

METHODS FOR SAMPLE QUALITY ASSESSMENT

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims the benefit of priority of US Provisional Application No. 63/389,061, filed July 14, 2022, which is incorporated by reference herein in its entirety for any purpose.

FIELD OF THE INVENTION

[0002] The present application relates generally to the detection of biomarkers and a method of evaluating the quality or suitability of a sample or set of samples for use in biomarker discovery or in a medical evaluation of a subject, such as a diagnostic assay.

BACKGROUND

[0003] Blood contains various cellular and humoral systems for reacting to injury or foreign and infectious agents. Small challenges can induce the innate immune system (complement system and cells such as macrophages) to release signals and enzymes, lead to activation of the platelets and trigger the coagulation of the blood. These signals are of interest because they can be directly involved in defense and repair systems and serve as markers for disease. However, such process signals may also be responsive to the effects of blood sample preparation or processing. If cells in the samples lyse, if platelets degranulate, or if the complement system is activated, then changes in analyte concentration may occur in the sample after it has been taken, and a “high fidelity” measurement technique may detect them. Merely exposing blood to air can result in unintended activation of these mechanisms. Thus, altering the time of sample processing steps can alter the apparent composition of serum or plasma such that physiologic information is masked by the pre-analytic variability imparted on the sample during collection and processing. The susceptibility of these processes and proteins to subtle alterations in sample handling can compromise the use of them as biomarkers.

[0004] Currently, researchers in multivariate biology are concerned about pre-analytical sample variation (often called “batch effects”). The extent to which sample quality can be determined is largely limited to visually obvious changes, such as red color indicating red cell lysis and cloudiness indicating high lipid or other contaminants. These relatively crude methods limit the reliability of all but the hardest and most robust protein

measurements. Ostroff, R. et al. (2010) J. Proteomics 73:649-666, documents some of the complex and nonlinear effects of variations in serum and plasma preparation.

[0005] Specific techniques to determine compliance with sample processing protocols are needed to monitor compliance, reject low quality samples, and/or make corrections in analytes of interest. Such techniques would improve evaluation of the quality of human or animal blood samples used in biomarker research, clinical diagnostic applications, bio-banks, and drug development.

SUMMARY OF THE INVENTION

[0006] The present application includes biomarkers, methods, reagents, devices, systems, and kits for assessing sample quality. The biomarkers of the present application were identified using linear regression of measurements of a specific set of proteins affected by variation in sample processing protocol. In some embodiments, the biomarker panels comprise proteins that are sensitive to sample handling.

[0007] In some embodiments, the methods comprise detecting the biomarkers using a multiplex slow off-rate aptamer-based assay described herein, for example, for assessing sample quality. In some embodiments, the sample is a blood sample, plasma sample, serum sample, or urine sample. In some embodiments, the sample is a plasma sample or a serum sample.

[0008] In some embodiments, the number of times a sample is frozen and thawed (freeze-thaw cycle) is predicted or estimated. In some embodiments, the sample processing steps include one or more of sample centrifugation, decanting or aspirating of the centrifuged supernatant, and freezing the decanted or aspirated sample.

[0009] In some embodiments, methods of assessing quality of a sample collected from a subject are provided, the method comprising detecting the level of each of N biomarker proteins in the sample, wherein N is at least 1, and wherein at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, or at least 7 of the N biomarker proteins are selected from LKHA4, EIF1B, RAGP1, SCFD1, PLD3, Lectin, mannose-binding 2, ILK1, and RPL5, wherein the sample is a serum sample.

[0010] In some embodiments, methods comprises measuring the level of each of N biomarker proteins in a serum sample from a subject are provided, wherein N is at least 1, and wherein at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, or at least 7 of the N biomarker proteins are selected from LKHA4, EIF1B, RAGP1, SCFD1, PLD3, Lectin, mannose-binding 2, ILK1, and RPL5, and identifying the sample as an analysis sample or

negative sample based on the level of the N biomarker proteins, wherein the analysis sample is a sample that is suitable for use in one or more of the following: protein biomarker discovery analysis, protein expression level analysis, a diagnostic method or a prognostic method, and the negative sample is a sample that is not suitable for use as an analysis sample. In some embodiments, the sample was frozen and thawed prior to the detecting. In some such embodiments, the number of times the sample was frozen and thawed prior to the detecting is determined.

[0011] In some embodiments, methods are provided. The methods comprising a) contacting a serum sample from a subject with a set of capture reagents, wherein each capture reagent has affinity for a different biomarker protein of N biomarker proteins wherein N is at least 1, and wherein at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, or at least 7 of the N biomarker proteins is selected from LKHA4, EIF1B, RAGP1, SCFD1, PLD3, Lectin, mannose-binding 2, ILK1, and RPL5; and b) measuring the level of each N biomarker protein with the set of capture reagents.

[0012] In some embodiments, methods are provided for comparing a plurality of samples collected from a plurality of subjects comprising detecting the level of each of N biomarker proteins in each of the plurality of samples, wherein N is at least 1, and wherein at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, or at least 7 of the N biomarker proteins are LKHA4, EIF1B, RAGP1, SCFD1, PLD3, Lectin, mannose-binding 2, ILK1, and RPL5, wherein the sample is a serum sample. In some embodiments, methods comprise a) determining the number of freeze-thaw cycles and b) comparing the determined number freeze-thaw cycles for each of the plurality of samples. In some embodiments, methods comprise identifying the plurality of samples as handled consistently or inconsistently, wherein samples handled consistently all have a determined number of freeze-thaw cycles within 1, 2, 3, 4 or 5 freeze-thaw cycles of each other. In some embodiments, the determining is based on comparing the detected levels of each of the N biomarker proteins to reference levels, wherein the reference levels are average levels of each of the N biomarker proteins present in samples having 1 freeze-thaw cycle or 2 freeze-thaw cycles. In some embodiments, the detected levels of each of the N biomarker proteins compared to the reference levels indicates that the number of freeze-thaw cycles was greater than 1, greater than 2, greater than 3, greater than 4, greater than 5, greater than 7 or greater than 10 freeze-thaw cycles; or wherein the levels of each of the N biomarker proteins used in a linear regression model predicts the number of freeze-thaw cycles was greater than 1, greater than 2, greater than 3, greater than 4, greater than 5, greater than 7 or greater than 10 freeze-thaw

cycles. In some embodiments, the determining is based on a panel of N biomarker proteins having an R^2 value of at least 0.65, at least 0.7, at least 0.75, at least 0.8, at least 0.85, at least 0.9, or at least 0.95.

[0013] In some embodiments, the methods comprise performing protein biomarker discovery analysis, protein expression level analysis, a diagnostic method or a prognostic method on the plurality of samples. In some embodiments, the methods comprise modifying a panel of proteins in the protein biomarker discovery analysis, the protein expression level analysis, the diagnostic method or the prognostic method based on the determined number of freeze-thaw cycles for each of the plurality of samples; or identifying one or more proteins in the sample as being affected by the number of freeze-thaw cycles; or identifying the level of one or more proteins in the sample as being affected by the number of freeze-thaw cycles; or changing the proteins used in a diagnostic, a prognostic or a health assessment related test based on the predicted number of freeze-thaw cycles; removing the proteins used in a diagnostic, a prognostic or a health assessment related test based on the predicted number of freeze-thaw cycles. In some embodiments, the panel of biomarker proteins is reduced in number of biomarker proteins measured.

[0014] In some embodiments, the determining measures compliance with a clinical trial sample collection and processing protocol. In some embodiments, the plurality of samples are collected at more than one sample collection site. In some embodiments, the plurality of samples from a first sample collection site are compared to a second plurality of samples from a second sample collection site. In some embodiments, one or more of the plurality of samples may be removed based on the number of freeze-thaw cycles.

[0015] In some embodiments, methods are provided comprising detecting the level of each of N biomarker proteins in a sample, wherein N is at least 1, and wherein at least 1 of the N biomarker proteins is selected from LKHA4, EIF1B, RAGP1, SCFD1, PLD3, Lectin, mannose-binding 2, ILK1, and RPL5. In some embodiments, the sample is a serum sample. In some embodiments, the serum sample is a human serum sample. In some embodiments, the number of freeze-thaw cycles is determined with the level of each of the N biomarkers. In some embodiments, the number of freeze-thaw cycles was greater than 1, greater than 2, greater than 3, greater than 4, greater than 5, greater than 7 or greater than 10 freeze-thaw cycles. In some embodiments, the determined number of freeze-thaw cycles is derived from the input of the level of each of N biomarker proteins in a statistical model. In some embodiments, the statistical model is a linear regression model.

[0016] In some embodiments, methods are provided comprising detecting the level of each of at least 1, 2, 3, 4, 5, 6, 7, or 8 biomarker proteins in a sample, wherein biomarker proteins are selected from LKHA4, EIF1B, RAGP1, SCFD1, PLD3, Lectin, mannose-binding 2, ILK1, and RPL5. In some embodiments, the sample is a serum sample. In some embodiments, the serum sample is a human serum sample. In some embodiments, a number of freeze-thaw cycles is determined with the level of each of the at least 1, 2, 3, 4, 5, 6, 7, or 8 biomarker proteins. In some embodiments, the number of freeze-thaw cycles was greater than 1, greater than 2, greater than 3, greater than 4, greater than 5, greater than 7 or greater than 10 freeze-thaw cycles. In some embodiments, the determined number of freeze-thaw cycles was greater than 1, greater than 2, greater than 3, greater than 4, greater than 5, greater than 7 or greater than 10 freeze-thaw cycles is derived from the input of the level each of the at least 1, 2, 3, 4, 5, 6, 7, or 8 biomarker proteins in a statistical model. In some embodiments, the statistical model is a linear regression model. In some embodiments, the methods further comprise modifying a panel of proteins in a protein biomarker discovery analysis, a protein expression level analysis, a diagnostic method or a prognostic method; identifying one or more proteins in the sample as being affected; identifying the level of one or more proteins in the sample as being affected; changing the proteins used in a diagnostic, a prognostic or a health assessment related test; or removing one or more proteins used in a diagnostic, a prognostic or a health assessment related test, each based on the outcome of the linear regression model.

[0017] In some embodiments, methods of assessing quality of a sample collected from a subject comprise detecting N biomarker proteins, wherein one or more of the N biomarker proteins are associated with the number of times the sample was frozen and thawed prior to the detecting.

[0018] In some embodiments, N is 1, N is 2, N is 3, N is 4, N is 5, N is 6, N is 7, or N is 8. In embodiments, additional biomarkers are measured and N is 9, N is 10, N is 11, N is 12, N is 13, N is 14, N is 15, N is 16, N is 17, N is 18, N is 19, or N is 20 or greater. In some embodiments, all of the N biomarker proteins are selected from the list in Table 1.

[0019] In some embodiments, the subject is a human subject. In some embodiments, the sample is a serum sample obtained from a whole blood sample. In some embodiments, the number of freeze-thaw cycles determined by a method herein is 1, 2, 3, 4, 5, 7, 10 or greater freeze-thaw cycles. In some embodiments, the method is performed *in vitro*. In some embodiments, the approximate number of freeze-thaw cycles is derived from the input of the

level of each of N biomarker proteins in a statistical model. In some embodiments, the statistical model is a linear regression model.

[0020] In some embodiments, the sample is identified as passing a quality assessment or failing a quality assessment. In some such embodiments, the identifying is based, at least in part, on the detected levels of the N biomarker proteins in the sample. In some embodiments, the identifying is based, at least in part, of the determined number of freeze-thaw cycles. In some embodiments, the sample undergoes further analysis if it is identified as passing quality assessment, or the sample is discarded if it is identified as failing the quality assessment.

[0021] In some embodiments, methods of assessing the quality of a plurality of samples are provided herein, comprising detecting the level of each of N biomarkers in a plurality of samples from a plurality of subjects. In some such embodiments, the number of freeze-thaw cycles is determined and compared across the plurality of samples. In some such embodiments, the consistency of sample handling across the plurality of samples is determined.

[0022] In some embodiments, the method comprises contacting biomarker proteins of the sample from the subject with a set of capture reagents, wherein each capture reagent of the set of capture reagents specifically binds to one biomarker protein being detected. In some embodiments, the method comprises contacting biomarker proteins of the sample from the subject with a set of capture reagents, wherein each capture reagent of the set of capture reagents specifically binds to a different biomarker protein being detected. In some embodiments, each capture reagent is an antibody or an aptamer. In some embodiments, each biomarker capture reagent is an aptamer. In some embodiments, at least one aptamer is a slow off-rate aptamer. In some embodiments, at least one slow off-rate aptamer comprises at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, or at least 10 nucleotides with modifications. In some embodiments, each slow off-rate aptamer binds to its target protein with an off rate ($t_{1/2}$) of ≥ 20 minutes, ≥ 30 minutes, ≥ 60 minutes, ≥ 90 minutes, ≥ 120 minutes, ≥ 150 minutes, ≥ 180 minutes, ≥ 210 minutes, or ≥ 240 minutes. In some embodiments, the level of each protein measured is determined from a relative fluorescence unit (RFU) or a protein concentration.

[0023] In some embodiments, a kit is provided, the kit comprising N biomarker protein capture reagents, wherein N is at least 1 and at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, or at least 7 of the capture reagents bind to LKHA4, EIF1B, RAGP1, SCFD1, PLD3, Lectin, mannose-binding 2, ILK1, or RPL5. In some embodiments, each

capture reagent binds to a different biomarker protein. In some embodiments, N is 1, N is 2, N is 3, N is 4, N is 5, N is 6, N is 7, N is 8, N is 9, N is 10, N is 11, N is 12, N is 13, N is 14, N is 15, N is 16, N is 17, N is 18, N is 19, or N is 20 or greater. In some embodiments, the kit comprises capture reagents from multiple sample processing panels.

[0024] In some embodiments, each of the N biomarker protein capture reagents specifically binds to a biomarker protein selected from Table 1. In some embodiments, each of the N biomarker capture reagents is an antibody or an aptamer. In some embodiments, each biomarker capture reagent is an aptamer. In some embodiments, at least one aptamer is a slow off-rate aptamer. In some embodiments, at least one slow off-rate aptamer comprises at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, or at least 10 nucleotides with modifications. In some embodiments, each slow off-rate aptamer binds to its target protein with an off rate ($t_{1/2}$) of ≥ 20 minutes, ≥ 30 minutes, ≥ 60 minutes, ≥ 90 minutes, ≥ 120 minutes, ≥ 150 minutes, ≥ 180 minutes, ≥ 210 minutes, or ≥ 240 minutes. In some embodiments, the kit is for use in detecting the N biomarker proteins in a sample from a subject. In some embodiments, the kit is for use in assessing the quality of the sample or a plurality of samples.

BRIEF DESCRIPTION OF THE DRAWINGS

[0025] Fig. 1 illustrates a nonlimiting exemplary computer system for use with various computer-implemented methods described herein.

[0026] Fig. 2 illustrates a nonlimiting exemplary aptamer assay that can be used to detect one or more biomarkers in a biological sample.

[0027] Figs. 3-5 show certain exemplary modified pyrimidines that may be incorporated into aptamers, such as slow off-rate aptamers.

[0028] Fig. 6 shows the Rsquared values for models with increasing number of features.

DETAILED DESCRIPTION

[0029] While the invention will be described in conjunction with certain representative embodiments, it will be understood that the invention is defined by the claims, and is not limited to those embodiments.

[0030] One skilled in the art will recognize many methods and materials similar or equivalent to those described herein may be used in the practice of the present invention. The present invention is in no way limited to the methods and materials described.

[0031] Unless defined otherwise, technical and scientific terms used herein have the meaning commonly understood by one of ordinary skill in the art to which this invention belongs. Although any methods, devices, and materials similar or equivalent to those described herein can be used in the practice of the invention, certain methods, devices, and materials are described herein.

[0032] All publications, published patent documents, and patent applications cited herein are hereby incorporated by reference to the same extent as though each individual publication, published patent document, or patent application was specifically and individually indicated as being incorporated by reference.

[0033] As used herein, the terms “comprises,” “comprising,” “includes,” “including,” “contains,” “containing,” and any variations thereof, are intended to cover a non-exclusive inclusion, such that a process, method, product-by-process, or composition of matter that comprises, includes, or contains an element or list of elements may include other elements not expressly listed.

[0034] “Biological sample”, “sample”, and “test sample” are used interchangeably herein to refer to any material, biological fluid, tissue, or cell obtained or otherwise derived from an individual. This includes blood (including whole blood, leukocytes, peripheral blood mononuclear cells, buffy coat, plasma, and serum), sputum, tears, mucus, nasal washes, nasal aspirate, urine, saliva, peritoneal washings, ascites, cystic fluid, glandular fluid, lymph fluid, bronchial aspirate, synovial fluid, joint aspirate, organ secretions, cells, a cellular extract, and cerebrospinal fluid. This also includes experimentally separated fractions of all of the preceding. For example, a blood sample can be fractionated into serum, plasma, or into fractions containing particular types of blood cells, such as red blood cells or white blood cells (leukocytes). In some embodiments, a sample is a plasma sample. As used herein, a “plasma sample” comprises plasma and, optionally, one or more preservatives or additives. A plasma sample is isolated from whole blood and therefore does not comprise a substantial amount of any other blood component. In some embodiments, a sample is a serum sample. As used herein, a “serum sample” comprises serum and, optionally, one or more preservatives or additives. A serum sample is isolated from whole blood and therefore does not comprise a substantial amount of any other blood component. In some embodiments, a sample is a urine sample. As used herein, a “urine sample” comprises urine and, optionally, one or more preservatives or additives. In some embodiments, a blood sample is a dried blood spot. In some embodiments, a plasma sample is a dried plasma spot. In some embodiments, a sample can be a combination of samples from an individual, such as a

combination of a tissue and fluid sample. The term “biological sample” also includes materials containing homogenized solid material, such as from a stool sample, a tissue sample, or a tissue biopsy, for example. The term “biological sample” also includes materials derived from a tissue culture or a cell culture. Any suitable methods for obtaining a biological sample can be employed; exemplary methods include, e.g., phlebotomy, swab (e.g., buccal swab), and a fine needle aspirate biopsy procedure. Exemplary tissues susceptible to fine needle aspiration include lymph node, lung, thyroid, breast, pancreas, and liver. Samples can also be collected, e.g., by micro dissection (e.g., laser capture micro dissection (LCM) or laser micro dissection (LMD)), bladder wash, smear (e.g., a PAP smear), or ductal lavage. A “biological sample” obtained or derived from an individual includes any such sample that has been processed in any suitable manner after being obtained from the individual.

