samples. The probe sequences were derived from different screening experiments and are listed herein in SEQ ID NOs 137-969 and 1072-3150. The oligonucleotides listed in these tables were synthesized and pooled together. The samples were plasma samples from 212 breast cancer patients, 177 biopsy confirmed non-cancer patients, and 117 normal control patients (self-reported as non-cancer). Experiments details are as provided above. The plasma samples were contacted with the oligonucleotide probe library and microvesicles were isolated using PEG precipitation. Oligonucleotides that were recovered with the microvesicles were isolated. Next Generation Sequencing (Illumina HiSeq) was used to identify the isolated sequences for each sample.

[001238] Analysis of significance of difference identified 18 aptamers with p-values below 0.01 when compared Cancer/Normal, 15 aptamers with p-values below 0.001 when compared cancer/Non-Cancer, 28 aptamers with p-values below 0.001 when compared Non-Cancer/Normal.

[001239] Multi-oligonucleotide panels were next constructed using a cross-validation approach. Briefly, 50 samples were randomly withheld from the sample cohort. The performance of individual oligonucleotides to distinguish the remaining cancers and non-cancer/normals was determined using logistic regression methodology. Additional oligonucleotides were added iteratively and performance was assessed using logistic regression until further performance improvements were no longer obtained with additional oligonucleotides. The approach generally led to panels of approximately 20-100 different probe sequences. The constructed panels were then used to classify the 50 withheld samples and diagnostic performance was assessed using Receiver Operating Curve (ROC) analysis and estimation of the Area Under the Curve (AUC).

[001240] In approximately 300 rounds of cross-validation, the average AUC was 0.6, thus showing that the average performance was statistically better than random (i.e., AUC of 0.5) and that the probe library could distinguish breast cancer and non-breast cancer/normal patient samples. AUC values as high as 0.8 were observed for particular cross validations. FIGs. 16A-B illustrate a model generated using a training (FIG. 16A) and test (FIG. 16B) set from a round of cross validation. The AUC was 0.803. The variable regions of the sequences used to build this model are shown in Table 40. Another exemplary round of cross-validation is shown in FIGs. 16C-D. The AUC was 0.678.

[001241] The SEQ ID NOs. of the sequences used in the model in FIGs. 16A-B are listed in rank in Table 40. As in Table 27, the oligonucleotides were synthesized with a 5' region consisting of the sequence (5' -CTAGCATGACTGCAGTACGT (SEQ ID NO. 131)) and a 3' region consisting of the sequence (5' -CTGTCTCTTATACACATCTGACGCTGCCGACGA (SEQ ID NO. 132)) flanking the variable regions.

### Table 40: Oligonucleotide Probe Variable Regions

<table>
<thead>
<tr>
<th>Rank Ordered SEQ ID NOs</th>
</tr>
</thead>
</table>
The data presented in this Example demonstrate that an oligonucleotide pool comprising members having the variable regions listed in SEQ ID NOs 137-969 and 1072-3150, e.g., a pool of probes having the variable regions listed in Table 40, can be used to distinguish plasma from individuals having breast cancer versus plasma from non-breast cancer individuals.

Example 31: Updated oligonucleotide pools to characterize Breast Cancer samples

The profiling of 500 clinical samples with an oligonucleotide probe library comprising 2000 oligonucleotides showed significant ability to distinguish breast cancer from non-cancer control samples. See Example 30 above. When performing cross-validation in these experiments, it was observed that 85 out of 500 were misclassified at a higher rate compared to the other samples. We therefore identified another selection of oligonucleotide probe which were able to correctly classify the noted samples. These 85 samples were profiled with a naive oligonucleotide probe library (6-3S/8-3S), and enriched on breast cancer and non-cancer plasma using methodology presented in the Examples above. The selected oligonucleotides were compared to a positive controls cohort, which comprised of the cancer and non-cancer samples that were consistently classified correctly within the experiments in Example 30 above. Oligonucleotides were selected based on absolute number of copies sequenced of at least 50 and various criteria when comparing sample groups including: 1) copy number fold-change (fc) of at least 1.2; 2) effect size (es) above 0.6; 3) t-test (p-value <= 0.05); and 4) Kolmogorov-Smirnov test (ks) (p-value <= 0.05). Effect size is a population effect size calculated as (mean(group1) - mean(group2)) / standard deviation(group 1 and group2).

The profiling data were normalized by dividing the count of each particular oligonucleotide by the total counts for particular sample and multiplying by the global mean across the entire experiment. These normalized values were used to calculate the statistical criteria specified above to compare the following samples type: Misclassified Cancer ("C"), misclassified non-cancer ("NC"), positive control Cancer ("C-P"), positive control Non-Cancer ("NC-P"), misclassified Normal ("N"). The comparisons performed and numbers of resulting oligonucleotides are shown in Table 41. Negative controls are oligonucleotides that did not match any criteria and thus should not distinguish be any samples groups.

Table 41: Oligonucleotide probe candidates were selected from the following comparisons

<table>
<thead>
<tr>
<th>Specificity</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>C/NC-P* A** and C/NC-P B***</td>
<td>82</td>
</tr>
<tr>
<td>C/NC-P A</td>
<td>108</td>
</tr>
<tr>
<td>C/NC-P B</td>
<td>13</td>
</tr>
<tr>
<td>Negative control</td>
<td>60</td>
</tr>
<tr>
<td>C-P/NC A C-P/NC B</td>
<td>34</td>
</tr>
<tr>
<td>C-P/NC A</td>
<td>68</td>
</tr>
<tr>
<td>C-P/NC B</td>
<td>101</td>
</tr>
<tr>
<td>NC/C-P A NC/C-P B</td>
<td>7</td>
</tr>
<tr>
<td>NC/C-P A</td>
<td>196</td>
</tr>
<tr>
<td>NC-P/C A NC-P/C B</td>
<td>65</td>
</tr>
<tr>
<td>-------------------</td>
<td>----</td>
</tr>
<tr>
<td>NC-P/C A</td>
<td>83</td>
</tr>
<tr>
<td>NC-P/C B</td>
<td>55</td>
</tr>
<tr>
<td>C+C-P/N</td>
<td>103</td>
</tr>
<tr>
<td>N/C+C-P</td>
<td>25</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>1000</strong></td>
</tr>
</tbody>
</table>

[C/NC-P] means that oligonucleotides were selected with specificity toward cancer ("C") while compared to Non-Cancer positive ("NC-P") control.

**A** - selection criteria: counts > 50, fc ≥ 1.2, es > 0.6, ttest p<0.05.

**B** - selection criteria: counts >200, estimated sample size <71, fc >1.2, ttest/ks test p<0.05.

The sequences selected based on **Table 41** are shown in **Table 42**. As in **Table 27**, the oligonucleotides each have a 5' region consisting of the sequence (5' -CTAGCATGACTGCAGTACGT (SEQ ID NO. 131)) and a 3' region consisting of the sequence (5’ -CTGTCTCTTATACACATCT (SEQ ID NO. 132)). **Table 42** lists in rank order the SEQ ID NOs of the variable regions that are positioned between the 5' and 3' regions. The column “Specificity” shows the specificity of the oligonucleotides based on the comparisons in **Table 41**.

**Table 42: Oligonucleotide Probe Variable Regions**

<table>
<thead>
<tr>
<th>Specificity</th>
<th>SEQ ID NO.</th>
</tr>
</thead>
<tbody>
<tr>
<td>C/NC-P A and C/NC-P B</td>
<td>3151-3232</td>
</tr>
<tr>
<td>C/NC-P A</td>
<td>3233-3340</td>
</tr>
<tr>
<td>C/NC-P B</td>
<td>3341-3353</td>
</tr>
<tr>
<td>Negative control</td>
<td>3354-3368</td>
</tr>
<tr>
<td>C-P/NC A C-P/NC B</td>
<td>3369-3402</td>
</tr>
<tr>
<td>C-P/NC A</td>
<td>3403-3470</td>
</tr>
<tr>
<td>C-P/NC B</td>
<td>3471-3571</td>
</tr>
<tr>
<td>Negative control</td>
<td>3572-3586</td>
</tr>
<tr>
<td>NC/C-P A NC/C-P B</td>
<td>3587-3593</td>
</tr>
<tr>
<td>NC/C-P A</td>
<td>3594-3789</td>
</tr>
<tr>
<td>Negative control</td>
<td>3790-3804</td>
</tr>
<tr>
<td>NC-P/C A NC-P/C B</td>
<td>3805-3869</td>
</tr>
<tr>
<td>NC-P/C A</td>
<td>3870-3952</td>
</tr>
<tr>
<td>NC-P/C B</td>
<td>3953-4007</td>
</tr>
<tr>
<td>Negative control</td>
<td>4008-4022</td>
</tr>
<tr>
<td>C/N</td>
<td>4023-4125</td>
</tr>
<tr>
<td>N/C</td>
<td>4126-4150</td>
</tr>
</tbody>
</table>

The 1000 selected oligonucleotides are synthesized as DNA with a 5’ biotinylation. The synthesized oligonucleotides are pooled with 1000 of the most informative aptamers from the profiling performed in **Example 30** to create a further optimized oligonucleotide probing library that can be used to distinguish cancer and normal samples.

**Example 32: Comparison of Oligonucleotide Pools**

In this Example, the performance of the oligonucleotide pool from **Example 30** was compared to the performance of the oligonucleotide pool from **Example 31** to characterize breast cancer...
samples. In this Example, the oligonucleotide pool from Example 30 is referred to as the 2000vl and the oligonucleotide pool from Example 30 is referred to as the 2000v2. A random forest approach was used to analyze the data, which was obtained for the 2000vl library in Example 30. The 2000v2 data was obtained via tested on a cohort of 858 plasma samples which comprised some samples described in Example 30 with additional samples according to three groupings: 194 breast cancer positive, 382 breast cancer negative as confirmed by biopsy (“non-cancer”), and 282 self declared normals (“normal”).

[001251] ROC curves are shown in FIGs. 22A-D for the various settings. AUC values are shown above the curve and are statistically significant from random (i.e., AUC = 0.5) in all cases. In FIG. 22A, the 2000vl probe library was used to distinguish cancer samples versus biopsy confirmed non-cancer samples. In FIG. 22B, the 2000v2 probe library was used to distinguish the same sample groupings. A small improvement in AUC was observed with the 2000v2 probe library in this setting. In FIG. 22C, the 2000vl probe library was used to distinguish cancer samples versus self-declared normal samples. In FIG. 22D, the 2000v2 probe library was used to distinguish the same sample groupings. A small decline in AUC was observed with the 2000v2 probe library in this setting. Highly stringent

[001252] The SEQ ID NOs. of the variable regions of informative oligonucleotides from the 2000vl and 2000v2 libraries are listed in Table 43. As above, the oligonucleotides were synthesized with a 5’ region consisting of the sequence (5’-CTAGCATGCAGTGACT (SEQ ID NO. 131)) and a 3’ region consisting of the sequence (5’-CTGTCGCTACACTCT (SEQ ID NO. 132)) flanking the variable regions. The listed oligonucleotides were used in the Random Forest modeling to generate AUC values, e.g. as shown in FIGs. 22A-D, and are listed in rank order. The row “Common” in Table 43 shows overlapping informative sequences from the 2000vl and 2000v2 settings.

### Table 43: Informative oligonucleotide probes

<table>
<thead>
<tr>
<th>Setting</th>
<th>SEQ ID NO.</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000vl</td>
<td></td>
</tr>
<tr>
<td>2000v2</td>
<td></td>
</tr>
</tbody>
</table>
The data presented in this Example demonstrate that an oligonucleotide pool comprising members having the variable regions listed in SEQ ID NOs 137-969 and 1072-4150, e.g., a pool of probes having the variable regions listed in Table 43, can be used to distinguish plasma from individuals having breast cancer versus plasma from non-breast cancer individuals.

**Example 33: Photo-cleavable Biotin mediated purification of an oligonucleotide library**

Tag probe and capture probe detection systems may have advantages in the read-out of oligonucleotide-based assays since such systems allow direct quantification of oligonucleotides without amplification. Amplification may introduce amplification bias, e.g., wherein some sequences are preferentially amplified versus others. One such system is the nCounter nucleic acid detection system from NanoString Technologies, Inc. (Seattle, WA). This system can be used to detect oligonucleotide probes of the invention. See, e.g., FIG. 14G and related discussion. Current applications of tag and capture probe detection systems include gene expression analysis, RNA:Protein analysis; single cell gene expression analysis; miRNA expression analysis, miRGE analysis, copy number variation analysis, lncRNA expression analysis; ChiP-String expression analysis. See, e.g., www.nanostring.com/applications/.

In this Example, we improved using the Nanostring system to detect the oligonucleotide probe libraries of the invention. As described above, a pre-enriched oligonucleotide probe library is contacted with a biological sample such as plasma, and microvesicles bound by various oligonucleotide probes are precipitated, e.g., with PEG. Thus, the oligonucleotide probe library is present in the mix with PEG/proteins. This mixture is compatible with PCR steps used for next-generation sequencing (NGS) preparation, but the Nanostring system may not detect probes in these mixtures. Without being bound by theory, this interference may be due to noted impurities with hybridization of the probes and other oligonucleotides used in purifications steps. We developed a protocol for purification of post-probing oligonucleotide probe library, which utilizes photo-cleavable (PC) Biotin (see, e.g., www.idtdna.com/site/Catalog/Modifications/Product/2291; references 1-4) and thus allows oligonucleotide probe detection using the Nanostring tag-capture system.

