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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
19 November 2009 (19.11.2009)

PCT

(10) International Publication Number
WO 2009/140390 A2

(51) International Patent Classification:
G01N 33/00 (2006.01)

(21) International Application Number:
PCT/US2009/043794

(22) International Filing Date:
13 May 2009 (13.05.2009)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
12/120,043 13 May 2008 (13.05.2008) US

(71) Applicant *(for all designated States except US)*: **BATTELLE MEMORIAL INSTITUTE** [US/US]; 902 Battelle Boulevard, PO Box 999, Richland, WA 99352 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **METZ, Thomas, O.** [US/US]; 8809 West 1st Avenue, Kennewick, WA 99336 (US). **QIAN, Wei-Jun** [US/US]; 1711 April Loop, Richland, WA 99354 (US). **JACOBS, Jon, M.** [US/US]; 2100 N Road 76, Pasco, WA 99301 (US). **POLPITIYA, Ashoka, D.** [US/US]; 2784 Sawgrass Loop, Richland, WA 99354 (US). **CAMP, David, G.** [US/US]; 2056 Hudson Avenue, Richland, WA 99354 (US). **SMITH, Richard, D.** [US/US]; 2625 Thoroughbred Way, Richland, WA 99352 (US).

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

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

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

Declarations under Rule 4.17:

— *of inventorship (Rule 4.17(iv))*

Published:

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

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

WO 2009/140390 A2

(54) Title: SERUM MARKERS FOR TYPE II DIABETES MELLITUS

(57) Abstract: A method for identifying persons with increased risk of developing type 2 diabetes mellitus, or having type II diabetes mellitus, utilizing selected biomarkers described herein either alone or in combination. The present disclosure allows for broad based, reliable, screening of large population bases and provides other advantages, including the formulation of effective strategies for characterizing, archiving, and contrasting data from multiple sample types under varying conditions. Also provided are arrays and kits that can be used to perform such methods.

SERUM MARKERS FOR TYPE II DIABETES MELLITUS

CROSS-REFERENCE TO RELATED APPLICATION

This application claims priority to U.S. Application No. 12/120,043 filed May 13, 2008, herein incorporated by reference.

FIELD

The present disclosure generally relates to methods and systems for the screening and detection of disease and more particularly to methods and systems for the screening and detection of persons at risk for developing type II diabetes mellitus.

STATEMENT AS TO RIGHTS TO INVENTIONS MADE UNDER FEDERALLY-SPONSORED RESEARCH AND DEVELOPMENT

This invention was made with Government support under Contract DE-AC0576RLO1830 awarded by the U.S. Department of Energy. The Government has certain rights in the invention.

BACKGROUND

Type II diabetes mellitus is a life changing disease that affects millions of persons. While the disease may be clinically diagnosed and confirmed fairly easily in some cases, earlier detection may lead to the possibility of intervention(s) that would alter or lessen the onset of clinical symptoms or allow other forms of preventative care to be undertaken.

Currently, one of the best approaches for predicting who may be at risk for developing type II diabetes mellitus before onset of clinical symptoms is by the oral glucose tolerance test (OGTT). The OGTT, is inconvenient, requires fasting and is not highly reproducible. The fasting blood glucose (FBG) is less burdensome, but much less sensitive, particularly in older Americans who have the highest prevalence of diabetes and pre-diabetes. The quantification of Hemoglobin A1c (a glycated form of hemoglobin) from blood has been widely used as a test for assessing the adequacy of glycemic control and risk of complications in diabetic

patients, but this test is not sufficiently sensitive to detect the range of glucose values typically seen in pre-diabetes or new onset type II diabetes. Furthermore, there are many variants of hemoglobin present in blood. This is particularly applicable in minority populations disproportionately affected by diabetes, and this adds additional uncertainty to the use of this test. A simplified, less burdensome approach to the diagnosis of diabetes and pre-diabetes would facilitate increased recognition and improved care of these individuals. While progress has been made to improve the accuracy and reproducibility of the measurement of surrogate biomarkers predictive of those at high risk for developing type II diabetes, a set of candidate biomarkers would benefit the clinical community, particularly if such surrogate biomarkers result in higher sensitivity and specificity.

Accordingly, what is needed is a method and system of screening for persons with increased risk of developing type II diabetes mellitus that utilizes candidate biomarkers allowing for broad based, reliable screening of large population bases. In addition, effective strategies for characterizing, archiving, and contrasting data from multiple sample types under varying conditions (*e.g.*, control versus disease) are also needed.

SUMMARY

The present application describes methods for identifying persons with increased risk of developing type II diabetes mellitus (or having type II diabetes mellitus) utilizing selected type II diabetes mellitus-related biomarkers described herein either alone or in combination. The present methods allow for broad based, reliable screening of large population bases and provides other advantages including the formulation of effective strategies for characterizing, archiving, and contrasting data from multiple sample types under varying conditions.

In one example, the type II diabetes mellitus-related markers are selected from those shown in any of Tables 1-4, such as those shown in Table 3 or 4. Detecting the markers can include detecting the full-length proteins or the individual peptides shown in the Tables (for example by using antibodies or mass spectrometry). In some examples, the markers in a test sample are quantified, for example relative to a control or values (such as a value or range of values expected

in a subject having or not having type II diabetes, or predisposed to developing such a disorder).