[0035] Further, in some embodiments, a biological sample may be derived by taking biological samples from a number of individuals and pooling them, or pooling an aliquot of each individual’s biological sample. The pooled sample may be treated as described herein for a sample from a single individual, and, for example, if a poor sample quality is established in the pooled sample, then each individual biological sample can be re-tested to determine which must be discarded, or the entire group of samples may be discarded if known to have been handled or processed in the same way.

[0036] For purposes of this specification, the phrase “data attributed to a biological sample from an individual” is intended to mean that the data in some form derived from, or were generated using, the biological sample of the individual. The data may have been reformatted, revised, or mathematically altered to some degree after having been generated, such as by conversion from units in one measurement system to units in another measurement system; but, the data are understood to have been derived from, or were generated using, the biological sample.

[0037] “Target”, “target molecule”, and “analyte” are used interchangeably herein to refer to any molecule of interest that may be present in a biological sample. A “molecule of interest” includes any minor variation of a particular molecule, such as, in the case of a protein, for example, minor variations in amino acid sequence, disulfide bond formation, glycosylation, lipidation, acetylation, phosphorylation, or any other manipulation or modification, such as conjugation with a labeling component, which does not substantially alter the identity of the molecule. A “target molecule”, “target”, or “analyte” refers to a set of copies of one type or species of molecule or multi-molecular structure. Exemplary target molecules include proteins, polypeptides, nucleic acids, carbohydrates, lipids,

polysaccharides, glycoproteins, hormones, receptors, antigens, antibodies, affybodies, antibody mimics, viruses, pathogens, toxic substances, substrates, metabolites, transition state analogs, cofactors, inhibitors, drugs, dyes, nutrients, growth factors, cells, tissues, and any fragment or portion of any of the foregoing. In some embodiments, a target molecule is a protein, in which case the target molecule may be referred to as a “target protein.”

[0038] As used herein, a “capture agent” or “capture reagent” refers to a molecule that is capable of binding specifically to a biomarker. A “target protein capture reagent” refers to a molecule that is capable of binding specifically to a target protein. Nonlimiting exemplary capture reagents include aptamers, antibodies, adnectins, ankyrins, other antibody mimetics and other protein scaffolds, autoantibodies, chimeras, small molecules, nucleic acids, lectins, ligand-binding receptors, imprinted polymers, avimers, peptidomimetics, hormone receptors, cytokine receptors, synthetic receptors, and modifications and fragments of any of the aforementioned capture reagents. In some embodiments, a capture reagent is selected from an aptamer and an antibody.

[0039] As used herein, “polypeptide,” “peptide,” and “protein” are used interchangeably herein to refer to polymers of amino acids of any length. The polymer may be linear or branched, it may comprise modified amino acids, and it may be interrupted by non-amino acids. The terms also encompass an amino acid polymer that has been modified naturally or by intervention; for example, disulfide bond formation, glycosylation, lipidation, acetylation, phosphorylation, or any other manipulation or modification, such as conjugation with a labeling component. Also included within the definition are, for example, polypeptides containing one or more analogs of an amino acid (including, for example, unnatural amino acids, etc.), as well as other modifications known in the art. Polypeptides can be single chains or associated chains. Also included within the definition are preproteins and intact mature proteins; peptides or polypeptides derived from a mature protein; fragments of a protein; splice variants; recombinant forms of a protein; protein variants with amino acid modifications, deletions, or substitutions; digests; and post-translational modifications, such as glycosylation, acetylation, phosphorylation, and the like.

[0040] The term “antibody” refers to full-length antibodies of any species and fragments and derivatives of such antibodies, including Fab fragments, F(ab')₂ fragments, single chain antibodies, Fv fragments, and single chain Fv fragments. The term “antibody” also refers to synthetically-derived antibodies, such as phage display-derived antibodies and fragments, affybodies, nanobodies, etc.

[0041] As used herein, “marker” and “biomarker” are used interchangeably to refer to a target molecule that indicates or is a sign of a normal or abnormal process in an individual or of a high or low quality sample. More specifically, a “marker” or “biomarker” is an anatomic, physiologic, biochemical, or molecular parameter associated with the presence of a specific state or process, whether normal or abnormal, and, if abnormal, whether chronic or acute. Biomarkers are detectable and measurable by a variety of methods including laboratory assays and medical imaging. When a biomarker is a protein, it is also possible to use the expression of the corresponding gene as a surrogate measure of the amount or presence or absence of the corresponding protein biomarker in a biological sample or methylation state of the gene encoding the biomarker or proteins that control expression of the biomarker. In certain aspects, a feature is an analyte/ SOMAmer reagent of other predictors in a statistical model.

[0042] As used herein, “biomarker level” and “level” refer to a measurement that is made using any analytical method for detecting the biomarker in a biological sample and that indicates the presence, absence, absolute amount or concentration, relative amount or concentration, titer, a level, an expression level, a ratio of measured levels, or the like, of, for, or corresponding to the biomarker in the biological sample. The exact nature of the “level” depends on the specific design and components of the particular analytical method employed to detect the biomarker.

[0043] When a biomarker indicates or is a sign of a poor quality sample, that biomarker is generally described as being either over-expressed or under-expressed as compared to an expression level or value of the biomarker that indicates or is a sign of a normal or quality sample. “Up-regulation”, “up-regulated”, “over-expression”, “over-expressed”, and any variations thereof are used interchangeably to refer to a value or level of a biomarker in a biological sample that is greater than a value or level (or range of values or levels) of the biomarker that is typically detected in similar, properly handled biological samples.

[0044] “Down-regulation”, “down-regulated”, “under-expression”, “under-expressed”, and any variations thereof are used interchangeably to refer to a value or level of a biomarker in a biological sample that is less than a value or level (or range of values or levels) of the biomarker that is typically detected in similar, properly handled biological samples. The terms may also refer to a value or level of a biomarker in a biological sample that is less than a value or level (or range of values or levels) of the biomarker that may be detected at a different stage of a particular disease.

[0045] Further, a biomarker that is either over-expressed or under-expressed can also be referred to as being “differentially expressed” or as having a “differential level” or “differential value” as compared to a “normal” expression level or value of the biomarker that indicates or is a sign of a normal process or proper sample handling. Thus, “differential expression” of a biomarker can also be referred to as a variation from a “normal” expression level of the biomarker.

[0046] The term “differential gene expression” and “differential expression” are used interchangeably to refer to a gene (or its corresponding protein expression product) whose expression is activated to a higher or lower level in a subject suffering from a specific disease or condition, relative to its expression in a normal or control subject. The terms also include genes (or the corresponding protein expression products) whose expression is activated to a higher or lower level at different stages of the same disease or condition. It is also understood that a differentially expressed gene may be either activated or inhibited at the nucleic acid level or protein level, or may be subject to alternative splicing to result in a different polypeptide product. Such differences may be evidenced by a variety of changes including mRNA levels, surface expression, secretion or other partitioning of a polypeptide. Differential gene expression may include a comparison of expression between two or more genes or their gene products; or a comparison of the ratios of the expression between two or more genes or their gene products; or even a comparison of two differently processed products of the same gene, which differ between normal subjects and subjects suffering from a disease; or between various stages of the same disease. Differential expression includes both quantitative, as well as qualitative, differences in the temporal or cellular expression pattern in a gene or its expression products among, for example, normal and diseased cells, or among cells which have undergone different disease events or disease stages.

[0047] A “control level” of a target molecule refers to the level of the target molecule in a properly handled sample of the same sample type. Control level may refer to the average level of the target molecule in properly handled samples from a population of individuals.

[0048] As used herein, “individual,” “subject,” and “patient” are used interchangeably to refer to a mammal. A mammalian individual can be a human or non-human. In various embodiments, the individual is a human. A healthy or normal individual is an individual in which the disease or condition of interest (including, for example, chronic heart failure and cardiovascular events such as myocardial infarction, stroke and hospitalization for heart failure) is not detectable by conventional diagnostic methods.

[0049] As used herein, “detecting” or “determining” with respect to a biomarker value includes the use of both the instrument used to observe and record a signal corresponding to a biomarker level and the material/s required to generate that signal. In various embodiments, the biomarker level is detected using any suitable method, including fluorescence, chemiluminescence, surface plasmon resonance, surface acoustic waves, mass spectrometry, infrared spectroscopy, Raman spectroscopy, atomic force microscopy, scanning tunneling microscopy, electrochemical detection methods, nuclear magnetic resonance, quantum dots, and the like.

[0050] As used herein, “sample processing” and “sample handling” refer to steps or procedures performed on a sample, such as a blood sample, after sample collection in order to prepare it for storage or analysis. In some embodiments, sample processing steps include centrifugation of the sample and decanting or aspirating the supernatant. In some embodiments, sample quality is assessed by determining the approximate durations of time that elapsed between sample processing steps. A sample processing time of or near zero means that each sample processing step was performed immediately with minimal time elapsed in between sample processing steps.

[0051] As used herein, “time-to-centrifugation” means the time elapsed between the moment that a blood sample is collected from a subject into a tube to the moment that the tube begins spinning in a centrifuge. In some embodiments, time-to-centrifugation is measured in hours. In some embodiments, ideal time-to-centrifugation for optimal sample quality is two hours or less. In some embodiments, the time-to-centrifugation is rounded to the nearest hour. In some embodiments, the time-to-centrifugation is rounded to the nearest half hour. Thus, in some embodiments, a time-to-centrifugation of less than a half hour or less than a quarter of an hour is rounded to zero.

[0052] As used herein, “time-to-decant” means the time elapsed between the moment that centrifugation of a sample is complete to the moment that the supernatant of the centrifuged sample begins to be decanted or aspirated from the precipitate. In some embodiments, time-to-decant is measured in hours. In some embodiments, ideal time-to-decant for optimal sample quality is less than one hour or less than half an hour, or less than a quarter hour. In some embodiments, the time-to-decant is rounded to the nearest hour. In some embodiments, the time-to-decant is rounded to the nearest half hour. Thus, in some embodiments, a time-to-decant of less than a half hour or less than a quarter of an hour is rounded to zero.

[0053] As used herein, “time-to-freeze” means the time elapsed between the moment that decanting or aspirating of a centrifuged sample is complete to the moment that the decanted or aspirated sample is placed in conditions at or under -20 °C. In some embodiments, time-to-freeze is measured in hours. In some embodiments, ideal time-to-decant for optimal sample quality is less than one hour or less than half an hour, or less than a quarter hour. In some embodiments, the time-to-freeze is rounded to the nearest hour. In some embodiments, the time-to-freeze is rounded to the nearest half hour. Thus, in some embodiments, a time-to-freeze of less than a half hour or less than a quarter of an hour is rounded to zero.

[0054] As used herein, “freeze-thaw cycle” means one round of exposing the sample to freezing temperatures (anywhere from approximately 0°C to -20 °C), or as low as -80°C or the temperature of liquid nitrogen), and then allowing the sample to thaw to liquid (for example, on ice or at room temperature).

[0055] “Solid support” refers herein to any substrate having a surface to which molecules may be attached, directly or indirectly, through either covalent or non-covalent bonds. A “solid support” can have a variety of physical formats, which can include, for example, a membrane; a chip (e.g., a protein chip); a slide (e.g., a glass slide or coverslip); a column; a hollow, solid, semi-solid, pore- or cavity- containing particle, such as, for example, a bead; a gel; a fiber, including a fiber optic material; a matrix; and a sample receptacle. Exemplary sample receptacles include sample wells, tubes, capillaries, vials, and any other vessel, groove or indentation capable of holding a sample. A sample receptacle can be contained on a multi-sample platform, such as a microtiter plate, slide, microfluidics device, and the like. A support can be composed of a natural or synthetic material, an organic or inorganic material. The composition of the solid support on which capture reagents are attached generally depends on the method of attachment (e.g., covalent attachment). Other exemplary receptacles include microdroplets and microfluidic controlled or bulk oil/aqueous emulsions within which assays and related manipulations can occur. Suitable solid supports include, for example, plastics, resins, polysaccharides, silica or silica-based materials, functionalized glass, modified silicon, carbon, metals, inorganic glasses, membranes, nylon, natural fibers (such as, for example, silk, wool and cotton), polymers, and the like. The material composing the solid support can include reactive groups such as, for example, carboxy, amino, or hydroxyl groups, which are used for attachment of the capture reagents. Polymeric solid supports can include, e.g., polystyrene, polyethylene glycol tetraphthalate, polyvinyl acetate, polyvinyl chloride, polyvinyl pyrrolidone, polyacrylonitrile, polymethyl

methacrylate, polytetrafluoroethylene, butyl rubber, styrenebutadiene rubber, natural rubber, polyethylene, polypropylene, (poly)tetrafluoroethylene, (poly)vinylidene fluoride, polycarbonate, and polymethylpentene. Suitable solid support particles that can be used include, e.g., encoded particles, such as Luminex®-type encoded particles, magnetic particles, and glass particles.

Exemplary Uses of Biomarkers

[0056] In various exemplary embodiments, methods are provided for evaluating or assessing sample quality by detecting one or more biomarker values corresponding to one or more biomarkers present in a sample from an individual, such as a blood, serum, or plasma sample, by any number of analytical methods, including any of the analytical methods described herein. These biomarkers are, for example, present at different levels in samples with differing quality. In some embodiments, the differing sample quality is due to differences in sample processing. Detection of the different levels of a biomarker in a sample can be used, for example, to estimate the time that elapsed between sample processing steps, such as time-to-centrifugation, time-to-decant, and/or time-to-freeze.

[0057] In addition to detecting biomarkers to assess sample quality, biomarkers can be used in diagnostic applications or to determine if a disease or condition is present in the subject from whom the sample was taken. In some embodiments, only samples that pass the quality assessment are further analyzed for diagnostic applications.

Detection and Determination of Biomarkers and Biomarker Levels

[0058] Levels of the biomarkers described herein can be detected using any of a variety of known analytical methods. In one embodiment, a biomarker level is detected using a capture reagent. In various embodiments, the capture reagent can be exposed to the biomarker in solution or can be exposed to the biomarker while the capture reagent is immobilized on a solid support. In other embodiments, the capture reagent contains a feature that is reactive with a secondary feature on a solid support. In these embodiments, the capture reagent can be exposed to the biomarker in solution, and then the feature on the capture reagent can be used in conjunction with the secondary feature on the solid support to immobilize the biomarker on the solid support. The capture reagent is selected based on the type of analysis to be conducted. Capture reagents include but are not limited to aptamers, antibodies, adnectins, ankyrins, other antibody mimetics and other protein scaffolds, autoantibodies, chimeras, small molecules, F(ab')₂ fragments, single chain antibody

fragments, Fv fragments, single chain Fv fragments, nucleic acids, lectins, ligand-binding receptors, affibodies, nanobodies, imprinted polymers, avimers, peptidomimetics, hormone receptors, cytokine receptors, and synthetic receptors, and modifications and fragments of these.

[0059] In some embodiments, a biomarker level is detected using a biomarker/capture reagent complex.

[0060] In some embodiments, the biomarker level is derived from the biomarker/capture reagent complex and is detected indirectly, such as, for example, as a result of a reaction that is subsequent to the biomarker/capture reagent interaction, but is dependent on the formation of the biomarker/capture reagent complex.

[0061] In some embodiments, the biomarker level is detected directly from the biomarker in a biological sample.

[0062] In some embodiments, biomarkers are detected using a multiplexed format that allows for the simultaneous detection of two or more biomarkers in a biological sample. In some embodiments of the multiplexed format, capture reagents are immobilized, directly or indirectly, covalently or non-covalently, in discrete locations on a solid support. In some embodiments, a multiplexed format uses discrete solid supports where each solid support has a unique capture reagent associated with that solid support, such as, for example quantum dots. In some embodiments, an individual device is used for the detection of each one of multiple biomarkers to be detected in a biological sample. Individual devices can be configured to permit each biomarker in the biological sample to be processed simultaneously. For example, a microtiter plate can be used such that each well in the plate is used to uniquely analyze one or more biomarkers to be detected in a biological sample.

[0063] In one or more of the foregoing embodiments, a fluorescent tag can be used to label a component of the biomarker/capture reagent complex to enable the detection of the biomarker level. In various embodiments, the fluorescent label can be conjugated to a capture reagent specific to any of the biomarkers described herein using known techniques, and the fluorescent label can then be used to detect the corresponding biomarker level. Suitable fluorescent labels include rare earth chelates, fluorescein and its derivatives, rhodamine and its derivatives, dansyl, allophycocyanin, PBXL-3, Qdot 605, Lissamine, phycoerythrin, Texas Red, and other such compounds.

[0064] In some embodiments, the fluorescent label is a fluorescent dye molecule. In some embodiments, the fluorescent dye molecule includes at least one substituted indolium ring system in which the substituent on the 3-carbon of the indolium ring contains a

chemically reactive group or a conjugated substance. In some embodiments, the dye molecule includes an AlexaFluor molecule, such as, for example, AlexaFluor 488, AlexaFluor 532, AlexaFluor 647, AlexaFluor 680, or AlexaFluor 700. In other embodiments, the dye molecule includes a first type and a second type of dye molecule, such as, e.g., two different AlexaFluor molecules. In some embodiments, the dye molecule includes a first type and a second type of dye molecule, and the two dye molecules have different emission spectra.

[0065] Fluorescence can be measured with a variety of instrumentation compatible with a wide range of assay formats. For example, spectrofluorimeters have been designed to analyze microtiter plates, microscope slides, printed arrays, cuvettes, etc. See Principles of Fluorescence Spectroscopy, by J.R. Lakowicz, Springer Science + Business Media, Inc., 2004. See Bioluminescence & Chemiluminescence: Progress & Current Applications; Philip E. Stanley and Larry J. Kricka editors, World Scientific Publishing Company, January 2002.

[0066] In one or more embodiments, a chemiluminescence tag can optionally be used to label a component of the biomarker/capture complex to enable the detection of a biomarker level. Suitable chemiluminescent materials include any of oxalyl chloride, Rodamin 6G, $\text{Ru}(\text{bipy})_3^{2+}$, TMAE (tetrakis(dimethylamino)ethylene), Pyrogallol (1,2,3-trihydroxybenzene), Lucigenin, peroxyoxalates, Aryl oxalates, Acridinium esters, dioxetanes, and others.