The steps of the purification of post-probing oligonucleotide probe library are shown in schematic 1700 in FIG. 17A and include the following: (i) modification of the 5’-end of each member of the oligonucleotide library with photocleavable-Biotin (PC-Biotin) 1701 via PCR with a PC-Biotinylated primer to create PC-Biotinylated aptamer library 1702; (ii) probing of plasma samples using the PC-Biotinylated oligonucleotide probe library followed by PEG precipitation as described herein, resulting in a pellet 1703 which contains, inter alia, bound oligonucleotide library members, proteins and microvesicles 1705, and PEG; (iii) heat denaturing of the re-suspended pellet to break
oligonucleotide/protein interactions and make PEG less viscous before mixing with streptavidin beads
1704; (iv) after short incubation with streptavidin beads 1704 to allow library 1702 to bind to the beads
1706, washing of the beads to remove impurities; (v) UV photo-cleavage 1707 of the spacer between the
PC-Biotin and oligonucleotide molecules, capturing the beads 1708 using a magnet and collecting the
released oligonucleotide library 1709 in the solution; and (vi) Nanostring 1710 sample preparation
according to the manufacturer's instructions. In FIG. 17A, "PC-B" refers to photocleavable biotin,
oligonucleotide library members are referred to as "AL," "Str. Beads" refers to streptavidin beads, and
"NS" refers to the Nanostring assay. The steps above are identified in the plot.

FIG. 17B show a 4% agarose gel with results of photocleavable biotin library cleavage
test with Streptavidin protein. Lane 1 is a 50bp DNA ladder and lane 10 is a Control with no biotin library
with Streptavidin protein. The gel shows that both biotinylated ("B") and photo-cleavable biotinylated
("PC-B") modified oligonucleotides were bound to Streptavidin, which resulted in slower mobility in gel
(lanes 4 and 5) as compared to the unbound oligonucleotides (lanes 2 and 3). After exposure to 365 nm
UV radiation, the PC-B oligonucleotides were photo-cleaved, which resulted in faster mobility in gel and
returning to the original molecular weight (i.e., before mixing with Streptavidin protein) (lanes 7-9, where
minutes UV exposure are indicated). In contrast, no change to the B oligonucleotides was observed after
10 min exposure to UV (lane 8). These results demonstrate that the photocleavable biotin library can be
captured by streptavidin beads and released by UV-mediated cleavage.

To test the workflow, a pool of plasma from patients with breast cancer and another pool
of plasma from healthy donors were probed with an oligonucleotide library comprised of 100
oligonucleotide probes, 50 of which were selected as cancer specific and other 50 were selected as
equivalent binders to both types of samples (i.e., "neutral"). The Nanostring probe system 1720 was
designed as in FIG. 17C. See FIG. 14G and description thereof for further details of the Nanostring probe
system. The "Nanostring Probe A" 1722 and "Nanostring Probe B" 1727 were designed in part as reverse
complements to the oligonucleotide probe library members of the invention (labeled "Aptamer" 1724 in
the figure). As can be seen in FIGs. 17D-E, relevant Nanostring counts were obtained for the
photocleavable oligonucleotide library (min relevant count is 300). Furthermore, the oligonucleotide
probes selected as binders ("Binder") for the cancer pool ("CP"), indeed show stronger binding to the
cancer pool as compared to counts from healthy pool ("HP"). The binding ratio between cancers and
normal was -5.25. For oligonucleotide probes selected as neutral binders ("Neutral"), there was no
relevant difference between the cancer pool and healthy pool.

The Nanostring protocol uses hybridization of target specific probes. We tested 100
different oligonucleotide probes, which are complementary to portion of each specific oligonucleotide, as
well as universal Primer probe for Nanostring detection using the above procedure. Following the
standard Nanostring detection protocol, all probes were mixed with recovered and purified library,
hybridized overnight and efficiently detected using the Nanostring tag-capture probe approach. In
contrast, without purification no quantification of oligonucleotides with Nanostring was observed (data
not shown).
In this Example, we established a process for oligonucleotide probe library purification after probing test samples. This process is efficient, low cost, can be accomplished in short time frame, and can be used for any purpose when library needs to be purified from probing product directly and quickly, given no biotin labeling is needed in downstream applications.

References:

Example 34: Single stranded DNA (ssDNA) oligonucleotide library preparation for library development

The preparation of high yield and high quality ssDNA libraries is a critical step in SELEX (Systematic Evolution of Ligands by Exponential enrichment) [1, 2] as well as in other biological applications, such as DNA chips and microarrays [3], and single-stranded conformation polymorphism technique (SSCP) [4]. The standard approach for preparing ssDNA libraries includes PCR amplification to first generate a double stranded (dsDNA) library, followed by ssDNA separation and purification. Several strategies of ssDNA preparation have been developed to date, each with advantages and disadvantages:

Lambda exonuclease digestion [2, 5-7]

The dsDNA standard PCR product is followed by Lambda exonuclease to digest the complementary strand and leave the target ssDNA. ssDNA purification is then performed to remove enzymes and unwanted buffer.

Advantages: Regular PCR amplification has high yield in generating dsDNA.

Disadvantages: The purity of final ssDNA is limited by enzyme digestion efficiency. Also dsDNA needs to be purified prior to digestion, together with post-digestion purification there will be two purifications, which results in substantial loss of input material. The digestion usually requires at least 2 hours. The digestion rate may not be consistent.

Asymmetric PCR [8, 9]

The procedure generates target ssDNA as the main product and less dsDNA products and non-target ssDNA. The band corresponding to the target ssDNA is cut from a native gel.

Advantages: The final ssDNA product potentially has high purity.

Disadvantages: Separation of strands is possible in the native gel, but the yield is typically low and the presence of non-target strand cannot be excluded. The yield cannot be increased on denaturing gel because the strands have the same length.
Biotin-streptavidin magnetic beads separation

The non-target PCR primer is biotinylated so final PCR products are Biotinylated-dsDNA, which can be captured by streptavidin magnetic beads and denatured to release the non-biotin labeled target ssDNA.

Advantages: The final ssDNA product has relatively high purity.

Disadvantages: In most cases, the input library needs to be biotinylated, but it may be difficult to replace or release the captured target strands from streptavidin beads. Post-denaturing purification is required to remove NaOH and/or acid used for neutralization.

Unequal primer length PCR

The non-target PCR primer has a chemical modified spacer and a few extra nucleotides following. In the PCR reaction, the DNA polymerase will stop at the spacer, resulting in unequal length of PCR dsDNA product. Then target ssDNA can be cut from a denaturing PAGE gel.

Advantages: The final ssDNA product has high purity because the target ssDNA is not mixed with non-target strands.

Disadvantages: ssDNA cannot be seen on native gel. Requires time consuming denaturing PAGE gel. It may be difficult to denature some dsDNA library, which can limit the final yield.

Indirect purification method

The indirect purification strategy combines Asymmetric PCR and Biotin-streptavidin magnetic beads separation. In short, regular PCR is used to generate sufficient template, then asymmetric PCR with excess of target primer and less biotinylated complementary primers, followed by biotin-streptavidin separation.

Advantages: May increase yield and purity of ssDNA product.

Disadvantages: It cannot produce biotinylated target ssDNA library. The process is relatively long and complicated and may be prone to generate mutants of the original sequence.

The invention provides methods of enriching oligonucleotide probe libraries against a target of interest. As the probes comprise ssDNA, the process may comprise PCR amplification then conversion back into ssDNA after each round of enrichment. In this Example, we developed a strategy for preparation of a ssDNA oligonucleotide library. The goals were to develop a process that is efficient and quick, while delivering high quality/purity ssDNA. We aimed to combine PCR and ssDNA prep in one step, remain efficient in the presence of selection buffer, target molecules, other sample components (e.g., highly abundant proteins for plasma samples) and other assay components (e.g., PEG precipitation solution that may be used to precipitate microvesicles). In addition, we desired the method to be able to generate ssDNA library with any modification, including without limitation Biotin.

We have used an optimized version of Lambda exonuclease digestion protocol for preparation of ssDNA oligonucleotide library. However, the digestion yield limits the overall recovery and is not consistent between different library preparations. In some cases, the ssDNA band is hardly visible on the gel following digestion. We have also observed incomplete digestion of dsDNA in the ssDNA product. In this Example, we developed an alternative protocol, termed "ssDNA by Unequal
length PRimer Asymmetric PCR,” or SUPRA. It lacks disadvantages from the known methods listed above, and provides high quality and yield up to 10x higher yield of ssDNA oligonucleotide library as compared to the previous methods. It is relatively fast and convenient technically, since target ssDNA can be distinguished from non-target DNA on a gel.

A schematic comparing standard PCR 1800 and unequal length PCR 1810 is shown in FIG. 18A. In regular PCR 1800, a forward primer 1801 and reverse primer 1803 are hybridized with the reverse strand of an aptamer library 1802. The PCR reaction is performed, thereby creating equal length forward 1804 and reverse strands 1802. The strands are denatured in equal length single strands 1805. In unequal length PCR 1801, a forward primer 1811 having a lengthener segment and terminator segment and a reverse primer 1813 are hybridized with the reverse strand of an aptamer library 1812. The PCR reaction is performed, thereby creating unequal length forward 1814 and reverse strands 1812. The strands are denatured into unequal length single strands 1814 and 1812 that can be separated by size, e.g., on a denaturing gel.

The steps of SUPRA include: (i) Modification of regular non-target primer with two Isp9 (Internal Spacer 9; triethylene glycol spacer) as terminator and 32 extra nucleotides (e.g., poly-A) as lengthener. It is referred as Unequal-Forward-Double isp9 primer (UF-D9); (ii) Perform asymmetric PCR, by mixing DNA template, UF-D9 and regular target (reverse) primer at ratio that favors the reverse primer, e.g., 1:37.5. The PCR program has longer elongation step (e.g., 3 min instead of standard 1 min) and more cycles due to linear amplification mode (instead of exponential). The PCR product contains a majority of target ssDNA and small portion of dsDNA. (iii) Mix PCR reaction products 1:1 with denaturing buffer (e.g., 180 mM NaOH and 6 mM EDTA) and denature samples by heating (e.g., 70°C for 10 min) and cooling (e.g., incubation on ice for 3 min); (iv) Run denatured products in denaturing buffer on an agarose gel stained with SybrGold. The non-target strand, which is longer due to the lengthener, will appear as upper band (if visible) and the target strand (strong lower band) is cut and purified. The process can include optional steps, including without limitation: (v) Weigh the gel pieces and purify ssDNA from the gel pieces (e.g., using the ssDNA Nucleospin kit or the like); (vi) quantification of the yield and native gel can be used to check the purity and yield of final product (e.g., using the ssDNA Qubit kit or the like).

The first step (i) uses a specific design of the forward primer with efficient terminator and lengthener, which creates non-target strand of unequal length. The DNA polymerase used to build the target strand will stop polymerization once it reaches the terminator, and the lengthener facilitates differentiation between the target and non-target strands. In the second step (ii), the ratio between the two primers is shifted toward the reverse primer, to produce a majority of target ssDNA. The ratio, however, should not limit double strand templates production to keep reaction running. FIG. 18B is a gel showing titration of forward and reverse primers input in asymmetrical PCR. The optimal condition, at which target strand is clearly visible, is in the range 1:20-1:50 F:R primers ratio. As shown in the figure, the ratio between two primers in asymmetric PCR can affect dsDNA and ssDNA amount in final products. The PCR thermocycler program is also adjusted to provide efficiency in the asymmetric PCR. In the third step
(iii), a reliable denaturing method is used to separate target ssDNA to ensure the final yield and high purity.

[001292] As desired, the final step (vi) estimates the ratio of residual dsDNA, e.g., using ssDNA Qubit kit. In cases where the yield is not critical, the denaturing steps (iii and iv) can be skipped and the PCR products can be directly run on native gel. There will be a dsDNA band, but lower MW target ssDNA band can be distinguished and purified from gel. This is also a way to visualize the target band directly after PCR for a quality check or purification without denaturing. The purity of final product will be the same but yield will be lower.

[001293] A comparison of native versus denatured gel purification is shown in FIG. 18C. A post-probing oligonucleotide probe library was PCRed using unequal length primers mixed at a ratio of 1:38 (Forward/Reverse). In the figure, the left lane on each gel is a 50 bp molecular weight ladder and the lower band is the reverse primer. The positions of the dsDNA and ssDNA are indicated. A native gel showed the presence of both dsDNA and ssDNA (target strand) (FIG. 18C, panel A). Here, part of the target reverse strand is migrating in dsDNA. Thus, using the native gel, one can purify target ssDNA with moderate recovery. When a higher yield is desired, the PCR products can be run on denaturing agarose gel as described above. This approach provides maximal recovery wherein only target strand is visible, and can be cut from gel and purified (FIG. 18C, panel B). In this case, the reverse strand ssDNA, which is part of the dsDNA on native gel (FIG. 18C, panel A), is denatured and migrates together with other free molecules of target ssDNA strand, while forward strand becomes invisible due to limited amplification.

[001294] Compared to standard asymmetric PCR, which has relatively low yield and does not allow to distinguish target and non-target strands on denaturing gel, SUPRA delivers different lengths of target and non-target that can be purified on both native gel and denaturing gels. Compared to unequal primer length PCR, which uses lengthy Urea-PAGE protocol and produces only dsDNA, SUPRA has less dsDNA and free target ssDNA can be cut even from native gel if yield is not critical.

[001295] SUPRA has been used in the oligonucleotide probe library enrichment methods provided by the invention. The method is robust. In the presence of enrichment buffer, target/non-target molecules, proteins, exosomes/microvesicles, PEG and other components, SUPRA provides high quality and quantity of the ssDNA oligonucleotide library.

[001296] References:


Example 35: Anti-CIq oligonucleotides

This Example presents identification of anti-CIq aptamers. In this Example, biotinylated forms of the oligonucleotides identified in Example 27 were synthesized and immobilized to streptavidin beads and incubated with the PEG precipitated plasma fraction. Proteins that immunoprecipitated with oligonucleotides from SEQ ID NO. 1472-1486 were separated by SDS-PAGE. Specific bands were excised from the gel and proteins therein identified by in-gel trypsin digestions and LC-MS/MS such as in Example 23 above.

Immunoprecipitation of specific oligonucleotides conjugated to magnetic beads was performed to identify oligonucleotide targets. Several oligonucleotides were identified that recognized CIq and subunits. The sequences of the variable regions of identified anti-CIq aptamers include 5’-ACTATAGAACAATCCACCTTTAAAATCAACTATCTTA (SEQ ID NO. 1472), also known by the identifier 83S-B1; 5’-TACCCCTGACAACTCCTCGCGGAGGCTCCCATCTCTGA (SEQ ID NO. 1475), also known by the identifier 83S-B4; 5’-GATTITAAAACCTTCACCTGATTGTGCCAATCCA (SEQ ID
AACC CAAT TCACATACACTCTCACCCCCACTAAACA (SEQ ID NO. 1482), also known by the identifier 83S-B6; and 5’ -AACCCAATTCACATACACTCTCGCCCTCACTAAACA 4177
AACCCAATTCACATACACTCTCTCCCTCACTAAACA 4178
AACCCAATTCACATACACTCTCACCCTCACTAAACA 4179

Further as noted below, an oligonucleotide comprising the reverse complement of this putative consensus sequence did not bind Clq. Several variants of the anti-Clq aptamers were also identified in the experiments of Example 27. The sequences of the variable regions thereof are shown in Table 44. As above, the full length oligonucleotide probes in Table 44 each had a 5’ region consisting of the sequence (5’ -CTAGCATGACTGCAGTACGT) (SEQ ID NO. 13 1)) and a 3’ region consisting of the sequence (5’ -CTGCTCTTTATACACATCTGGCAGCA (SEQ ID NO. 132)). Table 44 shows the variable region that is positioned between the 5’ and 3’ regions of the sequences.

Table 44: Anti-Clq Variable Regions

<table>
<thead>
<tr>
<th>Variable Region</th>
<th>SEQ ID NO.</th>
</tr>
</thead>
<tbody>
<tr>
<td>83S-B1 (SEQ ID NO. 1472) and variants below</td>
<td></td>
</tr>
<tr>
<td>ACTATAGAACATCCACCGCTTTAAATCAACTATCTTAA</td>
<td>1472</td>
</tr>
<tr>
<td>ACTATAGAACATCCACCGCTTTAAATCAACTATCTGA</td>
<td>4152</td>
</tr>
<tr>
<td>ACTATAGAACATCCACCGCTTTAAATCAACTATCTCA</td>
<td>4153</td>
</tr>
<tr>
<td>ACTATAGAACATCCACCGCTTTAAATCAACTATCTTA</td>
<td>4154</td>
</tr>
<tr>
<td>ACTATAGAACATCCACCGCTTTAAATCAACTATCTGA</td>
<td>4155</td>
</tr>
<tr>
<td>ACTATAGAACATCCACCGCTTTAAATCAACTATCTTA</td>
<td>4156</td>
</tr>
<tr>
<td>ACTATAGAACATCCACCGCTTTAAATCAACTATCTGA</td>
<td>4157</td>
</tr>
<tr>
<td>ACTATAGAACATCCACCGCTTTAAATCAACTATCTCA</td>
<td>4158</td>
</tr>
<tr>
<td>ACTATAGAACATCCACCGCTTTAAATCAACTATCTTA</td>
<td>4159</td>
</tr>
<tr>
<td>ACTATAGAACATCCACCGCTTTAAATCAACTATCTGA</td>
<td>4160</td>
</tr>
<tr>
<td>ACTATAGAACATCCACCGCTTTAAATCAACTATCTCA</td>
<td>4161</td>
</tr>
<tr>
<td>ACTATAGAACATCCACCGCTTTAAATCAACTATCTTA</td>
<td>4162</td>
</tr>
<tr>
<td>ACTATAGAACATCCACCGCTTTAAATCAACTATCTGA</td>
<td>4163</td>
</tr>
<tr>
<td>ACTATAGAACATCCACCGCTTTAAATCAACTATCTCA</td>
<td>4164</td>
</tr>
</tbody>
</table>

| 83S-B4 (SEQ ID NO. 1475) and variants below | |
| TACCCTTGAACATCCCTTCGCGCGCGAGGCCCACTCCCTGGA | 1475 |
| TACCCTTGAACATCCCTTCGCGCGCGAGGCCCACTCCCTCA | 1485 |
| TACCCTTGAACATCCCTTCGCGCGCGAGGCCCACTCCCTAA | 4165 |
| TACCCTTGAACATCCCTTCGCGCGCGAGGCCCACTCCCTAA | 4166 |

| 83S-B6 (SEQ ID NO. 1477) and variants below | |
| GATTAAAAACCTTGGACACCGATTTTGCCAATCCA | 1477 |
| GATTTTAAAAACCTTGGACACCGATTTTGCCAATCAA | 4167 |
| GATTTAAAAACCTTGGACACCGATTTTGCCAATCAG | 4168 |
| GATTTAAAAACCTTGGACACCGATTTTGCCAATCTA | 4169 |

<p>| 83S-B11 (SEQ ID NO. 1482) and variants below | |
| AACCCAAATTACATACACTCTCACCCCCACCTAAACA | 1482 |
| AACCCACCCATAACACTCTCACCCCCAAACACTAAACA | 4170 |
| AACCCCAATTACATACACTCTCACCCCCAAACACTAAACA | 4171 |
| AACCCCAATTACATACACTCTCACCCCCAAACACTAAACA | 4172 |
| AACCCCAATTACATACACTCTCACCCCCAAACACTAAACA | 4173 |
| AACCCCAATTACATACACTCTCACCCCCAAACACTAAACA | 4174 |
| AACCCCAATTACATACACTCTCACCCCCAAACACTAAACA | 4175 |
| AACCCCAATTACATACACTCTCACCCCCAAACACTAAACA | 4176 |
| AACCCCAATTACATACACTCTCACCCCCAAACACTAAACA | 4177 |
| AACCCCAATTACATACACTCTCACCCCCAAACACTAAACA | 4178 |
| AACCCCAATTACATACACTCTCACCCCCAAACACTAAACA | 4179 |</p>
<table>
<thead>
<tr>
<th>Sequence</th>
<th>4234</th>
</tr>
</thead>
<tbody>
<tr>
<td>AACCCAAATCATAACCTCTACACCTCATAAACAA</td>
<td>4180</td>
</tr>
<tr>
<td>AACCCAAATCCATAACCTCTACACCTCATAAACAA</td>
<td>4181</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4182</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4183</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4184</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4185</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4186</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4187</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4188</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4189</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4190</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4191</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4192</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4193</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4194</td>
</tr>
<tr>
<td>ATCCCAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4195</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4196</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4197</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4198</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4199</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4200</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4201</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4202</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4203</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4204</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4205</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4206</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4207</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4208</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4209</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4210</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4211</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4212</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4213</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4214</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4215</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4216</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4217</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4218</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4219</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4220</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4221</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4222</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4223</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4224</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4225</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4226</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4227</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4228</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4229</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4230</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4231</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4232</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4233</td>
</tr>
<tr>
<td>AACCCAAATTCAATACACCTCTACACCTCATAAACAA</td>
<td>4234</td>
</tr>
<tr>
<td>Oligonucleotide Sequence</td>
<td>Sequence Length</td>
</tr>
<tr>
<td>--------------------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>AACCCAAATTCCATACAGCTTCTACCCCTCAGTAAAAA</td>
<td>4290</td>
</tr>
<tr>
<td>AACCCAAATTCCATACAGCTTCTACCCCTCAGTAAAAA</td>
<td>4291</td>
</tr>
<tr>
<td>AACCCAAATTCCATACAGCTTCTACCCCTCAGTAAAAA</td>
<td>4292</td>
</tr>
<tr>
<td>AACCCAAATTCCATACAGCTTCTACCCCTCAGTAAAAA</td>
<td>4293</td>
</tr>
<tr>
<td>AACCCAAATTCCATACAGCTTCTACCCCTCAGTAAAAA</td>
<td>4294</td>
</tr>
<tr>
<td>AACCCAAATTCCATACAGCTTCTACCCCTCAGTAAAAA</td>
<td>4295</td>
</tr>
<tr>
<td>AACCCAAATTCCATACAGCTTCTACCCCTCAGTAAAAA</td>
<td>4296</td>
</tr>
<tr>
<td>AACCCAAATTCCATACAGCTTCTACCCCTCAGTAAAAA</td>
<td>4297</td>
</tr>
<tr>
<td>AACCCAAATTCCATACAGCTTCTACCCCTCAGTAAAAA</td>
<td>4298</td>
</tr>
<tr>
<td>AACCCAAATTCCATACAGCTTCTACCCCTCAGTAAAAA</td>
<td>4299</td>
</tr>
<tr>
<td>AACCCAAATTCCATACAGCTTCTACCCCTCAGTAAAAA</td>
<td>4300</td>
</tr>
<tr>
<td>AACCCAAATTCCATACAGCTTCTACCCCTCAGTAAAAA</td>
<td>4301</td>
</tr>
<tr>
<td>AACCCAAATTCCATACAGCTTCTACCCCTCAGTAAAAA</td>
<td>4302</td>
</tr>
<tr>
<td>AACCCAAATTCCATACAGCTTCTACCCCTCAGTAAAAA</td>
<td>4303</td>
</tr>
<tr>
<td>AACCCAAATTCCATACAGCTTCTACCCCTCAGTAAAAA</td>
<td>4304</td>
</tr>
<tr>
<td>AACCCAAATTCCATACAGCTTCTACCCCTCAGTAAAAA</td>
<td>4305</td>
</tr>
<tr>
<td>AACCCAAATTCCATACAGCTTCTACCCCTCAGTAAAAA</td>
<td>4306</td>
</tr>
<tr>
<td>AACCCAAATTCCATACAGCTTCTACCCCTCAGTAAAAA</td>
<td>4307</td>
</tr>
<tr>
<td>AACCCAAATTCCATACAGCTTCTACCCCTCAGTAAAAA</td>
<td>4308</td>
</tr>
<tr>
<td>AACCCAAATTCCATACAGCTTCTACCCCTCAGTAAAAA</td>
<td>4309</td>
</tr>
<tr>
<td>AACCCAAATTCCATACAGCTTCTACCCCTCAGTAAAAA</td>
<td>4310</td>
</tr>
<tr>
<td>AACCCAAATTCCATACAGCTTCTACCCCTCAGTAAAAA</td>
<td>4311</td>
</tr>
<tr>
<td>AACCCAAATTCCATACAGCTTCTACCCCTCAGTAAAAA</td>
<td>4312</td>
</tr>
<tr>
<td>AACCCAAATTCCATACAGCTTCTACCCCTCAGTAAAAA</td>
<td>4313</td>
</tr>
<tr>
<td>AACCCAAATTCCATACAGCTTCTACCCCTCAGTAAAAA</td>
<td>4314</td>
</tr>
<tr>
<td>AACCCAAATTCCATACAGCTTCTACCCCTCAGTAAAAA</td>
<td>4315</td>
</tr>
<tr>
<td>AACCCAAATTCCATACAGCTTCTACCCCTCAGTAAAAA</td>
<td>4316</td>
</tr>
<tr>
<td>AACCCAAATTCCATACAGCTTCTACCCCTCAGTAAAAA</td>
<td>4317</td>
</tr>
<tr>
<td>AACCCAAATTCCATACAGCTTCTACCCCTCAGTAAAAA</td>
<td>4318</td>
</tr>
<tr>
<td>AACCCAAATTCCATACAGCTTCTACCCCTCAGTAAAAA</td>
<td>4319</td>
</tr>
<tr>
<td>AACCCAAATTCCATACAGCTTCTACCCCTCAGTAAAAA</td>
<td>4320</td>
</tr>
<tr>
<td>AACCCAAATTCCATACAGCTTCTACCCCTCAGTAAAAA</td>
<td>4321</td>
</tr>
<tr>
<td>AACCCAAATTCCATACAGCTTCTACCCCTCAGTAAAAA</td>
<td>4322</td>
</tr>
<tr>
<td>AACCCAAATTCCATACAGCTTCTACCCCTCAGTAAAAA</td>
<td>4323</td>
</tr>
<tr>
<td>AACCCAAATTCCATACAGCTTCTACCCCTCAGTAAAAA</td>
<td>4324</td>
</tr>
<tr>
<td>AACCCAAATTCCATACAGCTTCTACCCCTCAGTAAAAA</td>
<td>4325</td>
</tr>
</tbody>
</table>

[001312] Specificity for Clq as a target was performed by immunoprecipitation of Clq depleted or spiked serum and plasma samples. SDS-PAGE analysis of the immunoprecipitation products demonstrated that bands corresponding to Clq subunits (i.e., A, B, C) were only present in the samples spiked with Clq. As a control, oligonucleotides comprising the reverse complements to SEQ ID NOs. 1472 and 1482 were synthesized. For example, the construct comprising the reverse complement to SEQ ID NOs. 1472 comprised the variable region sequence 5’ - TTAAGATAGTTGATTTAAGCGGTTGATTTCTATAGT (SEQ ID NO. 4326).