[0067] In some embodiments, the detection method includes an enzyme/substrate combination that generates a detectable signal that corresponds to the biomarker level. Generally, the enzyme catalyzes a chemical alteration of the chromogenic substrate which can be measured using various techniques, including spectrophotometry, fluorescence, and chemiluminescence. Suitable enzymes include, for example, luciferases, luciferin, malate dehydrogenase, urease, horseradish peroxidase (HRPO), alkaline phosphatase, beta-galactosidase, glucoamylase, lysozyme, glucose oxidase, galactose oxidase, and glucose-6-phosphate dehydrogenase, uricase, xanthine oxidase, lactoperoxidase, microperoxidase, and the like.

[0068] In some embodiments, the detection method can be a combination of fluorescence, chemiluminescence, radionuclide or enzyme/substrate combinations that generate a measurable signal. In some embodiments, multimodal signaling could have unique and advantageous characteristics in biomarker assay formats.

[0069] In some embodiments, the biomarker levels for the biomarkers described herein can be detected using any analytical methods including, singleplex aptamer assays, multiplexed aptamer assays, singleplex or multiplexed immunoassays, mRNA expression

profiling, miRNA expression profiling, mass spectrometric analysis, histological/cytological methods, etc. as discussed below.

Determination of Biomarker Levels using Aptamer-Based Assays

[0070] Assays directed to the detection and quantification of biomarker molecules in biological samples and other samples are important tools in scientific research and in the health care field. One class of such assays involves the use of a microarray that includes one or more aptamers immobilized on a solid support. The aptamers are each capable of binding to a target molecule in a highly specific manner and with very high affinity. See, e.g., U.S. Patent No. 5,475,096 entitled “Nucleic Acid Ligands”; see also, e.g., U.S. Patent No. 6,242,246, U.S. Patent No. 6,458,543, and U.S. Patent No. 6,503,715, each of which is entitled “Nucleic Acid Ligand Diagnostic Biochip”. Once the microarray is contacted with a sample, the aptamers bind to their respective target molecules present in the sample and thereby enable a determination of a biomarker level corresponding to a biomarker.

[0071] As used herein, an “aptamer” refers to a nucleic acid that has a specific binding affinity for a target molecule. It is recognized that affinity interactions are a matter of degree; however, in this context, the “specific binding affinity” of an aptamer for its target means that the aptamer binds to its target generally with a much higher degree of affinity than it binds to other components in a test sample. An “aptamer” is a set of copies of one type or species of nucleic acid molecule that comprises a particular nucleotide sequence. An aptamer can include any suitable number of nucleotides, including any number of chemically modified nucleotides. “Aptamers” refers to more than one such set of molecules. Different aptamers can have either the same or different numbers of nucleotides. Aptamers can be DNA or RNA or chemically modified nucleic acids and can be single-stranded, double-stranded, or contain double-stranded regions, and can include higher ordered structures. An aptamer can also be a photoaptamer, where a photoreactive or chemically reactive functional group is included in the aptamer to allow it to be covalently linked to its corresponding target. Any of the aptamer methods disclosed herein can include the use of two or more aptamers that specifically bind the same target molecule. As further described below, an aptamer may include a tag. If an aptamer includes a tag, all copies of the aptamer need not have the same tag. Moreover, if different aptamers each include a tag, these different aptamers can have either the same tag or a different tag.

[0072] An aptamer can be identified using any known method, including the SELEX process. Once identified, an aptamer can be prepared or synthesized in accordance with any known method, including chemical synthetic methods and enzymatic synthetic methods.

[0073] The terms “SELEX” and “SELEX process” are used interchangeably herein to refer generally to a combination of (1) the selection of aptamers that interact with a target molecule in a desirable manner, for example binding with high affinity to a protein, with (2) the amplification of those selected nucleic acids. The SELEX process can be used to identify aptamers with high affinity to a specific target or biomarker.

[0074] SELEX generally includes preparing a candidate mixture of nucleic acids, binding of the candidate mixture to the desired target molecule to form an affinity complex, separating the affinity complexes from the unbound candidate nucleic acids, separating and isolating the nucleic acid from the affinity complex, purifying the nucleic acid, and identifying a specific aptamer sequence. The process may include multiple rounds to further refine the affinity of the selected aptamer. The process can include amplification steps at one or more points in the process. See, e.g., U.S. Patent No. 5,475,096, entitled “Nucleic Acid Ligands”. The SELEX process can be used to generate an aptamer that covalently binds its target as well as an aptamer that non-covalently binds its target. See, e.g., U.S. Patent No. 5,705,337 entitled “Systematic Evolution of Nucleic Acid Ligands by Exponential Enrichment: Chemi-SELEX.”

[0075] The SELEX process can be used to identify high-affinity aptamers containing modified nucleotides that confer improved characteristics on the aptamer, such as, for example, improved *in vivo* stability or improved delivery characteristics. Examples of such modifications include chemical substitutions at the ribose and/or phosphate and/or base positions. SELEX process-identified aptamers containing modified nucleotides are described in U.S. Patent No. 5,660,985, entitled “High Affinity Nucleic Acid Ligands Containing Modified Nucleotides”, which describes oligonucleotides containing nucleotide derivatives chemically modified at the 5'- and 2'-positions of pyrimidines. U.S. Patent No. 5,580,737, see *supra*, describes highly specific aptamers containing one or more nucleotides modified with 2'-amino (2'-NH₂), 2'-fluoro (2'-F), and/or 2'-O-methyl (2'-Ome). See also, U.S. Patent Application Publication 20090098549, entitled “SELEX and PHOTOSELEX”, which describes nucleic acid libraries having expanded physical and chemical properties and their use in SELEX and photoSELEX.

[0076] SELEX can also be used to identify aptamers that have desirable off-rate characteristics. See U.S. Publication No. 20090004667, entitled “Method for Generating

Aptamers with Improved Off-Rates”, which describes improved SELEX methods for generating aptamers that can bind to target molecules. Methods for producing aptamers and photoaptamers having slower rates of dissociation from their respective target molecules are described. The methods involve contacting the candidate mixture with the target molecule, allowing the formation of nucleic acid-target complexes to occur, and performing a slow off-rate enrichment process wherein nucleic acid-target complexes with fast dissociation rates will dissociate and not reform, while complexes with slow dissociation rates will remain intact. Additionally, the methods include the use of modified nucleotides in the production of candidate nucleic acid mixtures to generate aptamers with improved off-rate performance. Nonlimiting exemplary modified nucleotides include, for example, the modified pyrimidines shown in Figs. 3-5. In some embodiments, an aptamer comprises at least one nucleotide with a modification, such as a base modification. In some embodiments, an aptamer comprises at least one nucleotide with a hydrophobic modification, such as a hydrophobic base modification, allowing for hydrophobic contacts with a target protein. Such hydrophobic contacts, in some embodiments, contribute to greater affinity and/or slower off-rate binding by the aptamer. Nonlimiting exemplary nucleotides with hydrophobic modifications are shown in Figure 3. In some embodiments, an aptamer comprises at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, or at least 10 nucleotides with hydrophobic modifications, where each hydrophobic modification may be the same or different from the others. In some embodiments, at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, or at least 10 hydrophobic modifications in an aptamer may be independently selected from the hydrophobic modifications shown in Figure 3.

[0077] In some embodiments, the aptamer is a slow off-rate aptamer. In some embodiments, a slow off-rate aptamer (including an aptamer comprising at least one nucleotide with a hydrophobic modification) has an off-rate ($t_{1/2}$) of ≥ 20 minutes, ≥ 30 minutes, ≥ 60 minutes, ≥ 90 minutes, ≥ 120 minutes, ≥ 150 minutes, ≥ 180 minutes, ≥ 210 minutes, or ≥ 240 minutes.

[0078] In some embodiments, an assay employs aptamers that include photoreactive functional groups that enable the aptamers to covalently bind or “photocrosslink” their target molecules. See, e.g., U.S. Patent No. 6,544,776 entitled “Nucleic Acid Ligand Diagnostic Biochip”. These photoreactive aptamers are also referred to as photoaptamers. See, e.g., U.S. Patent No. 5,763,177, U.S. Patent No. 6,001,577, and U.S. Patent No. 6,291,184, each of which is entitled “Systematic Evolution of Nucleic Acid Ligands by Exponential Enrichment:

Photoselection of Nucleic Acid Ligands and Solution SELEX”; see also, e.g., U.S. Patent No. 6,458,539, entitled “Photoselection of Nucleic Acid Ligands”. After the microarray is contacted with the sample and the photoaptamers have had an opportunity to bind to their target molecules, the photoaptamers are photoactivated, and the solid support is washed to remove any non-specifically bound molecules. Harsh wash conditions may be used, since target molecules that are bound to the photoaptamers are generally not removed, due to the covalent bonds created by the photoactivated functional group(s) on the photoaptamers. In this manner, the assay enables the detection of a biomarker level corresponding to a biomarker in the test sample.

[0079] In some assay formats, the aptamers are immobilized on the solid support prior to being contacted with the sample. Under certain circumstances, however, immobilization of the aptamers prior to contact with the sample may not provide an optimal assay. For example, pre-immobilization of the aptamers may result in inefficient mixing of the aptamers with the target molecules on the surface of the solid support, perhaps leading to lengthy reaction times and, therefore, extended incubation periods to permit efficient binding of the aptamers to their target molecules. Further, when photoaptamers are employed in the assay and depending upon the material utilized as a solid support, the solid support may tend to scatter or absorb the light used to effect the formation of covalent bonds between the photoaptamers and their target molecules. Moreover, depending upon the method employed, detection of target molecules bound to their aptamers can be subject to imprecision, since the surface of the solid support may also be exposed to and affected by any labeling agents that are used. Finally, immobilization of the aptamers on the solid support generally involves an aptamer-preparation step (i.e., the immobilization) prior to exposure of the aptamers to the sample, and this preparation step may affect the activity or functionality of the aptamers.

[0080] Aptamer assays that permit an aptamer to capture its target in solution and then employ separation steps that are designed to remove specific components of the aptamer-target mixture prior to detection have also been described (see U.S. Publication No. 20090042206, entitled “Multiplexed Analyses of Test Samples”). The described aptamer assay methods enable the detection and quantification of a non-nucleic acid target (e.g., a protein target) in a test sample by detecting and quantifying a nucleic acid (i.e., an aptamer). The described methods create a nucleic acid surrogate (i.e., the aptamer) for detecting and quantifying a non-nucleic acid target, thus allowing the wide variety of nucleic acid technologies, including amplification, to be applied to a broader range of desired targets, including protein targets.

[0081] Aptamers can be constructed to facilitate the separation of the assay components from an aptamer biomarker complex (or photoaptamer biomarker covalent complex) and permit isolation of the aptamer for detection and/or quantification. In one embodiment, these constructs can include a cleavable or releasable element within the aptamer sequence. In other embodiments, additional functionality can be introduced into the aptamer, for example, a labeled or detectable component, a spacer component, or a specific binding tag or immobilization element. For example, the aptamer can include a tag connected to the aptamer via a cleavable moiety, a label, a spacer component separating the label, and the cleavable moiety. In one embodiment, a cleavable element is a photocleavable linker. The photocleavable linker can be attached to a biotin moiety and a spacer section, can include an NHS group for derivatization of amines, and can be used to introduce a biotin group to an aptamer, thereby allowing for the release of the aptamer later in an assay method.

[0082] Homogenous assays, done with all assay components in solution, do not require separation of sample and reagents prior to the detection of signal. These methods are rapid and easy to use. These methods generate signal based on a molecular capture or binding reagent that reacts with its specific target. In some embodiments, the molecular capture reagents comprise one or more aptamers and/or antibodies or the like and the specific target of each of the one or more aptamers and/or antibodies or the like may be a biomarker shown in Table 1.

[0083] In some embodiments, a method for signal generation takes advantage of anisotropy signal change due to the interaction of a fluorophore-labeled capture reagent with its specific biomarker target. When the labeled capture reacts with its target, the increased molecular weight causes the rotational motion of the fluorophore attached to the complex to become much slower changing the anisotropy value. By monitoring the anisotropy change, binding events may be used to quantitatively measure the biomarkers in solutions. Other methods include fluorescence polarization assays, molecular beacon methods, time resolved fluorescence quenching, chemiluminescence, fluorescence resonance energy transfer, and the like.

[0084] An exemplary solution-based aptamer assay that can be used to detect a biomarker level in a biological sample includes the following: (a) preparing a mixture by contacting the biological sample with an aptamer that includes a first tag and has a specific affinity for the biomarker, wherein an aptamer affinity complex is formed when the biomarker is present in the sample; (b) exposing the mixture to a first solid support including a first capture element, and allowing the first tag to associate with the first capture element;

(c) removing any components of the mixture not associated with the first solid support; (d) attaching a second tag to the biomarker component of the aptamer affinity complex; (e) releasing the aptamer affinity complex from the first solid support; (f) exposing the released aptamer affinity complex to a second solid support that includes a second capture element and allowing the second tag to associate with the second capture element; (g) removing any non-complexed aptamer from the mixture by partitioning the non-complexed aptamer from the aptamer affinity complex; (h) eluting the aptamer from the solid support; and (i) detecting the biomarker by detecting the aptamer component of the aptamer affinity complex.

[0085] Any means known in the art can be used to detect a biomarker value by detecting the aptamer component of an aptamer affinity complex. A number of different detection methods can be used to detect the aptamer component of an affinity complex, such as, for example, hybridization assays, mass spectroscopy, or QPCR. In some embodiments, nucleic acid sequencing methods can be used to detect the aptamer component of an aptamer affinity complex and thereby detect a biomarker value. Briefly, a test sample can be subjected to any kind of nucleic acid sequencing method to identify and quantify the sequence or sequences of one or more aptamers present in the test sample. In some embodiments, the sequence includes the entire aptamer molecule or any portion of the molecule that may be used to uniquely identify the molecule. In other embodiments, the identifying sequencing is a specific sequence added to the aptamer; such sequences are often referred to as “tags,” “barcodes,” or “zipcodes.” In some embodiments, the sequencing method includes enzymatic steps to amplify the aptamer sequence or to convert any kind of nucleic acid, including RNA and DNA that contain chemical modifications to any position, to any other kind of nucleic acid appropriate for sequencing.

[0086] In some embodiments, the sequencing method includes one or more cloning steps. In other embodiments the sequencing method includes a direct sequencing method without cloning.

[0087] In some embodiments, the sequencing method includes a directed approach with specific primers that target one or more aptamers in the test sample. In other embodiments, the sequencing method includes a shotgun approach that targets all aptamers in the test sample.

[0088] In some embodiments, the sequencing method includes enzymatic steps to amplify the molecule targeted for sequencing. In other embodiments, the sequencing method directly sequences single molecules. An exemplary nucleic acid sequencing-based method that can be used to detect a biomarker value corresponding to a biomarker in a biological

sample includes the following: (a) converting a mixture of aptamers that contain chemically modified nucleotides to unmodified nucleic acids with an enzymatic step; (b) shotgun sequencing the resulting unmodified nucleic acids with a massively parallel sequencing platform such as, for example, the 454 Sequencing System (454 Life Sciences/Roche), the Illumina Sequencing System (Illumina), the ABI SOLiD Sequencing System (Applied Biosystems), the HeliScope Single Molecule Sequencer (Helicos Biosciences), or the Pacific Biosciences Real Time Single-Molecule Sequencing System (Pacific BioSciences) or the Polonator G Sequencing System (Dover Systems); and (c) identifying and quantifying the aptamers present in the mixture by specific sequence and sequence count.

[0089] A nonlimiting exemplary method of detecting biomarkers in a biological sample using aptamers is described in Example 1. *See also* Kraemer et al., 2011, *PloS One* 6(10): e26332.

Determination of Biomarker Levels using Immunoassays

[0090] Immunoassay methods are based on the reaction of an antibody to its corresponding target or analyte and can detect the analyte in a sample depending on the specific assay format. To improve specificity and sensitivity of an assay method based on affibody-reactivity, monoclonal antibodies and fragments thereof are often used because of their specific epitope recognition. Polyclonal antibodies have also been successfully used in various immunoassays because of their increased affinity for the target as compared to monoclonal antibodies. Immunoassays have been designed for use with a wide range of biological sample matrices. Immunoassay formats have been designed to provide qualitative, semi-quantitative, and quantitative results.

[0091] Quantitative results are generated through the use of a standard curve created with known concentrations of the specific analyte to be detected. The response or signal from an unknown sample is plotted onto the standard curve, and a quantity or level corresponding to the target in the unknown sample is established.

[0092] Numerous immunoassay formats have been designed. ELISA or EIA can be quantitative for the detection of an analyte. This method relies on attachment of a label to either the analyte or the antibody and the label component includes, either directly or indirectly, an enzyme. ELISA tests may be formatted for direct, indirect, competitive, or sandwich detection of the analyte. Other methods rely on labels such as, for example, radioisotopes (I^{125}) or fluorescence. Additional techniques include, for example, agglutination, nephelometry, turbidimetry, Western blot, immunoprecipitation,

immunocytochemistry, immunohistochemistry, flow cytometry, Luminex assay, and others (see *ImmunoAssay: A Practical Guide*, edited by Brian Law, published by Taylor & Francis, Ltd., 2005 edition).

[0093] Exemplary assay formats include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay, fluorescent, chemiluminescence, and fluorescence resonance energy transfer (FRET) or time resolved-FRET (TR-FRET) immunoassays. Examples of procedures for detecting biomarkers include biomarker immunoprecipitation followed by quantitative methods that allow size and peptide level discrimination, such as gel electrophoresis, capillary electrophoresis, planar electrochromatography, and the like.

[0094] Methods of detecting and/or for quantifying a detectable label or signal generating material depend on the nature of the label. The products of reactions catalyzed by appropriate enzymes (where the detectable label is an enzyme; see above) can be, without limitation, fluorescent, luminescent, or radioactive or they may absorb visible or ultraviolet light. Examples of detectors suitable for detecting such detectable labels include, without limitation, x-ray film, radioactivity counters, scintillation counters, spectrophotometers, colorimeters, fluorometers, luminometers, and densitometers.