[001313] FIG. 24A shows a silver-stained reducing sodium dodecyl sulfate-polyacrylamide (SDS-PA) gel of immunoprecipitated proteins from plasma samples with the indicated oligonucleotides. In the figure, B1 comprises the variable region SEQ ID NO. 1472, B11 comprises the variable region SEQ ID NO. 1482, and L4 and L15 comprises variable regions without specific recognition of Clq. The individual biotinylated oligonucleotides were immobilized on streptavidin magnetic beads, contacted with PEG-
precipitated plasma isolates from a pool of healthy donors, and the beads were isolated. Proteins that immunoprecipitated with the beads are shown on the gel. Controls included the reverse complements of B1 and B1 1 (lanes B1rc and B1lrc, respectively) and a control using beads without oligonucleotides (Lane "NO"). Dashed arrows indicate pulled down proteins ClqA, ClqB, and ClqC. The dotted arrow indicates the heavy chain of IgM (IgMHC). The band labeled "SA" is streptavidin and the lane kDa is a protein ladder. As shown in FIG. 24A, specific bands indicated by the dashed arrows were observed in the B1 and B1l immunoprecipitation experiments. These bands were excised and analyzed by LC-MS/MS. This analysis revealed the complement protein complex Clq as the target of sequences B1 and B1l, and the figure indicates the positions of Clq subunits A, B and C. In contrast, reverse complement versions of B1 and B1l (i.e., B1rc, B1lrc) did not immunoprecipitate Clq. Also shown in the figure, IgM co-precipitated with Clq. Without being bound by theory, these findings are consistent with previous observations where IgM was found to remain complexed with Clq in sera that were precipitated with PEG. See Krieger, et al. Characterization of immune complexes detected by the 1251-Clq binding assay in breast cancer. Clin Immunol Immunopathol 46, 14-23 (1988); which publication is incorporated by reference herein in its entirety.

To further investigate the interaction of aptamer B1 and B1l with Clq, we performed a filter retention assay. Radioactively labelled B1 and B1l were incubated with increasing concentrations of purified Clq and subsequently passed through a nitrocellulose membrane. The amount of oligonucleotides retained on membrane-immobilized Clq was then quantified by autoradiography. Results are shown in FIG. 24B. These experiments demonstrate that B1 and B1l exhibit concentration dependent binding to Clq, whereas B1lrc did not. Similar results were obtained with B1l and B1lrc in an Enzyme-Linked Oligonucleotide Assay (ELONA; see US Patent 5789163, which patent is incorporated herein in its entirety). See FIG. 24C. The figure shows ELONA analysis of B1l (circles) at indicated concentrations and a fixed Clq concentration (0.625 nM). As a control, the reverse complement of B1l, B1lRC, was used (squares). "No aptamer" control (triangles) shows low background binding of detector Streptavidin-HRP. B1l specifically binds Clq (estimated KD around 40 nM). This binding behaviour indicates that B1 and B1l specifically interact with Clq with high affinity.

A second control comprised the reverse complement to the consensus motif (i.e., SEQ ID NO. 4151). This control comprised the variable region sequence 5' - TACTATAGATGGATTTGCTCCGCTTTAAATCAAATCTATCTTATA - 3' (SEQ ID NO. 4327) and did not immunoprecipitate Clq from plasma.

Table 45 lists proteins that were identified in the immunoprecipitation with B1 (i.e., SEQ ID NO. 1472) using PSMs (peptide spectrum matches) sorted by area. The proteins in the list were observed only in the immunoprecipitation with B1 (i.e., not observed with negative controls) or at levels at least 2-fold over negative controls. The mass spectrometry results were obtained from a single band that was run a short distance in a gel in order to clean up the sample. Accession numbers in the table are from the UniProt database (www.uniprot.org). "GN=" is followed by the gene name. As seen in Table 45, Clq subunits A-C were prominent in the results. IgM was also observed at high levels. Without being bound
by theory, this may be due to complexes between Clq and immunoglobulins as noted above. Complement C4b-binding protein alpha chain was also observed at 10-fold higher levels versus controls.

![Table 45: Oligonucleotide probe targets](image)

<table>
<thead>
<tr>
<th>Accession number</th>
<th>Protein name</th>
</tr>
</thead>
<tbody>
<tr>
<td>P02747</td>
<td>Complement Clq subcomponent subunit C GN=C1QC</td>
</tr>
<tr>
<td>P02746</td>
<td>Complement Clq subcomponent subunit B GN=C1QB</td>
</tr>
<tr>
<td>P01871</td>
<td>Ig mu chain C region GN=IGHM</td>
</tr>
<tr>
<td>P04220</td>
<td>Ig mu heavy chain disease protein</td>
</tr>
<tr>
<td>P02745</td>
<td>Complement Clq subcomponent subunit A GN=C1QA</td>
</tr>
<tr>
<td>P01834</td>
<td>Ig kappa chain C region GN=IGKC</td>
</tr>
<tr>
<td>P0CGG05</td>
<td>Ig lambda-2 chain C regions GN=IGLC2</td>
</tr>
<tr>
<td>B9A064</td>
<td>Immunoglobulin lambda-like polypeptide 5 GN=IGLL5</td>
</tr>
<tr>
<td>P01857</td>
<td>Ig gamma-1 chain C region GN=IGHG1</td>
</tr>
<tr>
<td>P01860</td>
<td>Ig gamma-3 chain C region GN=IGHG3</td>
</tr>
<tr>
<td>P01859</td>
<td>Ig gamma-2 chain C region GN=IGHG2</td>
</tr>
<tr>
<td>P01766</td>
<td>Ig heavy chain V-III region BRO</td>
</tr>
<tr>
<td>043866</td>
<td>CD5 antigen-like GN=CD5L</td>
</tr>
<tr>
<td>P01877</td>
<td>Ig alpha-2 chain C region GN=IGHA2</td>
</tr>
<tr>
<td>P01617</td>
<td>Ig kappa chain V-II region TEW</td>
</tr>
<tr>
<td>P01767</td>
<td>Ig heavy chain V-III region BUT</td>
</tr>
<tr>
<td>P06310</td>
<td>Ig kappa chain V-II region RPMI 6410</td>
</tr>
<tr>
<td>P01768</td>
<td>Ig heavy chain V-III region CAM</td>
</tr>
<tr>
<td>P01779</td>
<td>Ig heavy chain V-III region TUR</td>
</tr>
<tr>
<td>P01620</td>
<td>Ig kappa chain V-III region SIE</td>
</tr>
<tr>
<td>P01606</td>
<td>Ig kappa chain V-I region OU</td>
</tr>
<tr>
<td>P04208</td>
<td>Ig lambda chain V-I region WAI</td>
</tr>
<tr>
<td>P18135</td>
<td>Ig kappa chain V-III region HAI</td>
</tr>
<tr>
<td>P01825</td>
<td>Ig heavy chain V-II region NEWM</td>
</tr>
<tr>
<td>P01591</td>
<td>Immunoglobulin J chain GN=IGJ</td>
</tr>
<tr>
<td>P01598</td>
<td>Ig kappa chain V-I region EU</td>
</tr>
<tr>
<td>P01605</td>
<td>Ig kappa chain V-I region Lay</td>
</tr>
<tr>
<td>P04433</td>
<td>Ig kappa chain V-III region VG (Fragment)</td>
</tr>
<tr>
<td>P01610</td>
<td>Ig kappa chain V-I region WEA</td>
</tr>
<tr>
<td>Q8WZ74</td>
<td>Cortactin-binding protein 2 GN=CTTNBP2</td>
</tr>
<tr>
<td>P01781</td>
<td>Ig heavy chain V-III region GAL</td>
</tr>
<tr>
<td>P04003</td>
<td>C4b-binding protein alpha chain GN=C4BPA</td>
</tr>
<tr>
<td>P00739</td>
<td>Haptoglobin-related protein GN=HPR</td>
</tr>
<tr>
<td>P06312</td>
<td>Ig kappa chain V-IV region (Fragment) GN=IGKV4-1</td>
</tr>
<tr>
<td>P00738-2</td>
<td>Isoform 2 of Haptoglobin GN=HP</td>
</tr>
<tr>
<td>P01625</td>
<td>Ig kappa chain V-IV region Len</td>
</tr>
<tr>
<td>P01717</td>
<td>Ig lambda chain V-IV region Hil</td>
</tr>
<tr>
<td>Q04695</td>
<td>Keratin, type I cytoskeletal 17 GN=KRT17</td>
</tr>
<tr>
<td>Q96PE2</td>
<td>Rho guanine nucleotide exchange factor 17 GN=ARHGEF17</td>
</tr>
<tr>
<td>P04211</td>
<td>Ig lambda chain V region 4A</td>
</tr>
<tr>
<td>P05909</td>
<td>Apolipoprotein D GN=APOD</td>
</tr>
<tr>
<td>P34096</td>
<td>Ribonuclease 4 GN=RNASE4</td>
</tr>
<tr>
<td>P01714</td>
<td>Ig lambda chain V-III region SH</td>
</tr>
<tr>
<td>0 14490-3</td>
<td>Isoform 3 of Disks large-associated protein 1 GN=DLGAP1</td>
</tr>
<tr>
<td>P04196</td>
<td>Histidine-rich glycoprotein GN=HRG</td>
</tr>
<tr>
<td>P35030-5</td>
<td>Isoform 5 of Trypsin-3 GN=PRSS3</td>
</tr>
<tr>
<td>P04430</td>
<td>Ig kappa chain V-I region BAN</td>
</tr>
<tr>
<td>P00748</td>
<td>Coagulation factor XII GN=F12</td>
</tr>
<tr>
<td>P06727</td>
<td>Apolipoprotein A-IV GN=APOA4</td>
</tr>
</tbody>
</table>
[001317] The oligonucleotide sequences in Table 44 can be further modified to assess the effects of various subsequences and structural conformation on the ability to bind C1q. Exemplary rationally designed modifications are shown in Table 46. For example, the modifications can be based on the secondary structure of the oligonucleotides as estimated using mFold. The sequences in the table may comprise a modification such as a 5' Biotin motif. The Name column indicates the sequence and/or modifications thereof. For example, “83s B1 Full” (SEQ ID NO. 4328) consists of the 83s B1 sequence (SEQ ID NO. 1472) shown above with 5' and 3' regions (i.e., SEQ ID NOs. 131 and 132, respectively).
Beneath are truncations ("TrXX-YY," where XX is the start position of the truncation and YY is the end position of the truncation as compared to the parent), substitutions (e.g., T(23)G indicating that T at position 23 of the parent sequence is modified to G), and combinations thereof. "RC" indicates reverse complement sequence, in this case for use as a negative control. The italicized and underlined sequences highlight the 5’-ACAATCC consensus sequence.

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence</th>
<th>SEQ ID NO.</th>
</tr>
</thead>
<tbody>
<tr>
<td>83s B1 Full</td>
<td>CTAGCATGACTGCAGTACTGACTATAGACAAATCCACCCGCTTTAAATCTACGCTGCTCCGCAAGCA</td>
<td>4328</td>
</tr>
<tr>
<td>83s B1_Tr17-62</td>
<td>ACGTACTATAGACAAATCCACCCGCTTTAAATCTATCTGCTGCTCCGCAAGCA</td>
<td>4329</td>
</tr>
<tr>
<td>83s B1_Tr19-59</td>
<td>GTACTATAGACAAATCCACCCGCTTTAAATCTATCTGCTGCTCCGCAAGCA</td>
<td>4330</td>
</tr>
<tr>
<td>83s B1 T(23)G</td>
<td>CTAGCATGACTGCAGTACTGACTATAGACAAATCCACCCGCTTTAAATCTACGCTGCTCCGCAAGCA</td>
<td>4331</td>
</tr>
<tr>
<td>83s B1_Tr17-62T(23)G</td>
<td>ACGTACGATAGACAAATCCACCCGCTTTAAATCTATCTGCTGCTCCGCAAGCA</td>
<td>4332</td>
</tr>
<tr>
<td>83s B1 T(57)G</td>
<td>CTAGCATGACTGCAGTACTGACTATAGACAAATCCACCCGCTTTAAATCTATCTGCTGCTCCGCAAGCA</td>
<td>4333</td>
</tr>
<tr>
<td>83s B1_Tr24-76</td>
<td>GATAACCAAATCCACCCGCTTTAAATCTATCTGCTGCTCCGCAAGCA</td>
<td>4334</td>
</tr>
<tr>
<td>83s_B11 Full</td>
<td>CTAGCATGACTGCAGTACTGACTATAGACAAATCCACCCGCTTTAAATCTACGCTGCTCCGCAAGCA</td>
<td>4335</td>
</tr>
<tr>
<td>B11_C(47)T</td>
<td>CTAGCATGACTGCAGTACTGACTATAGACAAATCCACCCGCTTTAAATCTACGCTGCTCCGCAAGCA</td>
<td>4336</td>
</tr>
<tr>
<td>B11_Tr19-35</td>
<td>GATAACCAAATCCACCCGCTTTAAATCTATCTGCTGCTCCGCAAGCA</td>
<td>4337</td>
</tr>
<tr>
<td>B11 Tr_G18_19- 35_36C</td>
<td>GATAACCAAATCCACCCGCTTTAAATCTATCTGCTGCTCCGCAAGCA</td>
<td>4338</td>
</tr>
<tr>
<td>83s B1RC (Negative Control)</td>
<td>CTAGCATGACTGCAGTACTGACTATAGACAAATCCACCCGCTTTAAATCTACGCTGCTCCGCAAGCA</td>
<td>4339</td>
</tr>
<tr>
<td>83s B4 Full</td>
<td>CTAGCATGACTGCAGTACTGACTATAGACAAATCCACCCGCTTTAAATCTACGCTGCTCCGCAAGCA</td>
<td>4340</td>
</tr>
<tr>
<td>83s B14 Full</td>
<td>CTAGCATGACTGCAGTACTGACTATAGACAAATCCACCCGCTTTAAATCTACGCTGCTCCGCAAGCA</td>
<td>4341</td>
</tr>
<tr>
<td>83s_B6 Full</td>
<td>CTAGCATGACTGCAGTACTGACTATAGACAAATCCACCCGCTTTAAATCTACGCTGCTCCGCAAGCA</td>
<td>4342</td>
</tr>
</tbody>
</table>

In further analysis, Clq was immunoprecipitated from female plasma samples using the 83s B1 aptamer (or its negative control). For sequences, see Table 46 above. Clq was eluted with TFA/Urea +/- acetonitrile then detected by either an in-solution trypsin digestion followed by mass spectrometry or SDS-PAGE stained for total protein as in FIG. 24A (gel not shown). Mass Spectrometry was performed using the Q Exactive HF Mass Spectrometer system from ThermoFisher Scientific (Waltham, MA) according to the manufacturer's instructions. Various post-translational modifications (PTM) have been observed for Clq A. For example, the UniProt database lists various oxidation sites (e.g., P45, K48, P54, P57, P73, P79, P85) and galactosyl sites (e.g., K48, K67). Other modifications were reported in Pfieger et al, Analysis of Human Clq by Combined Bottom-up and Top-down Mass
Spectrometry, Mol Cell Proteomics. 2010 Apr; 9(4): 593-610, which publication is incorporated by reference herein in its entirety. **Table 47** lists PTMs observed in this study. The modifications in bold italics are previously unreported to our knowledge.

**Table 47: Observed Clq A post-translational modifications**

<table>
<thead>
<tr>
<th>Position</th>
<th>Target</th>
<th>Modification</th>
<th>Classification</th>
<th>Sequence Motif</th>
<th>SEQ ID NO.</th>
</tr>
</thead>
<tbody>
<tr>
<td>45</td>
<td>P</td>
<td>Oxidation</td>
<td>Post-translational</td>
<td>PGRGGRpGLKGEQ</td>
<td>4343</td>
</tr>
<tr>
<td>48</td>
<td>K</td>
<td>Galactosyl</td>
<td>O-linked glycosylation</td>
<td>RGRFGGLKGGGP</td>
<td>4344</td>
</tr>
<tr>
<td>48</td>
<td>K</td>
<td>Oxidation</td>
<td>Post-translational</td>
<td>RGRFGGLKGGGP</td>
<td>4345</td>
</tr>
<tr>
<td>54</td>
<td>P</td>
<td>Oxidation</td>
<td>Post-translational</td>
<td>KEQGEpGAPGIR</td>
<td>4346</td>
</tr>
<tr>
<td>57</td>
<td>P</td>
<td>Oxidation</td>
<td>Post-translational</td>
<td>QGEpQApGIRTG</td>
<td>4346</td>
</tr>
<tr>
<td>67</td>
<td>K</td>
<td>Galactosyl</td>
<td>O-linked glycosylation</td>
<td>TGQGLGKGDQGP</td>
<td>4347</td>
</tr>
<tr>
<td>73</td>
<td>P</td>
<td>Oxidation</td>
<td>Post-translational</td>
<td>KGQGEpGNSNP</td>
<td>4348</td>
</tr>
<tr>
<td>75</td>
<td>P</td>
<td>Oxidation</td>
<td>Post-translational</td>
<td>DQGEPGPpNSNP</td>
<td>4349</td>
</tr>
<tr>
<td>79</td>
<td>P</td>
<td>Oxidation</td>
<td>Post-translational</td>
<td>PGEPGpNSNP</td>
<td>4350</td>
</tr>
<tr>
<td>81</td>
<td>K</td>
<td>Galactosyl</td>
<td>O-linked glycosylation</td>
<td>PGNPFGkGVyFG</td>
<td>4351</td>
</tr>
<tr>
<td>85</td>
<td>P</td>
<td>Oxidation</td>
<td>Post-translational</td>
<td>PGKVGYpGPSGL</td>
<td>4352</td>
</tr>
<tr>
<td>87</td>
<td>P</td>
<td>Oxidation</td>
<td>Post-translational</td>
<td>KGyQpGyPSGL</td>
<td>4353</td>
</tr>
<tr>
<td>106</td>
<td>P</td>
<td>Oxidation</td>
<td>Post-translational</td>
<td>KGyKGSpGyAcG</td>
<td>4354</td>
</tr>
<tr>
<td>126</td>
<td>M</td>
<td>Oxidation</td>
<td>Artefact</td>
<td>IRRNFpGyGyNV</td>
<td>4355</td>
</tr>
<tr>
<td>205</td>
<td>M</td>
<td>Oxidation</td>
<td>Artefact</td>
<td>QVVSyGyGyVL</td>
<td>4356</td>
</tr>
</tbody>
</table>

**Example 36: Clq immunoassay and isolation**

[001319] This Example illustrates Clq immunoassays using an anti-Clq oligonucleotide. A nucleic acid construct is synthesized comprising an anti-Clq oligonucleotide region, for example, as described in **Example 35** above. The anti-Clq construct comprises a 5’ biotin modification to facilitate specific recognition by a desired moiety attached to streptavidin.

[001320] A labeled anti-Clq aptamer construct is constructed. The anti-Clq construct is contacted with fluorescently labeled streptavidin such as a streptavidin - Alexa Fluor® 488 conjugate from Thermo Fisher Scientific, Catalog number: SI 1223. This creates a fluorescently labeled anti-Clq construct which is used to detect Clq in various immunoassay formats. In one scenario, a biological sample known or suspected to contain Clq is contacted with an ELISA plate. The plate is washed and contacted with the fluorescently labeled anti-Clq construct. The fluorescent signal is read from the wells in the plate, thereby providing an indication of the presence or amount of Clq in the biological sample. In another scenario, a biological sample is directly contacted with the fluorescently labeled anti-Clq construct. The contacted sample is subjected to flow cytometry to detect fluorescent particles of the size of microvesicles, thereby providing an indication of the presence or amount of microvesicle associated Clq in the biological sample. Alternate labels such as disclosed herein or known in the art can be used in such formats.

[001321] Various modifications of the above scenarios are performed. For example, the anti-Clq aptamer is directly labeled with Alexa Fluor during the oligonucleotide synthesis process.

[001322] An immobilized anti-Clq aptamer is constructed. In one scenario, the anti-Clq aptamer construct is contacted with streptavidin conjugated beads. The beads are contacted with a biological sample known or suspected to contain Clq. The beads are precipitated (e.g., by centrifugation or magnetism) and washed. Proteins that precipitate with the beads are analyzed, thereby providing an
indication of the presence or amount of Clq and associated proteins in the biological sample. In another scenario, the anti-Clq aptamer construct is contacted with streptavidin agarose resin, e.g., Pierce™ Streptavidin Agarose, Thermo Fisher Scientific Catalog number: 20347 or Pierce™ High Capacity Streptavidin Agarose Thermo Fisher Scientific Catalog number: 20357. The resins are placed in a spin column or chromatography column, respectively. The anti-Clq aptamer is contacted with the resin where it is bound by the streptavidin. A biological sample known or suspected to comprise Clq is allowed to pass through the resin. Clq and associated proteins in the biological sample are retained by the anti-Clq aptamer within the resin and are then analyzed after elution. In either scenario, if desired, the anti-Clq aptamer is contacted with the biological sample in solution and then the sample is contacted with the beads or resin. Clq and associated proteins are analyzed as above. This modification allows the Clq and aptamer to bind freely in solution prior to aptamer immobilization.

[001323] Various modifications of the above scenarios are performed. For example, the anti-Clq aptamer is directly conjugated to a bead or other desired surface.


Example 37: Clq Detection in Bodily Fluids

[001325] This Example describes using an anti-Clq oligonucleotide to detect Clq in bodily fluids. A bodily fluid such as blood or a derivative thereof, including without limitation sera or plasma, is obtained from a subject. An assay such as described in Example 36 is used to detect Clq in the bodily fluid. As desired, such detection may assist in the diagnosis, prognosis or theranosis of a disease or disorder. See, e.g., Examples 13-14 herein.

[001326] In one scenario, the anti-Clq aptamer is used to capture and quantitate immune complexes in clinical samples based on the ability of Clq to preferentially bind immune complexes versus monomeric immunoglobulins.

Example 38: Clq Inhibition

[001327] This Example describes using an anti-Clq oligonucleotide to inhibit Clq mediated apoptosis.

[001328] Deficiency of the host Clq inhibitor protein, encoded by the SERPING1 gene (also known as C1NH and others), is associated with hereditary angioedema ("hereditary angioneurotic edema"; "HAE"), which comprises swelling due to leakage of fluid from blood vessels into connective tissue. In addition to facial swelling and/or abdominal pain, Clq inhibitor deficiency also predisposes to autoimmune diseases, e.g., lupus erythematosus, and age related macular degeneration. Clq inhibitors
CINRYZE® (C1 esterase inhibitor [human]) and RUCONEST® (C1 esterase inhibitor [recombinant]) are approved by the US FDA for HAE symptoms. See, e.g., Bernstein, Hereditary angioedema: a current state-of-the-art review, VIII: current status of emerging therapies. Ann. Allergy Asthma Immunol. 100 (1 Suppl 2): S41-6 (January 2008). A pharmaceutical composition comprising an anti-C1q oligonucleotide of the invention is administered to an HAE patient in sufficient dosage to ameliorate symptoms of HAE.

As the complement cascade damages cells, C1q inhibitors have potential functionality for diseases other than HAE. See e.g., Caliezi et al, C1-Esterase inhibitor: an anti-inflammatory agent and its potential use in the treatment of diseases other than hereditary angioedema. Pharmacol. Rev. 52 (1): 91-112 (March 2000). During heart attack, lack of oxygen in heart cells causes necrosis in heart cells. Dying heart cells spill their contents in the extracellular environment, which triggers the complement cascade and may increase the damage for the surviving heart cells. A pharmaceutical composition comprising an anti-C1q oligonucleotide of the invention is administered to a heart attack victim in sufficient dosage to ameliorate complement-mediated heart damage. Similarly, a pharmaceutical composition comprising an anti-C1q oligonucleotide of the invention is administered to a patient following organ transplantation to ameliorate delayed graft function.

The complement cascade is a first defense against non-self cells and is also activated by malignant cells. Thus, complement appears to play a role in cancer surveillance by the immune system. Recently, however, it has been reported that the complement system can also play a role in promoting tumor growth, inflammation and angiogenesis. Thus, complement inhibition has been proposed for cancer treatment. See, e.g., Pio et al, Complement inhibition: a promising concept for cancer treatment, Semin Immunol. 2013 Feb; 25(1): 54-64; Pio et al., The role of complement in tumor growth. Adv Exp Med Biol. 2014;772:229-62. It has further been reported that C1q may act in the tumor microenvironment to promote cancer independently of its complement function. Bulla et al. report that C1q may promote cancer cell adhesion, migration and proliferation as well as angiogenesis and metastasis. See Bulla et al., C1q acts in the tumour microenvironment as a cancer-promoting factor independently of complement activation. Nat Commun. 2016; 7: 10346. A pharmaceutical composition comprising an anti-C1q oligonucleotide of the invention is administered to a cancer victim in sufficient dosage to treat the cancer in the victim. As desired, the composition is administered in combination with traditional cancer therapeutics or in addition to immune therapies, e.g., cell-based tumor immunotherapies, or antitumor vaccines.

Example 39: Complement cascade initiation

This Example describes using an anti-C1q oligonucleotide to direct complement mediated cell killing.

Antibody immunotherapies may direct the killing of target cells through several mechanisms, including without limitation complement-dependent cytotoxicity (CDC), antibody-dependent cellular cytotoxicity (ADCC), apoptosis, and direct growth arrest. See, e.g., Taylor and Lindorfer, The role of complement in mAb-based therapies of cancer. Methods. 2014 Jan 1;65(1): 18-27. doi: 10.1016/j.ymeth.2013.07.027. Epub 2013 Jul 22; Rogers et al., Complement in monoclonal antibody

[001333] The Fc regions of membrane-bound therapeutic antibodies interact with the heterooligomeric Clq complex and activate the classical complement pathway to initiate CDC. A multipartite construct comprising an anti-Clq oligonucleotide of the invention is provided in a pharmaceutical composition. The pharmaceutical composition is administered to a cancer victim in sufficient dosage to treat the cancer in the victim. The construct comprises the anti-Clq oligonucleotide connected to an anti-tumor domain. See, e.g., FIGs. 23A-D herein and related discussion. The anti-tumor domain can be to a general tumor antigen or specific to a tumor of a given origin. See, e.g., Section "Anti-target and multivalent oligonucleotides" herein for illustrative targets. As desired, the multipartite construct comprising an anti-Clq oligonucleotide is used to direct complement mediated cell killing where cancer cells and tumors have adapted to evade the complement system.

[001334] As noted herein, Clq may play a role in both preventing cancer, e.g., through complement mediated cell killing, and facilitating the growth of cancer, e.g., through cancer cell adhesion, migration and proliferation as well as angiogenesis and metastasis, perhaps in a non-complement fashion. Such disparate roles allow for combination treatment with oligonucleotides of the invention that both inhibit the roles of Clq in promoting cancer while targeting Clq mediated cell killing to cancer cells. A multipartite construct comprising an anti-Clq oligonucleotide of the invention is provided in a pharmaceutical composition. An anti-Clq oligonucleotide of the invention that inhibits non-complement functions of Clq is provided in a pharmaceutical composition. The pharmaceutical compositions are administered to a cancer victim in sufficient dosage to treat the cancer in the victim.

[001335] Although 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.
What is claimed is:

1. An oligonucleotide comprising a sequence selected from any one of SEQ ID NOs. 1472, 1475, 1477, 1482, and 4151-4325.

2. An oligonucleotide comprising a central sequence comprising a sequence according to any one of SEQ ID NOs. 1472, 1475, 1477, 1482, and 4151-4325, and a 5' region with sequence 5' - CTAGCATGACTGCAGTACGT (SEQ ID NO. 131), a 3' region with sequence 5' - CTGTCTCTTATACACATCT

3. The oligonucleotide of claim 1 or 2, wherein the oligonucleotide is capable of binding to complement protein complex Clq or a subunit thereof.

4. An oligonucleotide comprising the sequence 5' - ACAATCC (SEQ ID NO. 4151), wherein the oligonucleotide is capable of binding to complement protein complex Clq or a subunit thereof.