[0095] Any of the methods for detection can be performed in any format that allows for any suitable preparation, processing, and analysis of the reactions. This can be, for example, in multi-well assay plates (e.g., 96 wells or 386 wells) or using any suitable array or microarray. Stock solutions for various agents can be made manually or robotically, and all subsequent pipetting, diluting, mixing, distribution, washing, incubating, sample readout, data collection and analysis can be done robotically using commercially available analysis software, robotics, and detection instrumentation capable of detecting a detectable label.

Determination of Biomarker Levels using Gene Expression Profiling

[0096] Measuring mRNA in a biological sample may, in some embodiments, be used as a surrogate for detection of the level of the corresponding protein in the biological sample. Thus, in some embodiments, a biomarker or biomarker panel described herein can be detected by detecting the appropriate RNA.

[0097] In some embodiments, mRNA expression levels are measured by reverse transcription quantitative polymerase chain reaction (RT-PCR followed with qPCR). RT-PCR is used to create a cDNA from the mRNA. The cDNA may be used in a qPCR assay to produce fluorescence as the DNA amplification process progresses. By comparison to a standard curve, qPCR can produce an absolute measurement such as number of copies of

mRNA per cell. Northern blots, microarrays, Invader assays, and RT-PCR combined with capillary electrophoresis have all been used to measure expression levels of mRNA in a sample. See Gene Expression Profiling: Methods and Protocols, Richard A. Shimkets, editor, Humana Press, 2004.

Detection of Biomarkers Using *In Vivo* Molecular Imaging Technologies

[0098] In some embodiments, a biomarker described herein may be used in molecular imaging tests. For example, an imaging agent can be coupled to a capture reagent, which can be used to detect the biomarker *in vivo*.

[0099] *In vivo* imaging technologies provide non-invasive methods for determining the state of a particular disease in the body of an individual. For example, entire portions of the body, or even the entire body, may be viewed as a three dimensional image, thereby providing valuable information concerning morphology and structures in the body. Such technologies may be combined with the detection of the biomarkers described herein to provide information concerning the biomarker *in vivo*.

[00100] The use of *in vivo* molecular imaging technologies is expanding due to various advances in technology. These advances include the development of new contrast agents or labels, such as radiolabels and/or fluorescent labels, which can provide strong signals within the body; and the development of powerful new imaging technology, which can detect and analyze these signals from outside the body, with sufficient sensitivity and accuracy to provide useful information. The contrast agent can be visualized in an appropriate imaging system, thereby providing an image of the portion or portions of the body in which the contrast agent is located. The contrast agent may be bound to or associated with a capture reagent, such as an aptamer or an antibody, for example, and/or with a peptide or protein, or an oligonucleotide (for example, for the detection of gene expression), or a complex containing any of these with one or more macromolecules and/or other particulate forms.

[00101] The contrast agent may also feature a radioactive atom that is useful in imaging. Suitable radioactive atoms include technetium-99m or iodine-123 for scintigraphic studies. Other readily detectable moieties include, for example, spin labels for magnetic resonance imaging (MRI) such as, for example, iodine-123 again, iodine-131, indium-111, fluorine-19, carbon-13, nitrogen-15, oxygen-17, gadolinium, manganese or iron. Such labels are well known in the art and could easily be selected by one of ordinary skill in the art.

[00102] Standard imaging techniques include but are not limited to magnetic resonance imaging, computed tomography scanning, positron emission tomography (PET), single

photon emission computed tomography (SPECT), and the like. For diagnostic *in vivo* imaging, the type of detection instrument available is a major factor in selecting a given contrast agent, such as a given radionuclide and the particular biomarker that it is used to target (protein, mRNA, and the like). The radionuclide chosen typically has a type of decay that is detectable by a given type of instrument. Also, when selecting a radionuclide for *in vivo* diagnosis, its half-life should be long enough to enable detection at the time of maximum uptake by the target tissue but short enough that deleterious radiation of the host is minimized.

[00103] Exemplary imaging techniques include but are not limited to PET and SPECT, which are imaging techniques in which a radionuclide is synthetically or locally administered to an individual. The subsequent uptake of the radiotracer is measured over time and used to obtain information about the targeted tissue and the biomarker. Because of the high-energy (gamma-ray) emissions of the specific isotopes employed and the sensitivity and sophistication of the instruments used to detect them, the two-dimensional distribution of radioactivity may be inferred from outside of the body.

[00104] Commonly used positron-emitting nuclides in PET include, for example, carbon-11, nitrogen-13, oxygen-15, and fluorine-18. Isotopes that decay by electron capture and/or gamma-emission are used in SPECT and include, for example iodine-123 and technetium-99m. An exemplary method for labeling amino acids with technetium-99m is the reduction of pertechnetate ion in the presence of a chelating precursor to form the labile technetium-99m-precursor complex, which, in turn, reacts with the metal binding group of a bifunctionally modified chemotactic peptide to form a technetium-99m-chemotactic peptide conjugate.

[00105] Antibodies are frequently used for such *in vivo* imaging diagnostic methods. The preparation and use of antibodies for *in vivo* diagnosis is well known in the art. Similarly, aptamers may be used for such *in vivo* imaging diagnostic methods. For example, an aptamer that was used to identify a particular biomarker described herein may be appropriately labeled and injected into an individual to detect the biomarker *in vivo*. The label used will be selected in accordance with the imaging modality to be used, as previously described. Aptamer-directed imaging agents could have unique and advantageous characteristics relating to tissue penetration, tissue distribution, kinetics, elimination, potency, and selectivity as compared to other imaging agents.

[00106] Such techniques may also optionally be performed with labeled oligonucleotides, for example, for detection of gene expression through imaging with

antisense oligonucleotides. These methods are used for in situ hybridization, for example, with fluorescent molecules or radionuclides as the label. Other methods for detection of gene expression include, for example, detection of the activity of a reporter gene.

[00107] Another general type of imaging technology is optical imaging, in which fluorescent signals within the subject are detected by an optical device that is external to the subject. These signals may be due to actual fluorescence and/or to bioluminescence. Improvements in the sensitivity of optical detection devices have increased the usefulness of optical imaging for in vivo diagnostic assays.

[00108] For a review of other techniques, see N. Blow, *Nature Methods*, 6, 465-469, 2009.

Determination of Biomarker Levels using Mass Spectrometry Methods

[00109] A variety of configurations of mass spectrometers can be used to detect biomarker levels. Several types of mass spectrometers are available or can be produced with various configurations. In general, a mass spectrometer has the following major components: a sample inlet, an ion source, a mass analyzer, a detector, a vacuum system, and instrument-control system, and a data system. Difference in the sample inlet, ion source, and mass analyzer generally define the type of instrument and its capabilities. For example, an inlet can be a capillary-column liquid chromatography source or can be a direct probe or stage such as used in matrix-assisted laser desorption. Common ion sources are, for example, electrospray, including nanospray and microspray or matrix-assisted laser desorption. Common mass analyzers include a quadrupole mass filter, ion trap mass analyzer and time-of-flight mass analyzer. Additional mass spectrometry methods are well known in the art (see Burlingame et al. *Anal. Chem.* 70:647 R-716R (1998); Kinter and Sherman, New York (2000)).

[00110] Protein biomarkers and biomarker levels can be detected and measured by any of the following: electrospray ionization mass spectrometry (ESI-MS), ESI-MS/MS, ESI-MS/(MS)_n, matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS), surface-enhanced laser desorption/ionization time-of-flight mass spectrometry (SELDI-TOF-MS), desorption/ionization on silicon (DIOS), secondary ion mass spectrometry (SIMS), quadrupole time-of-flight (Q-TOF), tandem time-of-flight (TOF/TOF) technology, called ultraflex III TOF/TOF, atmospheric pressure chemical ionization mass spectrometry (APCI-MS), APCI-MS/MS, APCI-(MS)_N, atmospheric pressure photoionization mass spectrometry (APPI-MS), APPI-MS/MS, and APPI-(MS)_N,

quadrupole mass spectrometry, Fourier transform mass spectrometry (FTMS), quantitative mass spectrometry, and ion trap mass spectrometry.

[00111] Sample preparation and processing strategies are used to label and enrich samples before mass spectroscopic characterization of protein biomarkers and determination biomarker levels. Labeling methods include but are not limited to isobaric tag for relative and absolute quantitation (iTRAQ) and stable isotope labeling with amino acids in cell culture (SILAC). Capture reagents used to selectively enrich samples for candidate biomarker proteins prior to mass spectroscopic analysis include but are not limited to aptamers, antibodies, nucleic acid probes, chimeras, small molecules, an F(ab')₂ fragment, a single chain antibody fragment, an Fv fragment, a single chain Fv fragment, a nucleic acid, a lectin, a ligand-binding receptor, affibodies, nanobodies, ankyrins, domain antibodies, alternative antibody scaffolds (e.g. diabodies etc) imprinted polymers, avimers, peptidomimetics, peptoids, peptide nucleic acids, threose nucleic acid, a hormone receptor, a cytokine receptor, and synthetic receptors, and modifications and fragments of these.

Determination of Biomarker Levels using a Proximity Ligation Assay

[00112] A proximity ligation assay can be used to determine biomarker values. Briefly, a test sample is contacted with a pair of affinity probes that may be a pair of antibodies or a pair of aptamers, with each member of the pair extended with an oligonucleotide. The targets for the pair of affinity probes may be two distinct determinates on one protein or one determinate on each of two different proteins, which may exist as homo- or hetero-multimeric complexes. When probes bind to the target determinates, the free ends of the oligonucleotide extensions are brought into sufficiently close proximity to hybridize together. The hybridization of the oligonucleotide extensions is facilitated by a common connector oligonucleotide which serves to bridge together the oligonucleotide extensions when they are positioned in sufficient proximity. Once the oligonucleotide extensions of the probes are hybridized, the ends of the extensions are joined together by enzymatic DNA ligation.

[00113] Each oligonucleotide extension comprises a primer site for PCR amplification. Once the oligonucleotide extensions are ligated together, the oligonucleotides form a continuous DNA sequence which, through PCR amplification, reveals information regarding the identity and amount of the target protein, as well as, information regarding protein-protein interactions where the target determinates are on two different proteins. Proximity ligation can provide a highly sensitive and specific assay for real-time protein concentration and interaction information through use of real-time PCR. Probes that do not bind the

determinates of interest do not have the corresponding oligonucleotide extensions brought into proximity and no ligation or PCR amplification can proceed, resulting in no signal being produced.

[00114] The foregoing assays enable the detection of biomarker values that are useful in methods for assessing sample quality, wherein the methods comprise detecting, in a biological sample from an individual, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, or all least 8 biomarkers selected from the biomarkers in Table 1. A classification, as described below, using the biomarker levels indicates whether the sample has acceptable quality for use in subsequent analysis. In accordance with any of the methods described herein, biomarker levels can be detected and classified individually or they can be detected and classified collectively, as for example in a multiplex assay format.

Classification of Biomarkers and Calculation of Sample Freeze-Thaw Cycles

[00115] In some embodiments, a biomarker “signature” for a given sample quality test contains a set of biomarkers, each biomarker having characteristic levels in samples of acceptable quality or in sample of poor quality. Characteristic levels, in some embodiments, may refer to the mean or average of the biomarker levels for the samples in a particular group. In some embodiments, a method described herein can be used to assign a sample into one of two groups, either passing quality assessment or failing quality assessment.

[00116] The assignment of a sample into one of two or more groups is known as classification, and the procedure used to accomplish this assignment is known as a classifier or a classification method. Classification methods may also be referred to as scoring methods. There are many classification methods that can be used to construct a classifier from a set of biomarker levels. In some instances, classification methods are performed using supervised learning techniques in which a data set is collected using samples from two (or more, for multiple classification states) distinct groups one wishes to distinguish. Since the class (group or population) to which each sample belongs is known in advance for each sample, the classification method can be trained to give the desired classification response. It is also possible to use unsupervised learning techniques to produce a quality classifier.

[00117] Common approaches for developing classifiers include decision trees; bagging + boosting + forests; rule inference based learning; Parzen Windows; linear models; logistic; neural network methods; unsupervised clustering; K-means; hierarchical ascending/descending; semi-supervised learning; prototype methods; nearest neighbor; kernel density estimation; support vector machines; hidden Markov models; Boltzmann Learning; and classifiers may be combined either simply or in ways which minimize particular objective

functions. For a review, see, e.g., Pattern Classification, R.O. Duda, et al., editors, John Wiley & Sons, 2nd edition, 2001; see also, The Elements of Statistical Learning – Data Mining, Inference, and Prediction, T. Hastie, et al., editors, Springer Science+Business Media, LLC, 2nd edition, 2009.

[00118] To produce a classifier using supervised learning techniques, a set of samples called training data are obtained. In the context of quality tests, training data includes samples from the distinct groups (classes) to which unknown samples will later be assigned. For example, samples that underwent processing with differing, set times between processing steps can constitute training data to develop a classifier that can classify unknown samples as either passing or failing quality assessment based on the elapsed time between sample processing steps. The development of the classifier from the training data is known as training the classifier. Specific details on classifier training depend on the nature of the supervised learning technique. Training a naïve Bayesian classifier is an example of such a supervised learning technique (see, e.g., Pattern Classification, R.O. Duda, et al., editors, John Wiley & Sons, 2nd edition, 2001; see also, The Elements of Statistical Learning – Data Mining, Inference, and Prediction, T. Hastie, et al., editors, Springer Science+Business Media, LLC, 2nd edition, 2009). Training of a naïve Bayesian classifier is described, e.g., in U.S. Publication Nos: 2012/0101002 and 2012/0077695.

[00119] Since typically there are many more potential biomarker levels than samples in a training set, care must be used to avoid over-fitting. Over-fitting occurs when a statistical model describes random error or noise instead of the underlying relationship. Over-fitting can be avoided in a variety of way, including, for example, by limiting the number of biomarkers used in developing the classifier, by assuming that the biomarker responses are independent of one another, by limiting the complexity of the underlying statistical model employed, and by ensuring that the underlying statistical model conforms to the data.

[00120] An illustrative example of the development of a test using a set of biomarkers includes the application of a naïve Bayes classifier, a simple probabilistic classifier based on Bayes theorem with strict independent treatment of the biomarkers. Each biomarker is described by a class-dependent probability density function (pdf) for the measured RFU values or log RFU (relative fluorescence units) values in each class. The joint pdfs for the set of biomarkers in one class is assumed to be the product of the individual class-dependent pdfs for each biomarker. Training a naïve Bayes classifier in this context amounts to assigning parameters (“parameterization”) to characterize the class dependent pdfs. Any underlying

model for the class-dependent pdfs may be used, but the model should generally conform to the data observed in the training set.

[00121] The performance of the naïve Bayes classifier is dependent upon the number and quality of the biomarkers used to construct and train the classifier. A single biomarker will perform in accordance with its KS-distance (Kolmogorov-Smirnov). The addition of subsequent biomarkers with good KS distances (>0.3 , for example) will, in general, improve the classification performance if the subsequently added biomarkers are independent of the first biomarker. Using the sensitivity plus specificity as a classifier score, many high scoring classifiers can be generated with a variation of a greedy algorithm. (A greedy algorithm is any algorithm that follows the problem solving metaheuristic of making the locally optimal choice at each stage with the hope of finding the global optimum.)

[00122] Another way to depict classifier performance is through a receiver operating characteristic (ROC), or simply ROC curve or ROC plot. The ROC is a graphical plot of the sensitivity, or true positive rate, vs. false positive rate ($1 - \text{specificity}$ or $1 - \text{true negative rate}$), for a binary classifier system as its discrimination threshold is varied. The ROC can also be represented equivalently by plotting the fraction of true positives out of the positives (TPR = true positive rate) vs. the fraction of false positives out of the negatives (FPR = false positive rate). Also known as a Relative Operating Characteristic curve, because it is a comparison of two operating characteristics (TPR & FPR) as the criterion changes. The area under the ROC curve (AUC) is commonly used as a summary measure of diagnostic accuracy. It can take values from 0.0 to 1.0. The AUC has an important statistical property: the AUC of a classifier is equivalent to the probability that the classifier will rank a randomly chosen positive instance higher than a randomly chosen negative instance (Fawcett T, 2006. An introduction to ROC analysis. *Pattern Recognition Letters* .27: 861–874). This is equivalent to the Wilcoxon test of ranks (Hanley, J.A., McNeil, B.J., 1982. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 143, 29–36.). Another way of describing performance of a diagnostic test in relation to a known reference standard is the net reclassification index: the ability of the new test to correctly upgrade or downgrade risk when compared with the reference standard test. See, e.g., Pencina et al., 2011, *Stat. Med.* 30: 11-21. While the AUC under the ROC curve is optimal for assessing performance of a 2-class classifier, stratified and personalized medicine relies upon the inference that the population contains more classes than 2. For such comparisons the hazard ratio of the upper vs. lower quartiles (or other stratifications such as deciles) can be used more appropriately.

Kits

[00123] Any combination of the biomarkers described herein can be detected using a suitable kit, such as for use in performing the methods disclosed herein. Furthermore, any kit can contain one or more detectable labels as described herein, such as a fluorescent moiety, etc.

[00124] In some embodiments, a kit includes (a) one or more capture reagents (such as, for example, at least one aptamer or antibody) for detecting one or more biomarkers in a biological sample, wherein the biomarkers include at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, or all 8 biomarkers selected from the biomarkers in Table 1; and optionally (b) one or more software or computer program products for classifying the sample was obtained as either passing or failing quality assessment or for determining the approximate time or times of sample processing steps, as further described herein. Alternatively, rather than one or more computer program products, one or more instructions for manually performing the above steps by a human can be provided.

[00125] In some embodiments, a kit comprises a solid support, at least one capture reagent, and a signal generating material. The kit can also include instructions for using the devices and reagents, handling the sample, and analyzing the data. Further the kit may be used with a computer system or software to analyze and report the result of the analysis of the biological sample.

[00126] The kit may further comprise reagents for diagnostic analysis of samples, particularly samples that pass quality assessment.

[00127] The kits can also contain one or more reagents (e.g., solubilization buffers, detergents, washes, or buffers) for processing a biological sample. Any of the kits described herein can also include, e.g., buffers, blocking agents, mass spectrometry matrix materials, antibody capture agents, positive control samples, negative control samples, software and information such as protocols, guidance and reference data.