5. An oligonucleotide comprising the sequence 5' - ACTATAGAACAATCCACCGCTTTAAATCAACTATCTTA (SEQ ID NO. 1472) or 5' - AACCCCAATTCCACATACACTCTCACCCTTCCACTAAACA (SEQ ID NO. 1482), wherein the oligonucleotide is capable of binding to complement protein complex Clq or a subunit thereof.

6. The oligonucleotide of claim 5, wherein the Clq subunit is the A, B or C subunit.

7. The oligonucleotide of claim 6, wherein the Clq A subunit has at least one modification selected from Table 46.

8. An oligonucleotide comprising a nucleic acid sequence or a portion thereof that is at least 50, 55, 60, 65, 70, 75, 80, 85, 86, 86, 88, 89, 90, 95, 96, 97, 98, 99 or 100 percent homologous to an oligonucleotide sequence according to any preceding claim.

9. A plurality of oligonucleotides comprising at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130, 140, 150, 160, 170 or all oligonucleotide sequences according to any preceding claim.

10. A plurality of oligonucleotides comprising at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, or at least 1000 different oligonucleotide sequences according to claim 3 or 4.

11. The oligonucleotide or the plurality of oligonucleotides according to any preceding claim, wherein the oligonucleotide or the plurality of oligonucleotides comprises at least one of DNA, RNA, PNA, and any combination thereof.
12. The oligonucleotide or the plurality of oligonucleotides according to any preceding claim, wherein the oligonucleotide or the plurality of oligonucleotides comprises at least one functional modification selected from the group consisting of biotinylation, a non-naturally occurring nucleotide, a deletion, an insertion, an addition, and a chemical modification.

13. The oligonucleotide or plurality of oligonucleotides according to claim 12, wherein the chemical modification comprises at least one of C18, polyethylene glycol (PEG), PEG4, PEG6, PEG8, and PEG12.

14. A method comprising synthesizing the at least one oligonucleotide or the plurality of oligonucleotides according to any preceding claim.

15. A method comprising contacting a biological sample with the at least one oligonucleotide or the plurality of oligonucleotides according to any one of claims 1-13.

16. The method of claim 15, further comprising detecting a presence or level of complement protein complex C1q or a subunit thereof in the biological sample that is bound by the at least one oligonucleotide or at least one member of the plurality of oligonucleotides.

17. The method of claim 16, wherein the detecting comprises at least one of sequencing, amplification, hybridization, gel electrophoresis, chromatography, and any combination thereof.

18. The method of claim 17, wherein the sequencing comprises at least one of next generation sequencing, dye termination sequencing, pyrosequencing, and any combination thereof.

19. The method of claim 16, wherein the detecting comprises using at least one of an immunoassay, enzyme immunoassay (EIA), enzyme-linked immunoassorbent assay (ELISA), enzyme-linked oligonucleotide assay (ELONA), affinity isolation, immunoprecipitation, Western blot, gel electrophoresis, or flow cytometry.

20. The method of any one of claims 16-19, wherein the bound complement protein complex C1q or subunit thereof is associated with a microvesicle population.

21. The method of claim 20, further comprising isolating the microvesicle population prior to the contacting, after the contacting, or both.

22. The method of claim 21, wherein the isolating comprises affinity purification, filtration, concentration, polymer precipitation, PEG precipitation, ultracentrifugation, a molecular crowding reagent, affinity selection, chromatography, or any combination thereof.

23. The method of any one of claims 15-22, wherein the biological sample comprises a bodily fluid, tissue sample or cell culture.

24. The method of claim 23, wherein the bodily fluid comprises peripheral blood, sera, plasma, ascites, urine, cerebrospinal fluid (CSF), sputum, saliva, bone marrow, synovial fluid, aqueous
h humor, amniotic fluid, cerumen, breast milk, bronchoalveolar lavage fluid, semen, prostatic fluid, cowper’s fluid or pre-ejaculatory fluid, female ejaculate, sweat, fecal matter, hair oil, 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, blastocyl cavity fluid, or umbilical cord blood.

25. The method of claim 23, wherein the bodily fluid comprises whole blood, serum or plasma.

26. The method of any one of claims 15-22, wherein the biological sample comprises purified Clq, recombinant Clq, both, or subunits or fragments thereof.

27. The method of any one of claims 23-26, wherein the biological sample further comprises purified Clq, recombinant Clq, both, or fragments thereof.

28. The method of any one of claims 16-27, wherein the presence or level is used to characterize a phenotype.

29. The method of claim 28, wherein the phenotype is a disease or disorder.

30. The method of claim 29, wherein the characterizing comprises providing, or assisting in providing, at least one of diagnostic, prognostic and theranostic information for the disease or disorder.

31. The method of any one of claims 28-30, wherein the characterizing comprises comparing the presence or level to a reference.

32. The method of claim 31, wherein the reference comprises the presence or level determined in a sample from at least one individual without the phenotype or from at least one individual with a different phenotype.

33. The method any one of claims 28-32, wherein the sample is from a subject suspected of having or being predisposed to the disease or disorder.

34. The method of any of claims 28-33, wherein the disease or disorder comprises a cancer, a premalignant condition, an inflammatory disease, an immune disease, an autoimmune disease or disorder, a cardiovascular disease or disorder, neurological disease or disorder, infectious disease or pain.

35. The method of claim 34, wherein the cancer comprises an acute lymphoblastic leukemia; acute myeloid leukemia; adrenocortical carcinoma; AIDS-related cancers; AIDS-related lymphoma; anal cancer; appendix cancer; astrocytomomas; 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, astrocytomomas, 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; lung 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; Sezary 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; thyroid 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; Waldenstrom macroglobulinemia; or Wilm's tumor.

36. The method of claim 34, wherein the premalignant condition comprises Barrett's Esophagus.

37. The method of claim 34, wherein the autoimmune disease comprises 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.

-297-
38. The method of claim 34, wherein the cardiovascular disease comprises atherosclerosis, congestive heart failure, vulnerable plaque, stroke, ischemia, high blood pressure, stenosis, vessel occlusion or a thrombotic event.

39. The method of claim 34, wherein the neurological disease comprises Multiple Sclerosis (MS), Parkinson's Disease (PD), Alzheimer's Disease (AD), schizophrenia, bipolar disorder, depression, autism, Prion Disease, Pick's disease, dementia, Huntington disease (HD), Down's syndrome, cerebrovascular disease, Rasmussen's encephalitis, viral meningitis, neuropsychiatric systemic lupus erythematosus (NPSLE), amyotrophic lateral sclerosis, Creutzfeldt-Jacob disease, Gerstmann-Straussler-Scheinker disease, transmissible spongiform encephalopathy, ischemic reperfusion damage (e.g. stroke), brain trauma, microbial infection, or chronic fatigue syndrome.

40. The method of claim 34, wherein the pain comprises fibromyalgia, chronic neuropathic pain, or peripheral neuropathic pain.

41. The method of claim 34, wherein the infectious disease comprises a bacterial infection, viral infection, yeast infection, Whipple's Disease, Prion Disease, cirrhosis, methicillin-resistant staphylococcus aureus, HIV, HCV, hepatitis, syphilis, meningitis, malaria, tuberculosis, influenza.

42. A kit comprising a reagent for carrying out the method of any of claims 14-41.

43. Use of a reagent for carrying out the method of any of claims 14-41.

44. The kit of claim 42 or use of claim 43, wherein the reagent comprises the at least one oligonucleotide or the plurality of oligonucleotides, one or more primer for amplification or sequencing of such oligonucleotides, at least one binding agent to at least one protein, a binding buffer with or without MgCl₂, a sample processing reagent, a microvesicle isolation reagent, a detection reagent, a secondary detection reagent, a wash buffer, an elution buffer, a solid support, and any combination thereof.

45. The kit or use of claim 44, wherein the microvesicle isolation reagent comprises at least one of a concentrator unit, a filtration unit, a polymer, PEG, a size exclusion column, a binding agent to a microvesicle antigen, and any combination thereof; and/or the detection or secondary detection agent comprises streptavidin-horse radish peroxide (HRP), a streptavidin-conjugated fluorophore, a streptavidin-conjugated quantum dot, and any combination thereof.

46. A pharmaceutical composition comprising a therapeutically effective amount of the at least one oligonucleotide or the plurality of oligonucleotides according to any one of claims 1-13, or a salt thereof, and a pharmaceutically acceptable carrier, diluent, or both.

47. A method of treating or ameliorating a disease or disorder in a subject in need thereof, comprising administering the composition of claim 46 to the subject.

48. A method of inhibiting Clq in a subject, comprising administering the composition of claim 46 to the subject.
49. The method of claim 47 or 48, wherein the administering comprises at least one of intradermal, intramuscular, intraperitoneal, intravenous, subcutaneous, intranasal, epidural, oral, sublingual, intracerebral, intravaginal, transdermal, rectal, by inhalation, topical administration, or any combination thereof.

50. A multipartite construct that comprises a first segment that binds to a first target and a second segment that binds to a second target, wherein the first segment, the second segment, or both, comprises a sequence selected from any one of SEQ ID NOs. 137-969 and 1072-4325, or a region that is at least 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 96, 97, 98, 99 or 100 percent homologous thereto.

51. The multipartite construct of claim 50, wherein the first target comprises a cell-of-origin or disease specific protein marker.

52. The multipartite construct of any one of claims 50-51, wherein the first target comprises a cancer antigen, a tumor antigen, a cell surface antigen, a microvesicle surface antigen, a cell receptor and/or a membrane antigen.

53. The multipartite construct of any one of claims 50-52, wherein the first target is selected from any one of Table 3, Table 4, Tables 18-26, or Table 45.

54. The multipartite construct of any one of claims 50-52, wherein the first target is selected from the group consisting of CD2, CD4, CD19, CD20, CD22, CD23, CD30, CD33, CD37, CD40, CD44v6, CD52, CD56, CD70, CD74, CD79a, CD80, CD98, CD138, EGFR (Epidermal growth factor receptor), VEGF (Vascular endothelial growth factor), VEGFR1 (Vascular endothelial growth factor receptor 1), PDGFR (Platelet-derived growth factor receptor), RANKL (Receptor activator of nuclear factor kappa-B ligand), GPNMB (Transmembrane glycoprotein Neuromedin B), EphA2 (Ephrin type-A receptor 2), PSMA (Prostate-specific membrane antigen), Cripto (Cryptic family protein IB), EpCAM (Epithelial cell adhesion molecule), CTLA4 (Cytotoxic T-Lymphocyte Antigen 4), IGF-IR (Type 1 insulin-like growth factor receptor), GP3 (M13 bacteriophage), GP9 (Glycoprotein IX), CD42a, GP 40 (Glycoprotein 40kDa), GPC3 (glypican-3), GPC1 (glypican-1), TRAILRI (Tumor necrosis factor-related apoptosis-inducing ligand receptor 1), TRAILRII (Tumor necrosis factor-related apoptosis-inducing ligand receptor II), FAS (Type II transmembrane protein), PS (phosphatidyl serine) lipid, Mucl (Mucin 1, PEM), Mucl8, CD146, αβ1 integrin, α4β1 integrin, αv integrin (Vitronectin Receptor), Chondrolectin, CAIX (Carbonic anhydrase IX), GD2 gangloside, GD3 gangloside, GM1 gangloside, Lewis Y antigen, Mesothelin, HER2 (Human Epidermal Growth factor 2), HER3, HER4, FN14 (Fibroblast Growth Factor Inducible 14), CS1 (Cell surface glycoprotein, CD2 subset 1, CRACC, SLAMF7, CD319), 41BB CD137, SIP (Siah-1 Interacting Protein), CTGF (Connective tissue growth factor), HLADR (MHC class II cell surface receptor), PD-1 (Programmed Death 1, Type I membrane protein, PD-L1 (Programmed Death Ligand 1), PD-L2 (Programmed Death Ligand 2), IL-2 (Interleukin-2), IL-8 (Interleukin-8), IL-13 (Interleukin-13), PIGF (Phosphatidylinositol-glycan biosynthesis class F protein), NRPI (Neuropilin-1), ICAM1, CD54, GC182 (Claudin 18.2), Claudin, HGF (Hepatocyte growth factor), CEA
(Carcinoembryonic antigen), LTβR (lymphotoxin β receptor), Kappa Myeloma, Folate Receptor alpha, GRP78 (BIP, 78 kDa Glucose-regulated protein), A33 antigen, PSA (Prostate-specific antigen), CA 125 (Cancer antigen 125 or carbohydrate antigen 125), CA19.9, CA15.3, CA242, leptin, prolactin, osteopontin, IGF- II (Insulin-like growth factor 2), fascin, sPlgR (secreted chain of polymorphic immunoglobulin receptor), 14-3-3 eta protein, 5T4, ETA (epithelial tumor antigen), MAGE (Melanoma-associated antigen), MAPG (Melanoma-associated proteoglycan, NG2), vimentin, EPCA-1 (Early prostate cancer antigen-2), TAG-72 (Tumor-associated glycoprotein 72), factor VIII, Neprilysin (Membrane metallo-endopeptidase), 17-1 A (Epithelial cell surface antigen 17-1A), nucleolin, nucleophosmin, and any combination thereof.

55. The multipartite construct of claim 50, wherein the first target comprises CD20.

56. The multipartite construct of claim 50, wherein the first target comprises cMet.

57. The multipartite construct of any of claims 50-56, wherein the first segment comprises an oligonucleotide.

58. The multipartite construct of claim 57, further comprising a first oligonucleotide primer region and/or a second oligonucleotide primer region surrounding the first segment.

59. The multipartite construct of claim 50, wherein the second target comprises an immunomodulatory molecule.

60. The multipartite construct of claim 50, wherein the second target comprises at least one of a member of the innate immune system, a member of the complement system, Clq, Clr, C1s, CI, C3a, C3b, C3d, C5a, C2, C4, and any combination thereof.