[00128] In some embodiments, the kits comprise PCR primers for one or more aptamers specific to biomarkers described herein. In some embodiments, a kit may further include instructions for use and correlation of the biomarkers with estimation of sample processing times and/or sample quality. In some embodiments, a kit may also include a DNA array containing the complement of one or more of the aptamers specific for the biomarkers described herein, reagents, and/or enzymes for amplifying or isolating sample DNA. In some embodiments, kits may include reagents for real-time PCR, for example, TaqMan probes and/or primers, and enzymes.

[00129] For example, a kit can comprise (a) reagents comprising at least one capture reagent for determining the level of one or more biomarkers in a test sample, and optionally (b) one or more algorithms or computer programs for performing the steps of comparing the amount of each biomarker quantified in the test sample to one or more predetermined cutoffs. In some embodiments, an algorithm or computer program assigns a score for each biomarker quantified based on the comparison and, in some embodiments, combines the assigned scores for each biomarker quantified to obtain a total score. Further, in some embodiments, an algorithm or computer program compares the total score with a predetermined score, and uses the comparison to determine whether a sample passes or fails quality assessment. Alternatively, rather than one or more algorithms or computer programs, one or more instructions for manually performing the above steps by a human can be provided.

Biomarker Panels

[00130] In some embodiments, one or more of the biomarkers listed in Table 1 are detected. In some embodiments, the one or more biomarkers listed in Table 1 are detected in a serum sample from a subject. In some embodiments, all of the biomarkers listed in Table 1 are detected. In some embodiments, the level of each protein listed in Table 1 is detected. In some embodiments, the detecting of the one or more biomarkers or all of the biomarkers is performed in order to determine the number of times a sample is frozen and thawed.

Table 1: Panel for Freeze-thaw cycles (serum)

UniProt	Target	EntrezGeneSymbol	TargetFullName
P09960	LKHA4	LTA4H	Leukotriene A-4 hydrolase
O60739	EIF1B	EIF1B	Eukaryotic translation initiation factor 1b
P46060	RAGP1	RANGAP1	Ran GTPase-activating protein 1
Q8WVM8	SCFD1	SCFD1	Sec1 family domain-containing protein 1
Q8IV08	PLD3	PLD3	Phospholipase D3
Q12907	Lectin, mannose-binding 2	LMAN2	Vesicular integral-membrane protein VIP36
Q13418	ILK1	ILK	Integrin-linked protein kinase
P46777	RPL5	RPL5	60S ribosomal protein L5

Computer Methods and Software

[00131] A method for assessing sample quality, such as the number of freeze-thaw cycles, can comprise the following: 1) obtain a biological sample, such as a sample that has already undergone sample processing; 2) perform an analytical method to detect and measure a biomarker or set of biomarkers in a panel in the biological sample; 3) optionally perform any data normalization or standardization; 4) determine each biomarker level; and 5) report the results. In some embodiments, the results are calibrated to the sample type. In some embodiments, the biomarker levels are combined in some way and a single value for the combined biomarker levels is reported. In this approach, in some embodiments, the score may be a single number or identity determined from the integration of all the biomarkers that is compared to a pre-set threshold value that is an indication of satisfactory (passing) or unsatisfactory (failing) quality. Or the predictive score may be a series of bars that each represent a biomarker value and the pattern of the responses may be compared to a pre-set pattern for determination of the satisfactory (passing) or unsatisfactory (failing) quality.

[00132] At least some embodiments of the methods described herein can be implemented with the use of a computer. An example of a computer system 100 is shown in Figure 1. With reference to Figure 1, system 100 is shown comprised of hardware elements that are electrically coupled via bus 108, including a processor 101, input device 102, output device 103, storage device 104, computer-readable storage media reader 105a, communications system 106, processing acceleration (e.g., DSP or special-purpose processors) 107 and memory 109. Computer-readable storage media reader 105a is further coupled to computer-readable storage media 105b, the combination comprehensively representing remote, local, fixed and/or removable storage devices plus storage media, memory, etc. for temporarily and/or more permanently containing computer-readable information, which can include storage device 104, memory 109 and/or any other such accessible system 100 resource. System 100 also comprises software elements (shown as being currently located within working memory 191) including an operating system 192 and other code 193, such as programs, data and the like.

[00133] With respect to Figure 1, system 100 has extensive flexibility and configurability. Thus, for example, a single architecture might be utilized to implement one or more servers that can be further configured in accordance with currently desirable protocols, protocol variations, extensions, etc. However, it will be apparent to those skilled in

the art that embodiments may well be utilized in accordance with more specific application requirements. For example, one or more system elements might be implemented as sub-elements within a system 100 component (e.g., within communications system 106). Customized hardware might also be utilized and/or particular elements might be implemented in hardware, software or both. Further, while connection to other computing devices such as network input/output devices (not shown) may be employed, it is to be understood that wired, wireless, modem, and/or other connection or connections to other computing devices might also be utilized.

[00134] In one aspect, the system can comprise a database containing features of biomarkers characteristic of sample quality. The biomarker data (or biomarker information) can be utilized as an input to the computer for use as part of a computer implemented method. The biomarker data can include the data as described herein.

[00135] In one aspect, the system further comprises one or more devices for providing input data to the one or more processors.

[00136] The system further comprises a memory for storing a data set of ranked data elements.

[00137] In another aspect, the device for providing input data comprises a detector for detecting the characteristic of the data element, e.g., such as a mass spectrometer or gene chip reader.

[00138] The system additionally may comprise a database management system. User requests or queries can be formatted in an appropriate language understood by the database management system that processes the query to extract the relevant information from the database of training sets.

[00139] The system may be connectable to a network to which a network server and one or more clients are connected. The network may be a local area network (LAN) or a wide area network (WAN), as is known in the art. Preferably, the server includes the hardware necessary for running computer program products (e.g., software) to access database data for processing user requests.

[00140] The system may include an operating system (e.g., UNIX or Linux) for executing instructions from a database management system. In one aspect, the operating system can operate on a global communications network, such as the internet, and utilize a global communications network server to connect to such a network.

[00141] The system may include one or more devices that comprise a graphical display interface comprising interface elements such as buttons, pull down menus, scroll bars, fields

for entering text, and the like as are routinely found in graphical user interfaces known in the art. Requests entered on a user interface can be transmitted to an application program in the system for formatting to search for relevant information in one or more of the system databases. Requests or queries entered by a user may be constructed in any suitable database language.

[00142] The graphical user interface may be generated by a graphical user interface code as part of the operating system and can be used to input data and/or to display inputted data. The result of processed data can be displayed in the interface, printed on a printer in communication with the system, saved in a memory device, and/or transmitted over the network or can be provided in the form of the computer readable medium.

[00143] The system can be in communication with an input device for providing data regarding data elements to the system (e.g., expression values). In one aspect, the input device can include a gene expression profiling system including, e.g., a mass spectrometer, gene chip or array reader, and the like.

[00144] The methods and apparatus for analyzing sample quality biomarker information according to various embodiments may be implemented in any suitable manner, for example, using a computer program operating on a computer system. A conventional computer system comprising a processor and a random access memory, such as a remotely-accessible application server, network server, personal computer or workstation may be used. Additional computer system components may include memory devices or information storage systems, such as a mass storage system and a user interface, for example a conventional monitor, keyboard and tracking device. The computer system may be a stand-alone system or part of a network of computers including a server and one or more databases.

[00145] The sample quality biomarker analysis system can provide functions and operations to complete data analysis, such as data gathering, processing, analysis, reporting and/or sample quality identification. For example, in one embodiment, the computer system can execute the computer program that may receive, store, search, analyze, and report information relating to the sample quality assessment biomarkers. The computer program may comprise multiple modules performing various functions or operations, such as a processing module for processing raw data and generating supplemental data and an analysis module for analyzing raw data and supplemental data to generate a sample quality status and/or estimated sample processing time calculation. Calculation of sample processing time may optionally comprise generating or collecting additional information.

[00146] Some embodiments described herein can be implemented so as to include a computer program product. A computer program product may include a computer readable medium having computer readable program code embodied in the medium for causing an application program to execute on a computer with a database.

[00147] As used herein, a “computer program product” refers to an organized set of instructions in the form of natural or programming language statements that are contained on a physical media of any nature (e.g., written, electronic, magnetic, optical or otherwise) and that may be used with a computer or other automated data processing system. Such programming language statements, when executed by a computer or data processing system, cause the computer or data processing system to act in accordance with the particular content of the statements. Computer program products include without limitation: programs in source and object code and/or test or data libraries embedded in a computer readable medium. Furthermore, the computer program product that enables a computer system or data processing equipment device to act in pre-selected ways may be provided in a number of forms, including, but not limited to, original source code, assembly code, object code, machine language, encrypted or compressed versions of the foregoing and any and all equivalents.

[00148] In one aspect, a computer program product is provided for evaluation of sample quality. The computer program product includes a computer readable medium embodying program code executable by a processor of a computing device or system, the program code comprising: code that retrieves data attributed to a biological sample from an individual, wherein the data comprises biomarker levels that each correspond to one of the biomarkers in Table 1 and code that executes a classification method that indicates a sample quality status as a function of the biomarker levels.

[00149] In still another aspect, a computer program product is provided for determining sample processing time or times. The computer program product includes a computer readable medium embodying program code executable by a processor of a computing device or system, the program code comprising: code that retrieves data attributed to a biological sample from an individual, wherein the data comprises a biomarker value corresponding to at least one biomarker in the biological sample selected from the biomarkers provided in Table 1 and code that executes a classification method that indicates a sample quality status as a function of the biomarker levels.

[00150] While various embodiments have been described as methods or apparatuses, it should be understood that embodiments can be implemented through code coupled with a

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

[00151] It is also envisioned that embodiments could be accomplished as computer signals embodied in a carrier wave, as well as signals (e.g., electrical and optical) propagated through a transmission medium. Thus, the various types of information discussed above could be formatted in a structure, such as a data structure, and transmitted as an electrical signal through a transmission medium or stored on a computer readable medium.

[00152] It is also noted that many of the structures, materials, and acts recited herein can be recited as means for performing a function or step for performing a function. Therefore, it should be understood that such language is entitled to cover all such structures, materials, or acts disclosed within this specification and their equivalents, including the matter incorporated by reference.

[00153] The utilization of the biomarkers disclosed herein, and the various methods for determining biomarker values are described in detail above with respect to evaluation of sample quality and suitability for further analysis, such as diagnostic analysis. In some embodiments, the biomarkers, methods, and kits described herein are used to assess the absolute sample quality of one or more samples or the relative consistency of sample quality across a plurality of samples. In some such embodiments, the methods comprise identifying samples that pass or fail a quality assessment. In some embodiments, samples that pass quality assessment are analyzed and samples that fail quality assessment are discarded. In some embodiments, information obtained using the biomarkers, methods, and kits herein can be used to determine whether a sample collection and processing method or facility is suitable.

EXAMPLES

[00154] The following examples are provided for illustrative purposes only and are not intended to limit the scope of the application as defined by the appended claims. Routine molecular biology techniques described in the following examples can be carried out as described in standard laboratory manuals, such as Sambrook et al., *Molecular Cloning: A Laboratory Manual*, 3rd. ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., (2001).

Example 1: Exemplary Biomarker Detection Using Aptamers

[00155] An exemplary method of detecting one or more biomarker proteins in a sample is described, e.g., in Kraemer et al., *PloS One* 6(10): e26332, and is described below. Three different methods of quantification: microarray-based hybridization, a Luminex bead-based method, and qPCR, are described.

Reagents

[00156] HEPES, NaCl, KCl, EDTA, EGTA, MgCl₂ and Tween-20 may be purchased, e.g., from Fisher Biosciences. Dextran sulfate sodium salt (DxSO₄), nominally 8000 molecular weight, may be purchased, e.g., from AIC and is dialyzed against deionized water for at least 20 hours with one exchange. KOD EX DNA polymerase may be purchased, e.g., from VWR. Tetramethylammonium chloride and CAPSO may be purchased, e.g., from Sigma-Aldrich and streptavidin-phycoerythrin (SAPE) may be purchased, e.g., from Moss Inc. 4-(2-Aminoethyl)-benzenesulfonylfluoride hydrochloride (AEBSF) may be purchased, e.g., from Gold Biotechnology. Streptavidin-coated 96-well plates may be purchased, e.g., from Thermo Scientific (Pierce Streptavidin Coated Plates HBC, clear, 96-well, product number 15500 or 15501). NHS-PEO₄-biotin may be purchased, e.g., from Thermo Scientific (EZ-Link NHS-PEO₄-Biotin, product number 21329), dissolved in anhydrous DMSO, and may be stored frozen in single-use aliquots. Proteins may be purchased, e.g., from R&D Systems. Resistin and MCP-1 may be purchased, e.g., from PeproTech, and tPA may be purchased, e.g., from VWR.

Nucleic acids

[00157] Conventional (including amine- and biotin-substituted) oligodeoxynucleotides may be purchased, e.g., from Integrated DNA Technologies (IDT). Z-Block is a single-stranded oligodeoxynucleotide of sequence 5'-(AC-BnBn)₇-AC-3', where Bn indicates a benzyl-substituted deoxyuridine residue. Z-block may be synthesized using conventional phosphoramidite chemistry. Aptamer capture reagents may also be synthesized by

conventional phosphoramidite chemistry, and may be purified, for example, on a 21.5×75 mm PRP-3 column, operating at 80°C on a Waters Autopurification 2767 system (or Waters 600 series semi-automated system), using, for example, a timberline TL-600 or TL-150 heater and a gradient of triethylammonium bicarbonate (TEAB) / I to elute product. Detection is performed at 260 nm and fractions are collected across the main peak prior to pooling best fractions.

Buffers

[00158] Buffer SB18 is composed of 40 mM HEPES, 101 mM NaCl, 5 mM KCl, 5 mM MgCl₂, and 0.05% (v/v) Tween 20 adjusted to pH 7.5 with NaOH. Buffer SB17 is SB18 supplemented with 1 mM trisodium EDTA. Buffer PB1 is composed of 10 mM HEPES, 101 mM NaCl, 5 mM KCl, 5 mM MgCl₂, 1 mM trisodium EDTA and 0.05% (v/v) Tween-20 adjusted to pH 7.5 with NaOH. CAPSO elution buffer consists of 100 mM CAPSO pH 10.0 and 1 M NaCl. Neutralization buffer contains of 500 mM HEPES, 500 mM HCl, and 0.05% (v/v) Tween-20. Agilent Hybridization Buffer is a proprietary formulation that is supplied as part of a kit (Oligo aCGH/ChIP-on-chip Hybridization Kit). Agilent Wash Buffer 1 is a proprietary formulation (Oligo aCGH/ChIP-on-chip Wash Buffer 1, Agilent). Agilent Wash Buffer 2 is a proprietary formulation (Oligo aCGH/ChIP-on-chip Wash Buffer 2, Agilent). TMAC hybridization solution consists of 4.5 M tetramethylammonium chloride, 6 mM trisodium EDTA, 75 mM Tris-HCl (pH 8.0), and 0.15% (v/v) Sarkosyl. KOD buffer (10-fold concentrated) consists of 1200 mM Tris-HCl, 15 mM MgSO₄, 100 mM KCl, 60 mM (NH₄)₂SO₄, 1% v/v Triton-X 100 and 1 mg/mL BSA.

Sample preparation

[00159] Serum (stored at –80°C in 100 µL aliquots) is thawed in a 25°C water bath for 10 minutes, then stored on ice prior to sample dilution. Samples are mixed by gentle vortexing for 8 seconds. A 6% serum sample solution is prepared by dilution into 0.94× SB17 supplemented with 0.6 mM MgCl₂, 1 mM trisodium EGTA, 0.8 mM AEBSF, and 2 µM Z-Block. A portion of the 6% serum stock solution is diluted 10-fold in SB17 to create a 0.6% serum stock. 6% and 0.6% stocks are used, in some embodiments, to detect high- and low-abundance analytes, respectively.

Capture reagent (aptamer) and streptavidin plate preparation

[00160] Aptamers are grouped into 2 mixes according to the relative abundance of their cognate analytes (or biomarkers). Stock concentrations are 4 nM for each aptamer, and the final concentration of each aptamer is 0.5 nM. Aptamer stock mixes are diluted 4-fold in SB17 buffer, heated to 95 °C for 5 min and cooled to 37 °C over a 15-minute period prior to

use. This denaturation-renaturation cycle is intended to normalize aptamer conformer distributions and thus ensure reproducible aptamer activity in spite of variable histories. Streptavidin plates are washed twice with 150 μ L buffer PB1 prior to use.

Incubation and plate capture

[00161] Heat-cooled 2 $\times$ Aptamer mixes (55 μ L) are combined with an equal volume of 6% or 0.6% serum dilutions, producing mixes containing 3% and 0.3% serum. The plates are sealed with a Silicone Sealing Mat (AxyMat Silicone sealing mat, VWR) and incubated for 1.5 h at 37 °C. Mixes are then transferred to the wells of a washed 96-well streptavidin plate and further incubated on an Eppendorf Thermomixer set at 37 °C, with shaking at 800 rpm, for two hours.

Manual Assay

[00162] Unless otherwise specified, liquid is removed by dumping, followed by two taps onto layered paper towels. Wash volumes are 150 μ L and all shaking incubations are done on an Eppendorf Thermomixer set at 25°C, 800 rpm. Mixes are removed by pipetting, and plates are washed twice for 1 minute with buffer PB1 supplemented with 1 mM dextran sulfate and 500 μ M biotin, then 4 times for 15 seconds with buffer PB1. A freshly made solution of 1 mM NHS-PEO4-biotin in buffer PB1 (150 μ L/well) is added, and plates are incubated for 5 minutes with shaking. The NHS-biotin solution is removed, and plates washed 3 times with buffer PB1 supplemented with 20 mM glycine, and 3 times with buffer PB1. Eighty-five μ L of buffer PB1 supplemented with 1 mM DxSO_4 is then added to each well, and plates are irradiated under a BlackRay UV lamp (nominal wavelength 365 nm) at a distance of 5 cm for 20 minutes with shaking. Samples are transferred to a fresh, washed streptavidin-coated plate, or an unused well of the existing washed streptavidin plate, combining high and low sample dilution mixtures into a single well. Samples are incubated at room temperature with shaking for 10 minutes. Unabsorbed material is removed and the plates washed 8 times for 15 seconds each with buffer PB1 supplemented with 30% glycerol. Plates are then washed once with buffer PB1. Aptamers are eluted for 5 minutes at room temperature with 100 μ L CAPSO elution buffer. 90 μ L of the eluate is transferred to a 96-well HybAid plate and 10 μ L neutralization buffer is added.

Semi-Automated Assay

[00163] Streptavidin plates bearing adsorbed equilibration mixes are placed on the deck of a BioTek EL406 plate washer, which is programmed to perform the following steps: unadsorbed material is removed by aspiration, and wells are washed 4 times with 300 μ L of buffer PB1 supplemented with 1 mM dextran sulfate and 500 μ M biotin. Wells are then

washed 3 times with 300 μ L buffer PB1. One hundred fifty μ L of a freshly prepared (from a 100 mM stock in DMSO) solution of 1 mM NHS-PEO4-biotin in buffer PB1 is added. Plates are incubated for 5 minutes with shaking. Liquid is aspirated, and wells are washed 8 times with 300 μ L buffer PB1 supplemented with 10 mM glycine. One hundred μ L of buffer PB1 supplemented with 1 mM dextran sulfate are added. After these automated steps, plates are removed from the plate washer and placed on a thermoshaker mounted under a UV light source (BlackRay, nominal wavelength 365 nm) at a distance of 5 cm for 20 minutes. The thermoshaker is set at 800 rpm and 25 °C. After 20 minutes irradiation, samples are manually transferred to a fresh, washed streptavidin plate (or to an unused well of the existing washed plate). High-abundance (3% serum+3% aptamer mix) and low-abundance reaction mixes (0.3% serum+0.3% aptamer mix) are combined into a single well at this point. This “Catch-2” plate is placed on the deck of BioTek EL406 plate washer, which is programmed to perform the following steps: the plate is incubated for 10 minutes with shaking. Liquid is aspirated, and wells are washed 21 times with 300 μ L buffer PB1 supplemented with 30% glycerol. Wells are washed 5 times with 300 μ L buffer PB1, and the final wash is aspirated. One hundred μ L CAPSO elution buffer are added, and aptamers are eluted for 5 minutes with shaking. Following these automated steps, the plate is then removed from the deck of the plate washer, and 90 μ L aliquots of the samples are transferred manually to the wells of a HybAid 96-well plate that contains 10 μ L neutralization buffer.

Hybridization to custom Agilent 8 $\times$ 15k microarrays

[00164] 24 μ L of the neutralized eluate is transferred to a new 96-well plate and 6 μ L of 10 $\times$ Agilent Block (Oligo aCGH/ChIP-on-chip Hybridization Kit, Large Volume, Agilent 5188–5380), containing a set of hybridization controls composed of 10 Cy3 aptamers is added to each well. Thirty μ L 2 $\times$ Agilent Hybridization buffer is added to each sample and mixed. Forty μ L of the resulting hybridization solution is manually pipetted into each “well” of the hybridization gasket slide (Hybridization Gasket Slide, 8-microarray per slide format, Agilent). Custom Agilent microarray slides, bearing 10 probes per array complementary to 40 nucleotide random region of each aptamer with a 20 $\times$ dT linker, are placed onto the gasket slides according to the manufacturers’ protocol. The assembly (Hybridization Chamber Kit – SureHyb-enabled, Agilent) is clamped and incubated for 19 hours at 60 °C while rotating at 20 rpm.

Post Hybridization Washing

[00165] Approximately 400 mL Agilent Wash Buffer 1 is placed into each of two separate glass staining dishes. Slides (no more than two at a time) are disassembled and

separated while submerged in Wash Buffer 1, then transferred to a slide rack in a second staining dish also containing Wash Buffer 1. Slides are incubated for an additional 5 minutes in Wash Buffer 1 with stirring. Slides are transferred to Wash Buffer 2 pre-equilibrated to 37 °C and incubated for 5 minutes with stirring. Slides are transferred to a fourth staining dish containing acetonitrile, and incubated for 5 minutes with stirring.

Microarray Imaging

[00166] Microarray slides are imaged with an Agilent G2565CA Microarray Scanner System, using the Cy3-channel at 5 µm resolution at 100% PMT setting, and the XRD option enabled at 0.05. The resulting TIFF images are processed using Agilent feature extraction software version 10.5.1.1 with the GE1_105_Dec08 protocol.

Luminex probe design

[00167] Probes immobilized to beads have 40 deoxynucleotides complementary to the 3' end of the 40 nucleotide random region of the target aptamer. The aptamer complementary region is coupled to Luminex Microspheres through a hexaethyleneglycol (HEG) linker bearing a 5' amino terminus. Biotinylated detection deoxyoligonucleotides comprise 17–21 deoxynucleotides complementary to the 5' primer region of target aptamers. Biotin moieties are appended to the 3' ends of detection oligos.

Coupling of probes to Luminex Microspheres

[00168] Probes are coupled to Luminex Microplex Microspheres essentially per the manufacturer's instructions, but with the following modifications: amino-terminal oligonucleotide amounts are 0.08 nMol per 2.5×10^6 microspheres, and the second EDC addition is 5 µL at 10 mg/mL. Coupling reactions are performed in an Eppendorf ThermoShaker set at 25 °C and 600 rpm.

Microsphere hybridization

[00169] Microsphere stock solutions (about 40000 microspheres/µL) are vortexed and sonicated in a Health Sonics ultrasonic cleaner (Model: T1.9C) for 60 seconds to suspend the microspheres. Suspended microspheres are diluted to 2000 microspheres per reaction in $1.5 \times$ TMAC hybridization solutions and mixed by vortexing and sonication. Thirty-three µL per reaction of the bead mixture are transferred into a 96-well HybAid plate. Seven µL of 15 nM biotinylated detection oligonucleotide stock in $1 \times$ TE buffer are added to each reaction and mixed. Ten µL of neutralized assay sample are added and the plate is sealed with a silicon cap mat seal. The plate is first incubated at 96 °C for 5 minutes and incubated at 50°C without agitation overnight in a conventional hybridization oven. A filter plate (Dura pore, Millipore part number MSBVN1250, 1.2 µm pore size) is prewetted with 75 µL $1 \times$ TMAC

hybridization solution supplemented with 0.5% (w/v) BSA. The entire sample volume from the hybridization reaction is transferred to the filter plate. The hybridization plate is rinsed with 75 μ L 1 $\times$ TMAC hybridization solution containing 0.5% BSA and any remaining material is transferred to the filter plate. Samples are filtered under slow vacuum, with 150 μ L buffer evacuated over about 8 seconds. The filter plate is washed once with 75 μ L 1 $\times$ TMAC hybridization solution containing 0.5% BSA and the microspheres in the filter plate are resuspended in 75 μ L 1 $\times$ TMAC hybridization solution containing 0.5% BSA. The filter plate is protected from light and incubated on an Eppendorf Thermalmixer R for 5 minutes at 1000 rpm. The filter plate is then washed once with 75 μ L 1 $\times$ TMAC hybridization solution containing 0.5% BSA. 75 μ L of 10 μ g/mL streptavidin phycoerythrin (SAPE-100, MOSS, Inc.) in 1 $\times$ TMAC hybridization solution is added to each reaction and incubated on Eppendorf Thermalmixer R at 25 $^{\circ}$ C at 1000 rpm for 60 minutes. The filter plate is washed twice with 75 μ L 1 $\times$ TMAC hybridization solution containing 0.5% BSA and the microspheres in the filter plate are resuspended in 75 μ L 1 $\times$ TMAC hybridization solution containing 0.5% BSA. The filter plate is then incubated protected from light on an Eppendorf Thermalmixer R for 5 minutes, 1000 rpm. The filter plate is then washed once with 75 μ L 1 $\times$ TMAC hybridization solution containing 0.5% BSA. Microspheres are resuspended in 75 μ L 1 $\times$ TMAC hybridization solution supplemented with 0.5% BSA, and analyzed on a Luminex 100 instrument running Xponent 3.0 software. At least 100 microspheres are counted per bead type, under high PMT calibration and a doublet discriminator setting of 7500 to 18000.

qPCR read-out

[00170] Standard curves for qPCR are prepared in water ranging from 10⁸ to 10² copies with 10-fold dilutions and a no-template control. Neutralized assay samples are diluted 40-fold into diH₂O. The qPCR master mix is prepared at 2 $\times$ final concentration (2 $\times$ KOD buffer, 400 μ M dNTP mix, 400 nM forward and reverse primer mix, 2 $\times$ SYBR Green I and 0.5 U KOD EX). Ten μ L of 2 $\times$ qPCR master mix is added to 10 μ L of diluted assay sample. qPCR is run on a BioRad MyIQ iCycler with 2 minutes at 96 $^{\circ}$ C followed by 40 cycles of 96 $^{\circ}$ C for 5 seconds and 72 $^{\circ}$ C for 30 seconds.

Example 2. Freeze-Thaw Cycles Serum Model

[00171] Serum samples obtained from a subject may be initially obtained as a whole blood sample, clotted and then centrifuged. The resulting serum can then be collected by aspirating the serum from the precipitate. A serum sample must be analyzed or frozen at or less than -20 $^{\circ}$ C or -80 $^{\circ}$ C after collection. The serum sample may be frozen and thawed

multiple times for testing. Multiple freeze-thaw cycles can be a source of sample variability, potentially resulting in denatured protein or lysed cells.

[00172] In order to assess the quality of a serum sample, a linear regression elastic net model containing a panel of the 8 biomarker proteins listed in Table 1 was developed. The model provides a predicted value that is the number of freeze-thaw events a serum sample has undergone. The training and verification datasets were obtained from analysis of samples from adult human volunteers, as described in Model Development below. Table 2 shows the performance metrics for the model. “CCC” is Concordance Correlation Coefficient.”R²” is the degree of linear correlation, or the goodness-of-fit. CCC and R² are indicative of predictive performance of the model.

Table 2

Model	Dataset	# of Analytes	Predictive Performance	
			CCC (95% Conf)	R ² (95% Conf)
Freeze-thaw cycles (serum)	Training	8	0.980 (0.97, 0.99)	0.965 (0.95, 0.98)
	Verification		0.935 (0.89, 0.98)	0.945 (0.86, 0.98)

[00173] The freeze-thaw cycle serum model is an elastic net linear regression model. This model has 8 aptamers as its features, which bind the 8 different biomarker proteins listed in Table 1. The output of this model is the prediction of the number of times the sample was frozen and then thawed. Thus, the output is a number, 1 or higher, with 1 being immediate transfer of clotted and centrifuged serum samples after sample collection to freezing conditions.

Model Development

[00174] Multiple red-topped serum tubes of blood were collected from each of ten adult volunteers by applying a tourniquet and using a 21G butterfly needle set. The serum tubes were inverted 5x after collection and the blood was allowed to clot for 60 minutes after which the tubes were centrifuged at 2200 x g for 15 minutes in Beckman Coulter Allegra X-15R centrifuge. 0.75mL aliquots of serum were made in Eppendorf tubes for each donor, and “day 0” 70μL aliquots immediately transferred from the Eppendorf tubes to matrix tubes for storage at -80°C. The remaining Eppendorf aliquot was also placed at -80°C, and successively thawed each day for 9 days. On days 1, 2, 3, 4, 6, and 9 of thaw, 70μL aliquots were again transferred to matrix tubes for storage at -80°C until the day of the that analysis with the aptamers was performed, for example, according to a protocol described in Example

1. On the day of analysis, all matrix aliquots were thawed a final time such that the total number of freeze-thaws tested in the assay was 1x, 2x, 3x, 4x, 5x, 7x and 10x.

[00175] For this test, the data were split into training and verification sets with an 80/20 split independently for each number of freeze/thaw cycles. Individual donors were not assigned fully to training/verification, but rather each number of cycles for each donor was assigned randomly. This split results in 56 training samples and 14 verification samples.

[00176] Due to low sample numbers, a hold-out test dataset was not used for model development. The model was developed on the training dataset and then assessed on the verification set, but it was not used as a holdout validation or test set. The sample numbers are shown in Table 3 below.

Table 3: Sample numbers for freeze-thaw serum analysis

Dataset	Total Samples	Number of Freeze /Thaw Cycles						
		1	2	3	4	5	7	10
Training	56	8	8	8	8	8	8	8
Verification	14	2	2	2	2	2	2	2

[00177] After the aptamer assay was performed, the data was normalized. Control samples were internally median normalized, plate scaled, calibrated, and samples normalized using adaptive normalization by maximum likelihood (ANML). All samples were within normalization scale factor pass/fail criteria and subsequently, no samples were removed from analysis. Normalization scale factors were not significantly correlated with freeze-thaw cycles.

POC Results

[00178] The Proof of Concept (POC) results showed that a number of analytes were significant at different false-discovery rate (FDR) levels. Table 4 below shows those numbers and percentages calculated with the Pearson correlation test for linear models.

Table 4: Number and percentage of analytes significant at different Type I error cut-offs for freeze-thaw univariate results

FDR	# analytes (%) $\leq$ FDR level
0.10	354 (4.7 %)
0.05	228 (3.0 %)
0.01	117 (1.5 %)

Refinement and Validation

[00179] For the models developed in refinement, the training and verification sets described above were used.

[00180] Feature selection was accomplished by starting with the top 200 features by rank as identified in the POC univariate analysis. This list was refined through a series of elastic net regressions with alpha set to 0.15 and lambda set to 0.15, which were the optimal values identified in the POC analysis. Four rounds of elastic net regression with these parameters produced a list around 20 aptamers with non-zero coefficients. This list was further refined by removing all analytes that were already present in the serum time-to-freeze, time-to-clot, and time-to-decant models and retaining only the aptamers that are present in a panel of 3,000 markers. These steps produced a final set of 8 features.

[00181] The final model is an elastic net linear regression model with these 8 features, trained on the training dataset and evaluated on the verification set. Any predictions below zero are mapped to zero, as negative numbers of freeze/thaw cycles are not meaningful. Lin's correlation coefficient was calculated for the training and verification data. The model performs well on the verification set with an R-squared value of 0.945, a root mean square error (RMSE) of 0.76 and Concordance Correlation Coefficient (CCC) of 0.935 (Table 5).

Table 5: Performance of freeze-thaw serum model on training and verification data sets

Data set	R ²	RMSE	CCC
Training	0.965 (0.95, 0.98)	0.55 (0.47, 0.63)	0.980 (0.97, 0.99)
Verification	0.945 (0.86, 0.98)	0.76 (0.52, 0.99)	0.935 (0.89, 0.98)

[00182] The effect of winsorization and dropping features were evaluated to correct for outlier values in other datasets. For the original data, the RMSE was 0.55, with winsorization it was 1.93 and for feature removal it was 1.06. Outlier values were therefore replaced with zero for this model. These results are shown in Tables 6 and 7 below.

Table 6: RMSE Values

Model		RMSE		Final Selection
	Original (training)	Winsorized	Replacement with Zero	
Freeze-thaw cycles (serum)	0.55	1.93	1.06	Replacement with zero

Table 7: Predictive variation due to assay noise

Model	Bound	Error Bounds (95%)	Error Bounds (99%)
Freeze-Thaw (serum)	Upper Bound	0.909	1.211
	Lower Bound	-0.744	-2.499

Example 3: Use of the Sample Handling Model

[00183] The sample handling model may be used to censor individual samples for specific output that is important to the test. In one embodiment, samples that are identified as failing the quality assessment for the number of freeze-thaw cycles may be removed if the specific output of the number of freeze-thaw cycles is important to the test being conducted. In other embodiments, samples that are identified as failing the quality assessment for the number of freeze-thaw cycles may be included if the specific output of the number of freeze-thaw cycles is not important to the test being conducted. In some embodiments, a panel of biomarker proteins may be modified in a subsequent or simultaneous analysis such as protein biomarker discovery analysis, protein expression level analysis, diagnostic method or prognostic method based on the determined approximate times for each of the plurality of samples. In some embodiments, the panel of biomarker proteins is reduced in number of biomarker proteins measured.

[00184] The sample handling model may be used to identify biases within a plurality of samples collected. In one embodiment, the sample handling model may be used to identify biases between experimental samples and controls.

[00185] The sample handling model may be used to assess adherence to a clinical study protocol for sample collection and processing.

[00186] The sample handling model may be used to identify outlier samples in a plurality of samples. In one embodiment, the outlier sample may be 1 or 2 or 3 or more standard deviations from the other samples in the model. The outlier sample may have better or worse quality relative to the other samples in the model.

[00187] The sampling model may be used to compare a plurality of samples from a first site to one or more additional sites for sample collection and processing.

Example 4: Analysis of Freeze-Thaw Serum Model Biomarker Panels

[00188] Model biomarker panels comprising various combinations of the biomarkers listed in Table 1 were analyzed to determine the coefficient of determination (R^2) value for the various combinations. The model biomarker panels may be based on a panel of N biomarker proteins having an R^2 value of at least 0.65, at least 0.7, at least 0.75, at least 0.8, at least 0.85, at least 0.9, or at least 0.95, where N is 1, 2, 3, 4, 5, 6, 7, and/or 8 of the biomarker proteins listed in Table 1. Table 8 below shows the model results when various combinations comprising 1 to 8 biomarker proteins were measured. See also Fig. 6 showing the Rsquared values for models with increasing number of features.