61. The multipartite construct of claim 50, wherein the second target comprises Clq or a subunit thereof.

62. The multipartite construct of claim 62, wherein the C1q subunit is the A, B or C subunit.

63. The multipartite construct of claim 61, wherein the A subunit has at least one modification selected from Table 46.

64. The multipartite construct of any of claims 50-63, wherein the second segment comprises an antibody or oligonucleotide.

65. The multipartite construct of claim 64, further comprising a first oligonucleotide primer region and/or a second oligonucleotide primer region surrounding the second segment.

66. The multipartite construct of any of claims 50-65, wherein the second segment comprises an oligonucleotide according to any one of claims 1-8.
67. The multipartite construct of claim 50, further comprising a linker region between the first segment and second segment.

68. The multipartite construct of claim 67, wherein the linker region comprises an immunostimulatory sequence and/or an anti-proliferative or pro-apoptotic sequence.

69. The multipartite construct of claim 67, wherein the linker region comprises one or more CpG motif.

70. The multipartite construct of claim 67, wherein the linker region comprises a polyG sequence.

71. The multipartite construct of claim 50, wherein the multipartite construct is further modified to comprise at least one oligonucleotide chemical modification.

72. The multipartite construct of claim 71, wherein the modification is selected from the group consisting of: a chemical substitution at a sugar position; a chemical substitution at a phosphate position; and a chemical substitution at a base position of the nucleic acid.

73. The multipartite construct of claim 71, wherein the modification is selected from the group consisting of: incorporation of a modified nucleotide, 3’ capping, conjugation to an amine linker, conjugation to a high molecular weight, non-immunogenic compound, conjugation to a lipophilic compound, conjugation to a drug, conjugation to a cytotoxic moiety and labeling with a radioisotope.

74. The multipartite construct of claim 73, wherein the non-immunogenic, high molecular weight compound is polyalkylene glycol.

75. The multipartite construct of claim 74, wherein the polyalkylene glycol is polyethylene glycol.

76. The multipartite construct of claim 50, wherein the multipartite construct further comprises an immunostimulating moiety and/or a membrane disruptive moiety.

77. The multipartite construct of any of claims 50-76, wherein the multipartite construct comprises an oligonucleotide polymer, and optionally wherein the multipartite construct is flanked by a first oligonucleotide primer region and a second oligonucleotide primer region.

78. The multipartite construct of any of claims 50-77, wherein the first target comprises CD20 or cMET and the second segment comprises an oligonucleotide according to any one of claims 1-8.

79. A pharmaceutical composition comprising a therapeutically effective amount of the multipartite construct of any of claims 50-78, or a salt thereof, and a pharmaceutically acceptable carrier or diluent.
80. A method of treating or ameliorating a disease or disorder, comprising administering the composition of claim 79 to a subject in need thereof.

81. The method of claim 80, wherein the administering comprises at least one of intradermal, intramuscular, intraperitoneal, intravenous, subcutaneous, intranasal, epidural, oral, sublingual, intracerebral, intravaginal, transdermal, rectal, by inhalation, topical administration, or any combination thereof.

82. The method of claim 80, wherein the first target of the multipartite construct is determined by: a) contacting a biological sample from the subject with an oligonucleotide probe library; b) identifying members of the probe library that bound to the biological sample upon contact in a); and c) determining the first target as a binding target of a member of the probe library identified in b), thereby providing a personalized medical treatment to the subject.

83. The method of claim 82, wherein the biological sample comprises a biopsy sample, a tissue removed during a surgical procedure, a bodily fluid, a frozen section, a cell culture, or a fraction, derivative or components of any thereof.

84. The method of claim 83, wherein the bodily fluid comprises peripheral blood, sera, plasma, ascites, urine, cerebrospinal fluid (CSF), sputum, saliva, bone marrow, synovial fluid, aqueous humor, amniotic fluid, cerumen, breast milk, bronchoalveolar lavage fluid, semen (including prostatic fluid), Cowper's fluid or pre-ejaculatory fluid, female ejaculate, sweat, fecal matter, tears, cyst fluid, pleural fluid, 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.

85. A method of inducing killing of a cell, comprising contacting the cell with a multipartite construct of any of claims 50-78.

86. The method of claim 85, wherein the second segment comprises an oligonucleotide according to any one of claims 1-8.

87. The method of claim 85, wherein the cell comprises a disease or disorder.

88. The method of any one of claims 47-49, 80-84, or 87, wherein the disease or disorder comprises a cancer, a premalignant condition, an inflammatory disease, an immune disease, an autoimmune disease or disorder, a cardiovascular disease or disorder, a neurological disease or disorder, infectious disease or pain.

89. The method of claim 88, 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; lung 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; Sezary 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; Waldenstrom macroglobulinemia; or Wilm's tumor.

90. The method of claim 88, wherein the premalignant condition comprises Barrett's Esophagus or a colorectal polyp.

91. The method of claim 88, wherein the autoimmune disease comprises 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.

92. The method of claim 88, wherein the cardiovascular disease comprises atherosclerosis, congestive heart failure, vulnerable plaque, stroke, ischemia, high blood pressure, stenosis, vessel occlusion or a thrombotic event.

93. The method of claim 88, wherein the neurological disease comprises Multiple Sclerosis (MS), Parkinson's Disease (PD), Alzheimer's Disease (AD), schizophrenia, bipolar disorder, depression, autism, Prion Disease, Pick's disease, dementia, Huntington disease (HD), Down's syndrome, cerebrovascular disease, Rasmussen's encephalitis, viral meningitis, neurospsychiatric systemic lupus erythematosus (NPSLE), amyotrophic lateral sclerosis, Creutzfeldt-Jacob disease, Gerstmann-Straussler-Scheinker disease, transmissible spongiform encephalopathy, ischemic reperfusion damage (e.g. stroke), brain trauma, microbial infection, or chronic fatigue syndrome.

94. The method of claim 88, wherein the pain comprises fibromyalgia, chronic neuropathic pain, or peripheral neuropathic pain.

95. The method of claim 88, wherein the infectious disease comprises a bacterial infection, viral infection, yeast infection, Whipple's Disease, Prion Disease, cirrhosis, methicillin-resistant staphylococcus aureus, HIV, HCV, hepatitis, syphilis, meningitis, malaria, tuberculosis, or influenza.

96. A kit comprising a multipartite construct of any of claims 50-78, or a pharmaceutical composition of claim 79.

97. A kit comprising a reagent for carrying out the method of any of claims 80-95.

98. Use of a reagent for carrying out the method of any of claims 80-95.

99. Use of a reagent for the manufacture of a kit or reagent for carrying out the method of any of claims 80-95.

100. Use of a reagent for the manufacture of a medicament for carrying out the method of any of claims 80-95.

101. The kit of claim 97 or use of any of claims 98-100, wherein the reagent comprises a multipartite construct of any of claims 50-78, or a pharmaceutical composition of claim 79.

102. A method of detecting at least one microvesicle in a biological sample comprising contacting the biological sample with at least one binding agent to at least one protein in any one of Tables 18, 26 or 45 and detecting the at least one microvesicle recognized by the binding agent to the at least one protein.
103. The method of claim 102, wherein the at least one protein comprises a complement protein or an immunoglobulin.

104. The method of claim 103, wherein the complement protein comprises C1q, a subunit thereof (A, B or C), C4b, or any combination thereof.

105. The method of claim 103, wherein the immunoglobulin comprises IgM, IgG, or both.

106. The method of claim 102, wherein the at least one binding agent comprises a nucleic acid, DNA molecule, RNA molecule, antibody, antibody fragment, aptamer, peptoid, zDNA, peptide nucleic acid (PNA), locked nucleic acid (LNA), lectin, peptide, dendrimer, membrane protein labeling agent, chemical compound, or a combination thereof.

107. The method of claim 102, wherein the at least one binding agent comprises at least one oligonucleotide.

108. The method of any one of claims 102-107, wherein the at least one binding agent is used to capture and/or detect the at least one microvesicle.

109. The method of claim 108, wherein the at least one binding agent used to capture the at least one microvesicle is bound to a substrate.

110. The method of claim 109, wherein the substrate comprises a planar array, a column matrix, or a microbead.

111. The method of claim 108, wherein the at least one binding agent used to detect the at least one microvesicle is labeled.

112. The method of claim 111, wherein the label comprises at least one of a magnetic label, a fluorescent moiety, an enzyme, a chemilumine scent probe, a metal particle, a non-metal colloidal particle, a polymeric dye particle, a pigment molecule, a pigment particle, an electrochemically active species, a semiconductor nanocrystal, a nanoparticle, a quantum dot, a gold particle, a fluorophore, or a radioactive label.

113. The method of any one of claims 102-112, wherein the detected at least one microvesicle is found at higher levels in a healthy sample as compared to a diseased sample.

114. The method of any one of claims 102-112, wherein the detected at least one microvesicle is found at higher levels in a diseased sample as compared to a healthy sample.

115. The method of any one of claims 102-114, wherein the detecting is used to characterize or assist in characterizing a phenotype.

116. The method of claim 115, wherein the phenotype is a disease or disorder.
17. The method of claim 116, wherein the characterizing comprises providing or assisting in providing diagnostic, prognostic and/or theranostic information for the disease or disorder.

18. The method of any one of claims 115-117, wherein the characterizing comprises comparing a presence or level of the at least one microvesicle to a reference.

19. The method of claim 118, wherein the reference comprises a presence or level of the at least one microvesicle in a sample from an individual without the same phenotype.

20. The method of any one of claims 102-119, wherein the biological sample comprises a bodily fluid, tissue sample or cell culture.

21. The method of claim 115, wherein the bodily fluid or tissue sample are from a subject having or suspected of having a medical condition, a disease or a disorder.

22. The method of claim 115 or 121, wherein the bodily fluid comprises peripheral blood, sera, plasma, ascites, urine, cerebrospinal fluid (CSF), sputum, saliva, bone marrow, synovial fluid, aqueous humor, amniotic fluid, cerumen, breast milk, bronchoalveolar lavage fluid, semen, prostatic fluid, cowper's fluid or pre-ejaculatory fluid, female ejaculate, sweat, fecal matter, hair oil, 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, blastocyl cavity fluid, or umbilical cord blood.

23. The method of any one of claims 102-122, wherein the biological sample is from a subject suspected of having or being predisposed to a disease or disorder.

24. The method of any one of claims 116-119, or 123, wherein the disease or disorder comprises a cancer, a premalignant condition, an inflammatory disease, an immune disease, an autoimmune disease or disorder, a cardiovascular disease or disorder, neurological disease or disorder, infectious disease or pain.

25. The method of claim 124, 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; lung 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; penis 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; Sezary 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; Waldenstrom macroglobulinemia; or Wilm's tumor.

126. The method of claim 124, wherein the premalignant condition comprises Barrett's Esophagus.

127. The method of claim 124, wherein the autoimmune disease comprises 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 erythematosus (SLE), Hashimoto's Thyroiditis, Grave's disease, Ankylosing Spondylitis Sjogrens Disease, CREST syndrome, Scleroderma, Rheumatic Disease, organ rejection, Primary Sclerosing Cholangitis, or sepsis.

128. The method of claim 124, wherein the cardiovascular disease comprises atherosclerosis, congestive heart failure, vulnerable plaque, stroke, ischemia, high blood pressure, stenosis, vessel occlusion or a thrombotic event.
129. The method of claim 124, wherein the neurological disease comprises Multiple Sclerosis (MS), Parkinson's Disease (PD), Alzheimer's Disease (AD), schizophrenia, bipolar disorder, depression, autism, Prion Disease, Pick's disease, dementia, Huntington disease (HD), Down's syndrome, cerebrovascular disease, Rasmussen's encephalitis, viral meningitis, neuropsychiatric systemic lupus erythematosus (NPSLE), amyotrophic lateral sclerosis, Creutzfeldt-Jacob disease, Gerstmann-Straussler-Scheinker disease, transmissible spongiform encephalopathy, ischemic reperfusion damage (e.g. stroke), brain trauma, microbial infection, or chronic fatigue syndrome.

130. The method of claim 124, wherein the pain comprises fibromyalgia, chronic neuropathic pain, or peripheral neuropathic pain.

131. The method of claim 124, wherein the infectious disease comprises a bacterial infection, viral infection, yeast infection, Whipple's Disease, Prion Disease, cirrhosis, methicillin-resistant staphylococcus aureus, HIV, HCV, hepatitis, syphilis, meningitis, malaria, tuberculosis, influenza.

132. A kit comprising a reagent for carrying out the method of any of claims 102-131.

133. Use of a reagent for carrying out the method of any of claims 102-131.

134. The kit of claim 132 or use of claim 133, wherein the reagent comprises at least one binding agent to the at least one protein, a binding buffer with or without MgC\(^2\), a sample processing reagent, a microvesicle isolation reagent, a detection reagent, a secondary detection reagent, a wash buffer, an elution buffer, a solid support, and any combination thereof.

135. The kit or use of claim 134, wherein the microvesicle isolation reagent comprises at least one of a concentrator unit, a filtration unit, a polymer, PEG, a size exclusion column, a binding agent to a microvesicle antigen, and any combination thereof; and/or the detection or secondary detection agent comprises streptavidin-horse radish peroxide (HRP), a streptavidin-conjugated fluorophore, a streptavidin-conjugated quantum dot, and any combination thereof.
Selection strategy

Single round of binding with extensive washing “Rnd1 direct”

Additional round of binding, washing and competitive displacement with anti-EpCAM antibody “Rnd2 competitive”

FIG. 1
Screening Scheme

5 Detection Antibodies x 20 Capture Antibodies = 100 Combinations Screened

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

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
FIG. 4

1. Combine with exosomes from normal patients
2. Retain bound oligonucleotides and amplify
3. Separate oligonucleotides from cancer exosomes

Random oligonucleotide pool (1x10^6)
Cancer-specific oligonucleotide pool
CAR003 vs EPCAM antigen. Effect of denaturing vs MgCl2.