[00189] **Table 8: Freeze-Thaw Serum Biomarker Panels**

Rsquared	ILK	RANGAP1	EIF1B	LMAN2	SCFD1	LTA4H	RPL5	PLD3
0.376								PLD3
0.183							RPL5	
0.611							RPL5	PLD3
0.557						LTA4H		
0.628						LTA4H		PLD3
0.642						LTA4H	RPL5	
0.758						LTA4H	RPL5	PLD3
0.479					SCFD1			
0.588					SCFD1			PLD3
0.528					SCFD1		RPL5	
0.696					SCFD1		RPL5	PLD3
0.781					SCFD1	LTA4H		
0.792					SCFD1	LTA4H		PLD3
0.81					SCFD1	LTA4H	RPL5	
0.841					SCFD1	LTA4H	RPL5	PLD3
0.272				LMAN2				
0.484				LMAN2				PLD3
0.441				LMAN2			RPL5	
0.699				LMAN2			RPL5	PLD3
0.611				LMAN2		LTA4H		
0.664				LMAN2		LTA4H		PLD3
0.704				LMAN2		LTA4H	RPL5	
0.797				LMAN2		LTA4H	RPL5	PLD3
0.638				LMAN2	SCFD1			
0.687				LMAN2	SCFD1			PLD3
0.693				LMAN2	SCFD1		RPL5	
0.783				LMAN2	SCFD1		RPL5	PLD3
0.824				LMAN2	SCFD1	LTA4H		
0.829				LMAN2	SCFD1	LTA4H		PLD3

Rsquared	ILK	RANGAP1	EIF1B	LMAN2	SCFD1	LTA4H	RPL5	PLD3
0.858				LMAN2	SCFD1	LTA4H	RPL5	
0.879				LMAN2	SCFD1	LTA4H	RPL5	PLD3
0.538			EIF1B					
0.817			EIF1B					PLD3
0.546			EIF1B				RPL5	
0.818			EIF1B				RPL5	PLD3
0.641			EIF1B			LTA4H		
0.817			EIF1B			LTA4H		PLD3
0.652			EIF1B			LTA4H	RPL5	
0.819			EIF1B			LTA4H	RPL5	PLD3
0.774			EIF1B		SCFD1			
0.895			EIF1B		SCFD1			PLD3
0.801			EIF1B		SCFD1		RPL5	
0.897			EIF1B		SCFD1		RPL5	PLD3
0.836			EIF1B		SCFD1	LTA4H		
0.897			EIF1B		SCFD1	LTA4H		PLD3
0.836			EIF1B		SCFD1	LTA4H	RPL5	
0.897			EIF1B		SCFD1	LTA4H	RPL5	PLD3
0.629			EIF1B	LMAN2				
0.835			EIF1B	LMAN2				PLD3
0.629			EIF1B	LMAN2			RPL5	
0.839			EIF1B	LMAN2			RPL5	PLD3
0.69			EIF1B	LMAN2		LTA4H		
0.836			EIF1B	LMAN2		LTA4H		PLD3
0.708			EIF1B	LMAN2		LTA4H	RPL5	
0.84			EIF1B	LMAN2		LTA4H	RPL5	PLD3
0.84			EIF1B	LMAN2	SCFD1			
0.918			EIF1B	LMAN2	SCFD1			PLD3
0.849			EIF1B	LMAN2	SCFD1		RPL5	
0.918			EIF1B	LMAN2	SCFD1		RPL5	PLD3
0.875			EIF1B	LMAN2	SCFD1	LTA4H		
0.919			EIF1B	LMAN2	SCFD1	LTA4H		PLD3
0.875			EIF1B	LMAN2	SCFD1	LTA4H	RPL5	
0.919			EIF1B	LMAN2	SCFD1	LTA4H	RPL5	PLD3
0.518		RANGAP1						
0.782		RANGAP1						PLD3
0.519		RANGAP1					RPL5	
0.791		RANGAP1					RPL5	PLD3
0.692		RANGAP1				LTA4H		
0.812		RANGAP1				LTA4H		PLD3
0.698		RANGAP1				LTA4H	RPL5	
0.826		RANGAP1				LTA4H	RPL5	PLD3
0.763		RANGAP1			SCFD1			

Rsquared	ILK	RANGAP1	EIF1B	LMAN2	SCFD1	LTA4H	RPL5	PLD3
0.87		RANGAP1			SCFD1			PLD3
0.775		RANGAP1			SCFD1		RPL5	
0.87		RANGAP1			SCFD1		RPL5	PLD3
0.869		RANGAP1			SCFD1	LTA4H		
0.906		RANGAP1			SCFD1	LTA4H		PLD3
0.869		RANGAP1			SCFD1	LTA4H	RPL5	
0.906		RANGAP1			SCFD1	LTA4H	RPL5	PLD3
0.66		RANGAP1		LMAN2				
0.83		RANGAP1		LMAN2				PLD3
0.661		RANGAP1		LMAN2			RPL5	
0.842		RANGAP1		LMAN2			RPL5	PLD3
0.751		RANGAP1		LMAN2		LTA4H		
0.846		RANGAP1		LMAN2		LTA4H		PLD3
0.759		RANGAP1		LMAN2		LTA4H	RPL5	
0.862		RANGAP1		LMAN2		LTA4H	RPL5	PLD3
0.862		RANGAP1		LMAN2	SCFD1			
0.922		RANGAP1		LMAN2	SCFD1			PLD3
0.866		RANGAP1		LMAN2	SCFD1		RPL5	
0.922		RANGAP1		LMAN2	SCFD1		RPL5	PLD3
0.916		RANGAP1		LMAN2	SCFD1	LTA4H		
0.941		RANGAP1		LMAN2	SCFD1	LTA4H		PLD3
0.916		RANGAP1		LMAN2	SCFD1	LTA4H	RPL5	
0.942		RANGAP1		LMAN2	SCFD1	LTA4H	RPL5	PLD3
0.567		RANGAP1	EIF1B					
0.834		RANGAP1	EIF1B					PLD3
0.58		RANGAP1	EIF1B				RPL5	
0.834		RANGAP1	EIF1B				RPL5	PLD3
0.693		RANGAP1	EIF1B			LTA4H		
0.835		RANGAP1	EIF1B			LTA4H		PLD3
0.702		RANGAP1	EIF1B			LTA4H	RPL5	
0.836		RANGAP1	EIF1B			LTA4H	RPL5	PLD3
0.793		RANGAP1	EIF1B		SCFD1			
0.908		RANGAP1	EIF1B		SCFD1			PLD3
0.824		RANGAP1	EIF1B		SCFD1		RPL5	
0.912		RANGAP1	EIF1B		SCFD1		RPL5	PLD3
0.869		RANGAP1	EIF1B		SCFD1	LTA4H		
0.914		RANGAP1	EIF1B		SCFD1	LTA4H		PLD3
0.87		RANGAP1	EIF1B		SCFD1	LTA4H	RPL5	
0.915		RANGAP1	EIF1B		SCFD1	LTA4H	RPL5	PLD3
0.675		RANGAP1	EIF1B	LMAN2				
0.859		RANGAP1	EIF1B	LMAN2				PLD3
0.676		RANGAP1	EIF1B	LMAN2			RPL5	
0.861		RANGAP1	EIF1B	LMAN2			RPL5	PLD3

Rsquared	ILK	RANGAP1	EIF1B	LMAN2	SCFD1	LTA4H	RPL5	PLD3
0.752		RANGAP1	EIF1B	LMAN2		LTA4H		
0.86		RANGAP1	EIF1B	LMAN2		LTA4H		PLD3
0.77		RANGAP1	EIF1B	LMAN2		LTA4H	RPL5	
0.864		RANGAP1	EIF1B	LMAN2		LTA4H	RPL5	PLD3
0.871		RANGAP1	EIF1B	LMAN2	SCFD1			
0.939		RANGAP1	EIF1B	LMAN2	SCFD1			PLD3
0.882		RANGAP1	EIF1B	LMAN2	SCFD1		RPL5	
0.94		RANGAP1	EIF1B	LMAN2	SCFD1		RPL5	PLD3
0.916		RANGAP1	EIF1B	LMAN2	SCFD1	LTA4H		
0.944		RANGAP1	EIF1B	LMAN2	SCFD1	LTA4H		PLD3
0.917		RANGAP1	EIF1B	LMAN2	SCFD1	LTA4H	RPL5	
0.944		RANGAP1	EIF1B	LMAN2	SCFD1	LTA4H	RPL5	PLD3
0.189	ILK							
0.378	ILK							PLD3
0.548	ILK						RPL5	
0.637	ILK						RPL5	PLD3
0.647	ILK					LTA4H		
0.65	ILK					LTA4H		PLD3
0.835	ILK					LTA4H	RPL5	
0.835	ILK					LTA4H	RPL5	PLD3
0.48	ILK				SCFD1			
0.65	ILK				SCFD1			PLD3
0.587	ILK				SCFD1		RPL5	
0.698	ILK				SCFD1		RPL5	PLD3
0.783	ILK				SCFD1	LTA4H		
0.795	ILK				SCFD1	LTA4H		PLD3
0.851	ILK				SCFD1	LTA4H	RPL5	
0.853	ILK				SCFD1	LTA4H	RPL5	PLD3
0.372	ILK			LMAN2				
0.484	ILK			LMAN2				PLD3
0.677	ILK			LMAN2			RPL5	
0.725	ILK			LMAN2			RPL5	PLD3
0.68	ILK			LMAN2		LTA4H		
0.682	ILK			LMAN2		LTA4H		PLD3
0.865	ILK			LMAN2		LTA4H	RPL5	
0.865	ILK			LMAN2		LTA4H	RPL5	PLD3
0.639	ILK			LMAN2	SCFD1			
0.742	ILK			LMAN2	SCFD1			PLD3
0.72	ILK			LMAN2	SCFD1		RPL5	
0.785	ILK			LMAN2	SCFD1		RPL5	PLD3
0.824	ILK			LMAN2	SCFD1	LTA4H		
0.834	ILK			LMAN2	SCFD1	LTA4H		PLD3
0.885	ILK			LMAN2	SCFD1	LTA4H	RPL5	

Rsquared	ILK	RANGAP1	EIF1B	LMAN2	SCFD1	LTA4H	RPL5	PLD3
0.887	ILK			LMAN2	SCFD1	LTA4H	RPL5	PLD3
0.907	ILK		EIF1B					
0.912	ILK		EIF1B					PLD3
0.911	ILK		EIF1B				RPL5	
0.916	ILK		EIF1B				RPL5	PLD3
0.908	ILK		EIF1B			LTA4H		
0.912	ILK		EIF1B			LTA4H		PLD3
0.915	ILK		EIF1B			LTA4H	RPL5	
0.917	ILK		EIF1B			LTA4H	RPL5	PLD3
0.914	ILK		EIF1B		SCFD1			
0.923	ILK		EIF1B		SCFD1			PLD3
0.915	ILK		EIF1B		SCFD1		RPL5	
0.923	ILK		EIF1B		SCFD1		RPL5	PLD3
0.917	ILK		EIF1B		SCFD1	LTA4H		
0.923	ILK		EIF1B		SCFD1	LTA4H		PLD3
0.919	ILK		EIF1B		SCFD1	LTA4H	RPL5	
0.924	ILK		EIF1B		SCFD1	LTA4H	RPL5	PLD3
0.918	ILK		EIF1B	LMAN2				
0.921	ILK		EIF1B	LMAN2				PLD3
0.924	ILK		EIF1B	LMAN2			RPL5	
0.928	ILK		EIF1B	LMAN2			RPL5	PLD3
0.918	ILK		EIF1B	LMAN2		LTA4H		
0.921	ILK		EIF1B	LMAN2		LTA4H		PLD3
0.927	ILK		EIF1B	LMAN2		LTA4H	RPL5	
0.929	ILK		EIF1B	LMAN2		LTA4H	RPL5	PLD3
0.929	ILK		EIF1B	LMAN2	SCFD1			
0.937	ILK		EIF1B	LMAN2	SCFD1			PLD3
0.93	ILK		EIF1B	LMAN2	SCFD1		RPL5	
0.938	ILK		EIF1B	LMAN2	SCFD1		RPL5	PLD3
0.931	ILK		EIF1B	LMAN2	SCFD1	LTA4H		
0.937	ILK		EIF1B	LMAN2	SCFD1	LTA4H		PLD3
0.935	ILK		EIF1B	LMAN2	SCFD1	LTA4H	RPL5	
0.939	ILK		EIF1B	LMAN2	SCFD1	LTA4H	RPL5	PLD3
0.831	ILK	RANGAP1						
0.844	ILK	RANGAP1						PLD3
0.85	ILK	RANGAP1					RPL5	
0.863	ILK	RANGAP1					RPL5	PLD3
0.89	ILK	RANGAP1				LTA4H		
0.89	ILK	RANGAP1				LTA4H		PLD3
0.922	ILK	RANGAP1				LTA4H	RPL5	
0.922	ILK	RANGAP1				LTA4H	RPL5	PLD3
0.856	ILK	RANGAP1			SCFD1			
0.879	ILK	RANGAP1			SCFD1			PLD3

Rsquared	ILK	RANGAP1	EIF1B	LMAN2	SCFD1	LTA4H	RPL5	PLD3
0.861	ILK	RANGAP1			SCFD1		RPL5	
0.881	ILK	RANGAP1			SCFD1		RPL5	PLD3
0.919	ILK	RANGAP1			SCFD1	LTA4H		
0.92	ILK	RANGAP1			SCFD1	LTA4H		PLD3
0.93	ILK	RANGAP1			SCFD1	LTA4H	RPL5	
0.931	ILK	RANGAP1			SCFD1	LTA4H	RPL5	PLD3
0.882	ILK	RANGAP1		LMAN2				
0.888	ILK	RANGAP1		LMAN2				PLD3
0.904	ILK	RANGAP1		LMAN2			RPL5	
0.91	ILK	RANGAP1		LMAN2			RPL5	PLD3
0.916	ILK	RANGAP1		LMAN2		LTA4H		
0.916	ILK	RANGAP1		LMAN2		LTA4H		PLD3
0.948	ILK	RANGAP1		LMAN2		LTA4H	RPL5	
0.949	ILK	RANGAP1		LMAN2		LTA4H	RPL5	PLD3
0.914	ILK	RANGAP1		LMAN2	SCFD1			
0.927	ILK	RANGAP1		LMAN2	SCFD1			PLD3
0.919	ILK	RANGAP1		LMAN2	SCFD1		RPL5	
0.93	ILK	RANGAP1		LMAN2	SCFD1		RPL5	PLD3
0.949	ILK	RANGAP1		LMAN2	SCFD1	LTA4H		
0.951	ILK	RANGAP1		LMAN2	SCFD1	LTA4H		PLD3
0.96	ILK	RANGAP1		LMAN2	SCFD1	LTA4H	RPL5	
0.96	ILK	RANGAP1		LMAN2	SCFD1	LTA4H	RPL5	PLD3
0.928	ILK	RANGAP1	EIF1B					
0.931	ILK	RANGAP1	EIF1B					PLD3
0.93	ILK	RANGAP1	EIF1B				RPL5	
0.933	ILK	RANGAP1	EIF1B				RPL5	PLD3
0.931	ILK	RANGAP1	EIF1B			LTA4H		
0.933	ILK	RANGAP1	EIF1B			LTA4H		PLD3
0.937	ILK	RANGAP1	EIF1B			LTA4H	RPL5	
0.938	ILK	RANGAP1	EIF1B			LTA4H	RPL5	PLD3
0.933	ILK	RANGAP1	EIF1B		SCFD1			
0.94	ILK	RANGAP1	EIF1B		SCFD1			PLD3
0.934	ILK	RANGAP1	EIF1B		SCFD1		RPL5	
0.94	ILK	RANGAP1	EIF1B		SCFD1		RPL5	PLD3
0.94	ILK	RANGAP1	EIF1B		SCFD1	LTA4H		
0.943	ILK	RANGAP1	EIF1B		SCFD1	LTA4H		PLD3
0.942	ILK	RANGAP1	EIF1B		SCFD1	LTA4H	RPL5	
0.944	ILK	RANGAP1	EIF1B		SCFD1	LTA4H	RPL5	PLD3
0.944	ILK	RANGAP1	EIF1B	LMAN2				
0.946	ILK	RANGAP1	EIF1B	LMAN2				PLD3
0.948	ILK	RANGAP1	EIF1B	LMAN2			RPL5	
0.95	ILK	RANGAP1	EIF1B	LMAN2			RPL5	PLD3
0.946	ILK	RANGAP1	EIF1B	LMAN2		LTA4H		

Rsquared	ILK	RANGAP1	EIF1B	LMAN2	SCFD1	LTA4H	RPL5	PLD3
0.947	ILK	RANGAP1	EIF1B	LMAN2		LTA4H		PLD3
0.956	ILK	RANGAP1	EIF1B	LMAN2		LTA4H	RPL5	
0.956	ILK	RANGAP1	EIF1B	LMAN2		LTA4H	RPL5	PLD3
0.954	ILK	RANGAP1	EIF1B	LMAN2	SCFD1			
0.96	ILK	RANGAP1	EIF1B	LMAN2	SCFD1			PLD3
0.955	ILK	RANGAP1	EIF1B	LMAN2	SCFD1		RPL5	
0.96	ILK	RANGAP1	EIF1B	LMAN2	SCFD1		RPL5	PLD3
0.96	ILK	RANGAP1	EIF1B	LMAN2	SCFD1	LTA4H		
0.962	ILK	RANGAP1	EIF1B	LMAN2	SCFD1	LTA4H		PLD3
0.964	ILK	RANGAP1	EIF1B	LMAN2	SCFD1	LTA4H	RPL5	
0.965	ILK	RANGAP1	EIF1B	LMAN2	SCFD1	LTA4H	RPL5	PLD3

What is claimed is:

1. A method of assessing quality of a sample collected from a subject comprising detecting the level of each of N biomarker proteins in the sample, wherein N is at least 1, and wherein at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, or at least 7 of the N biomarker proteins are selected from LKHA4, EIF1B, RAGP1, SCFD1, PLD3, Lectin, mannose-binding 2, ILK1, and RPL5, wherein the sample is a serum sample.

2. A method comprising:

a) measuring the level of each of N biomarker proteins in a serum sample from a subject, wherein N is at least 1, and wherein at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, or at least 7 of the N biomarker proteins are selected from LKHA4, EIF1B, RAGP1, SCFD1, PLD3, Lectin, mannose-binding 2, ILK1, and RPL5; and

b) identifying the sample as an analysis sample or negative sample based on the level of the N biomarker proteins;

wherein the analysis sample is a sample that is suitable for use in one or more of the following: protein biomarker discovery analysis, protein expression level analysis, a diagnostic method or a prognostic method, and the negative sample is a sample that is not suitable for use as an analysis sample.

3. A method comprising:

a) contacting a serum sample from a subject with a set of capture reagents, wherein each capture reagent has affinity for a different biomarker protein of N biomarker proteins wherein N is at least 1, and wherein at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, or at least 7 of the N biomarker proteins is selected from LKHA4, EIF1B, RAGP1, SCFD1, PLD3, Lectin, mannose-binding 2, ILK1, and RPL5; and

b) measuring the level of each N biomarker protein with the set of capture reagents.

4. The method of any one of claims 1 to 3, wherein 2 of the N biomarker proteins are LKHA4 and EIF1B.

5. The method of any one of claims 1 to 3, wherein 2 of the N biomarker proteins are LKHA4 and RAGP1.

6. The method of any one of claims 1 to 3, wherein 2 of the N biomarker proteins are LKHA4 and SCFD1.

7. The method of any one of claims 1 to 3, wherein 2 of the N biomarker proteins are LKHA4 and PLD3.

8. The method of any one of claims 1 to 3, wherein 2 of the N biomarker proteins are LKHA4 and Lectin, mannose-binding 2.
9. The method of any one of claims 1 to 3, wherein 2 of the N biomarker proteins are LKHA4 and ILK1.
10. The method of any one of claims 1 to 3, wherein 2 of the N biomarker proteins are LKHA4 and RPL5.
11. The method of any one of claims 1 to 3, wherein 2 of the N biomarker proteins are EIF1B and RAGP1.
12. The method of any one of claims 1 to 3, wherein 2 of the N biomarker proteins are EIF1B and SCFD1.
13. The method of any one of claims 1 to 3, wherein 2 of the N biomarker proteins are EIF1B and PLD3.
14. The method of any one of claims 1 to 3, wherein 2 of the N biomarker proteins are EIF1B and Lectin, mannose-binding 2.
15. The method of any one of claims 1 to 3, wherein 2 of the N biomarker proteins are EIF1B and ILK1.
16. The method of any one of claims 1 to 3, wherein 2 of the N biomarker proteins are EIF1B and RPL5.
17. The method of any one of claims 1 to 3, wherein 2 of the N biomarker proteins are RAGP1 and SCFD1.
18. The method of any one of claims 1 to 3, wherein 2 of the N biomarker proteins are RAGP1 and PLD3.
19. The method of any one of claims 1 to 3, wherein 2 of the N biomarker proteins are RAGP1 and Lectin, mannose-binding 2.
20. The method of any one of claims 1 to 3, wherein 2 of the N biomarker proteins are RAGP1 and ILK1.
21. The method of any one of claims 1 to 3, wherein 2 of the N biomarker proteins are RAGP1 and RPL5.
22. The method of any one of claims 1 to 3, wherein 2 of the N biomarker proteins are SCFD1 and PLD3.
23. The method of any one of claims 1 to 3, wherein 2 of the N biomarker proteins are SCFD1 and Lectin, mannose-binding 2.
24. The method of any one of claims 1 to 3, wherein 2 of the N biomarker proteins are SCFD1 and ILK1.

25. The method of any one of claims 1 to 3, wherein 2 of the N biomarker proteins are SCFD1 and RPL5.
26. The method of any one of claims 1 to 3, wherein 2 of the N biomarker proteins are PLD3 and Lectin, mannose-binding 2.
27. The method of any one of claims 1 to 3, wherein 2 of the N biomarker proteins are PLD3 and ILK1.
28. The method of any one of claims 1 to 3, wherein 2 of the N biomarker proteins are PLD3 and RPL5.
29. The method of any one of claims 1 to 3, wherein 2 of the N biomarker proteins are Lectin, mannose-binding 2 and ILK1.
30. The method of any one of claims 1 to 3, wherein 2 of the N biomarker proteins are Lectin, mannose-binding 2 and RPL5.
31. The method of any one of claims 1 to 3, wherein 2 of the N biomarker proteins are ILK1 and RPL5.
32. The method of any one of claims 1 to 31, wherein N is 3, N is 4, N is 5, N is 6, N is 7, or N is 8.
33. The method of claim 32, wherein all of the N biomarker proteins are selected from LKHA4, EIF1B, RAGP1, SCFD1, PLD3, Lectin, mannose-binding 2, ILK1, and RPL5.
34. The method of any one of claims 1-33, wherein the subject is a human subject.
35. The method of any one of claims 1-34, wherein the sample was processed, frozen, and thawed after the sample collection and prior to the detecting.
36. The method of claim 35, wherein the sample processing steps comprised at least one freeze-thaw cycle.
37. The method of claim 35, wherein the sample processing steps comprised greater than one freeze-thaw cycle.
38. The method of any one of claims 35-37, comprising determining a number of freeze-thaw cycles for the sample.
39. The method of claim 38, wherein the determining is based on comparing the detected levels of the N biomarker proteins to reference levels, wherein the reference levels are average levels of the N biomarker proteins present in samples having 1 freeze-thaw cycle or 2 freeze-thaw cycles.
40. The method of claims 1-39, wherein the detected levels of the N biomarker proteins compared to reference levels indicates that the number of freeze-thaw cycles was

greater than 1, greater than 2, greater than 3, greater than 4, greater than 5, greater than 7 or greater than 10 freeze-thaw cycles.

41. The method of any one of claims 38-40, wherein the determining is based on a panel of N biomarker proteins having an R^2 value of at least 0.65, at least 0.7, at least 0.75, at least 0.8, at least 0.85, at least 0.9, or at least 0.95.

42. The method of any one of claims 38-41, comprising performing protein biomarker discovery analysis, protein expression level analysis, a diagnostic method or a prognostic method on the sample.

43. The method of any one of claims 1-42, comprising identifying the sample as passing a quality assessment or failing a quality assessment.

44. The method of claim 43, wherein the identifying is based, at least in part, on the detected levels of the N biomarker proteins.

45. The method of claim 43 or 44, wherein the sample is identified as passing if the number of freeze-thaw cycles is determined to be 1, less than 2, less than 3, less than 4, less than 5, or 5.

46. The method of claim 43 or 44, wherein the sample is identified as failing if the number of freeze-thaw cycles is determined to be greater than 5, greater than 6, greater than 7, greater than 8, greater than 9 or greater than 10.

47. The method of any one of claims 43-46, comprising either a) performing further analysis of the sample if it is identified as passing the quality assessment; or b) discarding the sample if it is identified as failing the quality assessment.

48. The method of any one of claims 1-47, comprising detecting the level of each of N biomarkers in a plurality of samples from a plurality of subjects.

49. A method for comparing a plurality of samples collected from a plurality of subjects comprising detecting the level of each of N biomarker proteins in each of the plurality of samples, wherein N is at least 1, and wherein at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, or at least 7 of the N biomarker proteins are selected LKHA4, EIF1B, RAGP1, SCFD1, PLD3, Lectin, mannose-binding 2, ILK1, and RPL5, wherein the sample is a serum sample.

50. The method of claim 49, comprising a) determining the number of freeze-thaw cycles and b) comparing the determined number freeze-thaw cycles for each of the plurality of samples.

51. The method of claim 50, comprising identifying the plurality of samples as handled consistently or inconsistently, wherein samples handled consistently all have a

determined number of freeze-thaw cycles within 1, 2, 3, 4 or 5 freeze-thaw cycles of each other.

52. The method of claim 50 or 51, wherein the determining is based on comparing the detected levels of each of the N biomarker proteins to reference levels, wherein the reference levels are average levels of each of the N biomarker proteins present in samples having 1 freeze-thaw cycle or 2 freeze-thaw cycles.

53. The method of any one of claims 49-52, wherein the detected levels of each of the N biomarker proteins compared to the reference levels indicates that the number of freeze-thaw cycles was greater than 1, greater than 2, greater than 3, greater than 4, greater than 5, greater than 7 or greater than 10 freeze-thaw cycles; or wherein the levels of each of the N biomarker proteins used in a linear regression model predicts the number of freeze-thaw cycles was greater than 1, greater than 2, greater than 3, greater than 4, greater than 5, greater than 7 or greater than 10 freeze-thaw cycles.

54. The method of any one of claims 49-53, wherein the determining is based on a panel of N biomarker proteins having an R^2 value of at least 0.65, at least 0.7, at least 0.75, at least 0.8, at least 0.85, at least 0.9, or at least 0.95.

55. The method of any one of claims 49-54, comprising performing protein biomarker discovery analysis, protein expression level analysis, a diagnostic method or a prognostic method on the plurality of samples.

56. The method of any one of claims 49-55, comprising modifying a panel of biomarker proteins in a protein biomarker discovery analysis, a protein expression level analysis, a diagnostic method or a prognostic method based on the determined number of freeze-thaw cycles for each of the plurality of samples; or identifying one or more proteins in the sample as being affected by the number of freeze-thaw cycles; or identifying the level of one or more proteins in the sample as being affected by the number of freeze-thaw cycles; or changing the panel of biomarker proteins used in a diagnostic, a prognostic or a health assessment related test based on the predicted number of freeze-thaw cycles; removing the proteins used in a diagnostic, a prognostic or a health assessment related test based on the predicted number of freeze-thaw cycles.

57. The method of claim 56, wherein the panel of biomarker proteins is reduced in number of biomarker proteins measured.

58. The method of any one of claims 49-57, wherein the determining measures compliance with a clinical trial sample collection and processing protocol.

59. The method of any one of claims 49-57, wherein the plurality of samples are collected at more than one sample collection site.

60. The method of claim 59, wherein the plurality of samples from a first sample collection site are compared to a second plurality of samples from a second sample collection site.

61. The method of any one of claims 49-60, wherein one or more of the plurality of samples may be removed based on the number of freeze-thaw cycles.

62. The method of any one of claims 49-61,
wherein 2 of the N biomarker proteins are LKHA4 and EIF1B; or
wherein 2 of the N biomarker proteins are LKHA4 and RAGP1; or
wherein 2 of the N biomarker proteins are LKHA4 and SCFD1; or
wherein 2 of the N biomarker proteins are LKHA4 and PLD3; or
wherein 2 of the N biomarker proteins are LKHA4 and Lectin, mannose-binding 2; or
wherein 2 of the N biomarker proteins are LKHA4 and ILK1; or
wherein 2 of the N biomarker proteins are LKHA4 and RPL5; or
wherein 2 of the N biomarker proteins are EIF1B and RAGP1; or
wherein 2 of the N biomarker proteins are EIF1B and SCFD1; or
wherein 2 of the N biomarker proteins are EIF1B and PLD3; or
wherein 2 of the N biomarker proteins are EIF1B and Lectin, mannose-binding 2; or
wherein 2 of the N biomarker proteins are EIF1B and ILK1; or
wherein 2 of the N biomarker proteins are EIF1B and RPL5; or
wherein 2 of the N biomarker proteins are RAGP1 and SCFD1; or
wherein 2 of the N biomarker proteins are RAGP1 and PLD3; or
wherein 2 of the N biomarker proteins are RAGP1 and Lectin, mannose-binding 2; or
wherein 2 of the N biomarker proteins are RAGP1 and ILK1; or
wherein 2 of the N biomarker proteins are RAGP1 and RPL5; or
wherein 2 of the N biomarker proteins are SCFD1 and PLD3; or
wherein 2 of the N biomarker proteins are SCFD1 and Lectin, mannose-binding 2; or
wherein 2 of the N biomarker proteins are SCFD1 and ILK1; or
wherein 2 of the N biomarker proteins are SCFD1 and RPL5; or
wherein 2 of the N biomarker proteins are PLD3 and Lectin, mannose-binding 2; or
wherein 2 of the N biomarker proteins are PLD3 and ILK1; or
wherein 2 of the N biomarker proteins are PLD3 and RPL5; or
wherein 2 of the N biomarker proteins are Lectin, mannose-binding 2 and ILK1; or

wherein 2 of the N biomarker proteins are Lectin, mannose-binding 2 and RPL5; or
wherein 2 of the N biomarker proteins are ILK1 and RPL5.

63. The method of any one of claims 49-62, wherein N is 3, N is 4, N is 5, N is 6, N is 7, or N is 8.

64. The method of claim 63, wherein all of the N biomarker proteins are selected from LKHA4, EIF1B, RAGP1, SCFD1, PLD3, Lectin, mannose-binding 2, ILK1, and RPL5.

65. The method of any one of claims 49-64, wherein the subject is a human subject.

66. The method of any one of claims 49-65, wherein the sample was processed, frozen, and thawed after the sample collection and prior to the detecting.

67. The method of claim 66, wherein the sample processing steps comprised at least one freeze-thaw cycle.

68. The method of claim 66, wherein the sample processing steps comprised greater than one freeze-thaw cycle.

69. The method of any one of claims 1-68, wherein the detecting comprises performing mass spectrometry, an aptamer based assay, and/or an antibody based assay.

70. The method of any one of claims 1-69, wherein the method comprises contacting biomarker proteins of the sample from the subject with a set of capture reagents, wherein each capture reagent of the set of capture reagents specifically binds to one biomarker protein being detected.

71. The method of claim 70, wherein each the capture reagents specifically binds to a different biomarker protein being detected.

72. The method of claim 70 or 71, wherein each capture reagent is an antibody or an aptamer.

73. The method of claim 72, wherein each capture reagent is an aptamer.

74. The method of claim 73, wherein at least one aptamer is a slow off-rate aptamer.

75. The method of claim 74, wherein at least one slow off-rate aptamer comprises at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, or at least 10 nucleotides with modifications.

76. The method of claim 73 or 74, wherein each slow off-rate aptamer binds to its target protein with an off rate ($t_{1/2}$) of ≥ 20 minutes, ≥ 30 minutes, ≥ 60 minutes, ≥ 90 minutes, ≥ 120 minutes, ≥ 150 minutes, ≥ 180 minutes, ≥ 210 minutes, or ≥ 240 minutes.

77. A kit comprising N biomarker protein capture reagents, wherein N is at least 1, and wherein at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, or at least 7 of the capture reagents bind to proteins selected from LKHA4, EIF1B, RAGP1, SCFD1, PLD3, Lectin, mannose-binding 2, ILK1, and RPL5.

78. The kit of claim 77, wherein N is at least 2, and wherein at least 2 of the capture reagents binds to a protein selected from LKHA4, EIF1B, RAGP1, SCFD1, PLD3, Lectin, mannose-binding 2, ILK1, and RPL5.

79. The kit of claim 77 or 78, wherein each of the capture reagents binds to a different protein.

80. The kit of any one of claims 77-79, wherein N is 2, N is 3, N is 4, N is 5, N is 6, N is 7, or N is 8.

81. The kit of claim 80, wherein each of the capture reagents binds to a protein selected from LKHA4, EIF1B, RAGP1, SCFD1, PLD3, Lectin, mannose-binding 2, ILK1, and RPL5.

82. The kit of any one of claims 77-81, wherein each of the capture reagents is an antibody or an aptamer.

83. The kit of claim 82, wherein each capture reagent is an aptamer.

84. The kit of claim 83, wherein at least one aptamer is a slow off-rate aptamer.

85. The kit of claim 84, wherein at least one slow off-rate aptamer comprises at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, or at least 10 nucleotides with modifications.

86. The kit of claim 84 or claim 85, wherein each slow off-rate aptamer binds to its target protein with an off rate ($t_{1/2}$) of ≥ 20 minutes, ≥ 30 minutes, ≥ 60 minutes, ≥ 90 minutes, ≥ 120 minutes, ≥ 150 minutes, ≥ 180 minutes, ≥ 210 minutes, or ≥ 240 minutes.

87. The kit of any one of claims 77-86, for use in detecting the N biomarker proteins in a sample from a subject.

88. The kit of claim 87, for use in assessing the quality of the sample based at least in part on the levels of the detected N biomarker proteins in the sample.

89. The kit of any one of claims 77-88, for use in determining the number of freeze-thaw cycles for the sample.

90. A method comprising detecting the level of each of N biomarker proteins in a sample, wherein N is at least 1, and wherein at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, or at least 7 of the N biomarker proteins is selected from LKHA4, EIF1B, RAGP1, SCFD1, PLD3, Lectin, mannose-binding 2, ILK1, and RPL5.

91. The method of claim 90, wherein the sample is a serum sample.
92. The method of claim 91, wherein the serum sample is a human serum sample.
93. The method of any one of claims 90-92, wherein the number of freeze-thaw cycles is determined with the level of each of the N biomarker proteins.
94. The method of claim 93, wherein the number of freeze-thaw cycles is determined to be greater than 1, greater than 2, greater than 3, greater than 4, greater than 5, greater than 7 or greater than 10 freeze-thaw cycles.
95. The method of claim 94, wherein the determined number of freeze-thaw cycles is derived from input of the level of each of the N biomarker proteins in a statistical model.
96. The method of claim 95, wherein the statistical model is a linear regression model.
97. The method of any one of claims 90-96, wherein the level of each protein measured is determined from a relative florescence unit (RFU) or a protein concentration.
98. A method comprising detecting the level of each of at least 1, 2, 3, 4, 5, 6, 7, or 8 biomarker proteins in a sample, wherein biomarker proteins are selected from LKHA4, EIF1B, RAGP1, SCFD1, PLD3, Lectin, mannose-binding 2, ILK1, and RPL5.
99. The method of claim 98, wherein the sample is a serum sample.
100. The method of claim 99, wherein the serum sample is a human serum sample.
101. The method of any one of claims 98-100, wherein a number of freeze-thaw cycles is determined with the level of each of the at least 1, 2, 3, 4, 5, 6, 7 or 8 biomarker proteins.
102. The method of claim 101, wherein the number of freeze-thaw cycles is determined to be greater than 1, greater than 2, greater than 3, greater than 4, greater than 5, greater than 7 or greater than 10 freeze-thaw cycles.
103. The method of claim 102, wherein the determined number of freeze-thaw cycles determined to be greater than 1, greater than 2, greater than 3, greater than 4, greater than 5, greater than 7 or greater than 10 freeze-thaw cycles is derived from input of the level of at least 1, 2, 3, 4, 5, 6, 7, or 8 biomarker proteins in a statistical model.
104. The method of claim 103, wherein the statistical model is a linear regression model.
105. The method of claim 103, further comprising modifying a panel of proteins in a protein biomarker discovery analysis, a protein expression level analysis, a diagnostic method or a prognostic method; identifying one or more proteins in the sample as being

affected; identifying the level of one or more proteins in the sample as being affected; changing the proteins used in a diagnostic, a prognostic or a health assessment related test; or removing one or more proteins used in a diagnostic, a prognostic or a health assessment related test, each based on the outcome of the linear regression model.