<table>
<thead>
<tr>
<th>Antigen, ug</th>
<th>0</th>
<th>1</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>75C, 5min, No MgCl2</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **Apt4 Control**
- **Pool 1**
- **Pool 2**
- **Pool 3**

FIG. 7H
FIG. 7J

- BSA - kDa
- EPCAM Histagged Exp. MW 29 kDa
- Appears as 36 kDa due to glycosilation
A) Direct conjugation of cMVs to a substrate

B) Biotin functionalized lipid crosslinking of cMVs to a substrate
C) Immuno-precipitation of cMVs with cMV specific antibodies immobilized to a substrate

D) Immuno-precipitation of cMVs with cMV specific aptamer immobilized to a substrate
### Cancer v Non-Cancer

<table>
<thead>
<tr>
<th></th>
<th>P-0.005</th>
<th>P-0.005</th>
</tr>
</thead>
<tbody>
<tr>
<td>d-50</td>
<td>2471</td>
<td>2.95%</td>
</tr>
<tr>
<td>d-100</td>
<td>1615</td>
<td>3.07%</td>
</tr>
<tr>
<td>d-200</td>
<td>914</td>
<td>3.18%</td>
</tr>
<tr>
<td>d-500</td>
<td>315</td>
<td>3.10%</td>
</tr>
<tr>
<td>Random</td>
<td></td>
<td>0.50%</td>
</tr>
</tbody>
</table>

### Random

<table>
<thead>
<tr>
<th></th>
<th>P-0.005</th>
<th>P-0.005</th>
</tr>
</thead>
<tbody>
<tr>
<td>d-50</td>
<td>49</td>
<td>0.06%</td>
</tr>
<tr>
<td>d-100</td>
<td>27</td>
<td>0.05%</td>
</tr>
<tr>
<td>d-200</td>
<td>17</td>
<td>0.06%</td>
</tr>
<tr>
<td>d-500</td>
<td>5</td>
<td>0.05%</td>
</tr>
<tr>
<td>Random</td>
<td></td>
<td>0.50%</td>
</tr>
</tbody>
</table>

**FIG. 12E**
FIGs. 17D-E
FIG. 18A
1910

Provide Aptamer Pool to Target of Interest

1911

Contact Pool with Sample

1912

Wash, Retain Aptamers that Bind Sample

1913

Identify Retained Aptamers

1914

Characterize Sample

1915

FIG. 19B
FIG. 20H

Sample 2012

Supernatant 2014

Probing Library 2011

Bind 2013

Microvesicle pellet 2015

Elute 2016

Sequencing 2017

Profile 2018
Patient Sample

Identify probes that bind sample 2353

Patient Sample 2352

Administer Construct 2356

Monitor Disease 2357

Synthesize Construct 2355

Identify Probe Targets 2354

Oligonucleotide probe library 2351

FIG. 23E
FIG. 24B

amount DNA bound [%]

- B1RC
- B1
- B11

conc. C1Q [nM]

0 6.6 66.0 330.0 660.0
Absorbance [450-540 nm]

• B11
• B11RC
• no aptamer

Aptamer conc. [nM]

[C1Q] = 0.625 nM

FIG. 24C
INTERNATIONAL SEARCH REPORT

A. CLASSIFICATION OF SUBJECT MATTER

IPC(B) - C12Q 1/66; C07K 14/47; A61K 35/17; C12N 15/11 (2016.01)

CPC - C12Q 1/6811, 1/6869; C07K 14/47, 14/70503; A61K 35/17; C12N 15/11

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC(B): C12Q 1/66; C07K 14/47, 14/525; A61K 35/17, 38/17, 38/19, 48/00; C12N 15/11, 15/62; A01K 67/027 (2016.01)

CPC: C12Q 1/6811, 1/6869; C07K 14/47, 14/70503, 14/47Z; 14/525; A61K 35/17; 38/00, 39/00, 48/00; C12N 15/11

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic database consulted during the international search (name of database and, where practicable, search terms used)

PatSeer (US, EP, WO, JP, DE, GB, CN, FR, KR, ES, AU, IN, CA, INPADOC Data); Lens sequence search database; Google Scholar; Pubmed; EBSCO; NCBI BLAST, ENA; oligonucleotide, nucleic acid, multipartite, construct, target, epitope, C1q

C. DOCUMENTS CONSIDERED TO BE RELEVANT

<table>
<thead>
<tr>
<th>Category</th>
<th>Citation of document, with indication, where appropriate, of the relevant passages</th>
<th>Relevant to claim No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>US 2003/0064402 A1 (EGHOLM, M et al.) 03 April, 2003; abstract; paragraph [0006]; claim 1</td>
<td>50-51, 52/50-51, 55, 56, 59-63, 67-76</td>
</tr>
</tbody>
</table>

Further documents are listed in the continuation of Box C. | See patent family annex.

Date of the actual completion of the international search: 28 October 2016 (28.10.2016)

Date of mailing of the international search report: 28 NOV 2016

Name and mailing address of the ISA/ Authorised officer
Mail Stop PCT, Attn: ISA/US, Commissioner for Patents
P.O. Box 1450, Alexandria, Virginia 22313-1450
Facsimile No. 571-273-8300
Authorized officer Shane Thomas
PCT Helpdesk: 571-272-4300
PCT OSP: 571-272-7774

Form PCT/ISA/210 (second sheet) (January 2015)
This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. Claim Nos.: because they relate to subject matter not required to be searched by this Authority, namely:

2. Claim Nos.: because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. Claim Nos.: 8-49, 53, 54, 57, 58, 64-66, 77-101, 113-135 because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

This International Searching Authority found multiple inventions in this international application, as follows:

"-Please See Supplemental Page-" *.

As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.

As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims: it is covered by claims Nos.:

Groups I, Claims 1-7, 50-52, 55, 56, 59-63, 67-76; SEQ ID NO: 1472 (oligonucleotide sequence); SEQ ID NO: 131 (5' extender sequence); Exemplary election: SEQ ID NO: 1475 (oligonucleotide sequence).

Remark on Protest

The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

No protest accompanied the payment of additional search fees.
This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees must be paid.

Groups I, Claims 1-7, 50-52, 55, 56, 59-63, 67-76; and SEQ ID NOs: 1472 and 131 are directed toward an oligonucleotide comprising a central sequence comprising a sequence according to any one of SEQ ID NOs. 1472, 1475, 1477, 1482, and 4151-4325, and a 5′ region with sequence ID NO: 131; wherein the oligonucleotide is capable of binding to complement protein C1q; and a multipartite construct that comprises a first segment that binds to a first target, and a second segment that binds to a second target.

The oligonucleotide and construct will be searched to the extent that they encompass SEQ ID NO: 1472 (first exemplary oligonucleotide sequence); and SEQ ID NO: 131 (first exemplary 5′ region extender sequence). Applicant is invited to elect additional oligonucleotide sequence(s), with specified SEQ ID NO: for each, and/or adaptor or extender sequence(s) with specified SEQ ID NO: for each, to be searched. Additional oligonucleotide sequence(s) and/or adaptor/extender sequence(s) will be searched upon the payment of additional fees. It is believed that claims 1 (in-part), 2 (in-part), 3 (in-part), 5 (in-part), 7 (in-part), 50 (in-part), 51 (in-part), 52 (in-part), 55 (in-part), 56 (in-part), 59 (in-part), 60 (in-part), 61 (in-part), 62 (in-part) and 63 (in-part) encompass this first named invention and thus these claims will be searched without fee to the extent that they encompass SEQ ID NO: 1472 (oligonucleotide sequence); and SEQ ID NO: 131 (5′ extender sequence). Applicants must specify the claims that encompass any additionally elected oligonucleotide and/or adaptor/extender sequence(s). Applicants must further indicate, if applicable, the claims which encompass the first named invention, if discovered, to clearly identify how any group. Failure to clearly identify how any paid additional invention fees are to be applied to the "*" group(s) will result in only the first claimed invention to be searched/examined. An exemplary election would be an oligonucleotide encompassing SEQ ID NO: 1475 (first exemplary elected oligonucleotide sequence).

Groups II, Claims 102-112 and KM16A are directed toward a method of detecting at least one microvesicle in a biological sample comprising contacting the biological sample with at least one binding agent.

The method can be searched to the extent that the at least one protein encompasses KRT6A (first exemplary target protein). Applicant is invited to elect additional target protein(s) to be searched. Additional target protein(s) can be searched upon the payment of additional fees. It is believed that claims 102 (in-part), 103 (in-part), 104 (in-part), 105 (in-part), 106 (in-part), 107 (in-part), 108 (in-part), 109 (in-part), 110 (in-part), 111 (in-part) and 112 (in-part) encompass this first named invention and thus these claims can be searched with payment of a fee for the search of Groups II, to the extent that they encompass KRT6A (target protein). Failure to clearly identify how any paid additional invention fees are to be applied to the "*" group(s) can result in only the first claimed invention to be searched/examined. An exemplary election would be a protein encompassing desmoplakin (DSP) (first exemplary elected target protein).

The inventions listed as Groups I, and II, do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: the special technical features of Groups I include SEQ ID NO: 1472, not present in any of Groups II; the special technical features of Groups II include detection of a microvesicle, not present in any of Groups I.

No technical features are shared between the oligonucleotide sequences of Groups I and, accordingly, these groups lack unity a priori. No technical features are shared between the adaptor/extends sequences of Groups I and, accordingly, these groups lack unity a priori. No technical features are shared between the proteins of Groups II and, accordingly, these groups lack unity a priori.

Groups I and II share the technical features including: binding. Groups I share the technical features including: an oligonucleotide comprising a central sequence, and a 5′ region with sequence 5′CTAGCATGACTGCAGTACGT [SEQ ID NO: 131], a 3′ region with sequence 5′- CTGTCTCCTTATACACATCTGAGCTGCAGCGCAGCA [SEQ ID NO: 132], or both; wherein the oligonucleotide is capable of binding to complement protein complex C1q or a subunit thereof; and a multipartite construct that comprises a first segment that binds to a first target and a second segment that binds to a second target. Groups II share the technical features including: a method of detecting at least one microvesicle in a biological sample comprising contacting the biological sample with at least one binding agent to at least one protein, and detecting the at least one microvesicle recognized by the binding agent to at least one protein.

However, these shared technical features are previously disclosed by WO 2014/068408 A2 to Caris Life Sciences Switzerland Holdings, S.A.R.L. (hereinafter ‘Caris’) in view of US 2003/0078221 A1 to Bieseker et al. (hereinafter ‘Bieseker’) and WO 2014/190214 A1 to Active Motif Inc. (hereinafter ‘Active’).

Caris discloses binding (binding; abstract); an oligonucleotide comprising a central sequence (an aptamer (oligonucleotide) comprising a central sequence; paragraph [00578], Table 11), and a 5′ region with sequence (a 5′ region with sequence); paragraph [00578], Table 11), and a 3′ region with sequence (a 3′ region with sequence); paragraph [00578], Table 11), wherein the oligonucleotide is capable of binding to a protein (wherein the aptamer (oligonucleotide) is capable of binding to a surface antigen (protein); abstract, paragraph [0019]); and a multipartite construct (an aptamer comprising a leader, a central sequence, and a tail (a multipartite construct); paragraphs [00297], [00578], Table 11) that comprises a first segment that binds to a target (that comprises a first segment that binds to a first target); paragraph [00297]) and a second segment that binds to a second target (and a leader or tail sequence that facilitates primer binding or capture (a second segment that binds to a second target); paragraph [00297]); a method of detecting at least one microvesicle in a biological sample (a method of detecting at least one microvesicle in a biological sample; paragraph [0027]) comprising contacting the biological sample with at least one binding agent to at least one protein (comprising contacting the biological sample with at least one aptamer (binding agent to at least one protein) abstract, paragraphs [0019], [0027]), and detecting the at least one microvesicle recognized by the binding agent to at least one protein (detecting the at least one microvesicle recognized by the binding agent to at least one protein; paragraphs [0019], [0027]).
Caris does not disclose 5’- CTGTCCTTATACATCTGACGCTGCGACGA (SEQ ID NO. 132); wherein the oligonucleotide is capable of binding to complement protein complex C1q or a subunit thereof.

Biesecker discloses high affinity nucleic acid ligands for complement C1q (high affinity nucleic acid ligands for complement C1q; abstract); wherein the complement protein is present in a vesicle (wherein the complement protein is present in a vesicle; paragraph [0043]).

Active discloses an oligonucleotide comprising a sequence complementary to SEQ ID NO: 132 (an Illumina oligonucleotide comprising SEQ ID NO: 10 (an oligonucleotide comprising a sequence complementary SEQ ID NO: 132); paragraph [0043], Figure 16; SEQ ID NO: 10), wherein the oligonucleotide comprises a commercially available oligonucleotide adapter (wherein the oligonucleotide comprises a commercially available oligonucleotide adapter; paragraph [0043]).

It would have been obvious to a person of ordinary skill in the art at the time of the invention was made to have modified the disclosure of Caris to have used the high affinity nucleic acids that bind to C1q, as disclosed by Biesecker, in order to enable the detection of vesicles including complement proteins, such as C1q, as disclosed by Biesecker. It further would have been obvious to a person of ordinary skill in the art at the time of the invention was made to have modified the disclosure of Caris to have ligated an adapter including a commercially available sequence, such as SEQ ID NO: 132, or its complement, as disclosed by Active, in order to enable capture of the aptamer for detection and/or amplification, as disclosed by Caris, using kits and reagents from the provider of the nucleic acid sequence.

Since none of the special technical features of the Groups I+ and II+ inventions is found in more than one of the inventions, and since all of the shared technical features are previously disclosed by a combination of the Caris, Biesecker and Active references, unity of invention is lacking.