The presence of the markers described herein may be determined and utilized in a variety of ways employing various methodologies and utilizing a variety of sample types. Thus while one embodiment related to the application of the markers of the present invention in serum or plasma is described, this description is intended to be illustrative only and not exclusive in any way. With appropriate modification such a method may also be utilized in other sample types. In this and comparing the quantity of the at least one serum constituent to a standardized range of levels for this constituent to determine whether the level of the serum constituent when compared to the normal range is indicative of a predisposition for type II diabetes mellitus. A system for performing the method of the present invention is made up of the requisite pieces and parts that would allow such a method to be performed.

Also provided are arrays and kits that include probes for detecting the disclosed type II diabetes mellitus-related markers.

Additional features of the present disclosure will be set forth as follows and will be readily apparent from the descriptions and demonstrations set forth herein. Accordingly, the following descriptions of the present disclosure should be seen as illustrative of the disclosure and not as limiting in any way.

SEQUENCE LISTING

The protein sequences listed in the accompanying sequence listing are shown using standard three-letter abbreviations for amino acids.

DETAILED DESCRIPTION OF SEVERAL EMBODIMENTS

Abbreviations and Terms

The following explanations of terms and methods are provided to better describe the present disclosure and to guide those of ordinary skill in the art in the practice of the present disclosure. The singular forms “a,” “an,” and “the” refer to one or more than one, unless the context clearly dictates otherwise. For example, the term “comprising a protein” includes single or plural proteins and is considered equivalent to the phrase “comprising at least one protein.” The term “or” refers to a

single element of stated alternative elements or a combination of two or more elements, unless the context clearly indicates otherwise. As used herein, "comprises" means "includes." Thus, "comprising A or B," means "including A, B, or A and B," without excluding additional elements.

Unless explained otherwise, all technical and scientific terms used herein have the same meaning as commonly understood to one of ordinary skill in the art to which this disclosure belongs. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present disclosure, suitable methods and materials are described below. The materials, methods, and examples are illustrative only and not intended to be limiting. While the methods have been described and were utilized in testing, it is to be distinctly understood that the disclosure is not limited to any particular form of testing utilized, but is intended to include all methods that are capable of detecting the materials that are set forth in the claims.

Antibody: A polypeptide ligand including at least a light chain or heavy chain immunoglobulin variable region which specifically recognizes and binds an epitope of an antigen, such as an endothelial marker or a fragment thereof. Antibodies are composed of a heavy and a light chain, each of which has a variable region, termed the variable heavy (V_H) region and the variable light (V_L) region. Together, the V_H region and the V_L region are responsible for binding the antigen recognized by the antibody. In one example, an antibody specifically binds to one of the proteins or peptides listed in Tables 1-4, but not other proteins (such as other proteins found in human serum or plasma).

This includes intact immunoglobulins and the variants and portions of them well known in the art, such as Fab' fragments, $F(ab)'_2$ fragments, single chain Fv proteins ("scFv"), and disulfide stabilized Fv proteins ("dsFv"). A scFv protein is a fusion protein in which a light chain variable region of an immunoglobulin and a heavy chain variable region of an immunoglobulin are bound by a linker, while in dsFvs, the chains have been mutated to introduce a disulfide bond to stabilize the association of the chains. The term also includes genetically engineered forms such as chimeric antibodies (for example, humanized murine antibodies), heteroconjugate

antibodies (such as, bispecific antibodies). See also, *Pierce Catalog and Handbook*, 1994-1995 (Pierce Chemical Co., Rockford, IL); Kuby, J., *Immunology*, 3rd Ed., W.H. Freeman & Co., New York, 1997.

Typically, a naturally occurring immunoglobulin has heavy (H) chains and light (L) chains interconnected by disulfide bonds. There are two types of light chain, lambda (λ) and kappa (k). There are five main heavy chain classes (or isotypes) which determine the functional activity of an antibody molecule: IgM, IgD, IgG, IgA and IgE.

Each heavy and light chain contains a constant region and a variable region, (the regions are also known as "domains"). In combination, the heavy and the light chain variable regions specifically bind the antigen. Light and heavy chain variable regions contain a "framework" region interrupted by three hypervariable regions, also called "complementarity-determining regions" or "CDRs". The extent of the framework region and CDRs have been defined (see, Kabat *et al.*, *Sequences of Proteins of Immunological Interest*, U.S. Department of Health and Human Services, 1991, which is hereby incorporated by reference). The Kabat database is now maintained online. The sequences of the framework regions of different light or heavy chains are relatively conserved within a species. The framework region of an antibody, that is the combined framework regions of the constituent light and heavy chains, serves to position and align the CDRs in three-dimensional space.

The CDRs are primarily responsible for binding to an epitope of an antigen. The CDRs of each chain are typically referred to as CDR1, CDR2, and CDR3, numbered sequentially starting from the N-terminus, and are also typically identified by the chain in which the particular CDR is located. Thus, a V_H CDR3 is located in the variable domain of the heavy chain of the antibody in which it is found, whereas a V_L CDR1 is the CDR1 from the variable domain of the light chain of the antibody in which it is found. An antibody that binds RET will have a specific V_H region and the V_L region sequence, and thus specific CDR sequences. Antibodies with different specificities (*i.e.* different combining sites for different antigens) have different CDRs. Although it is the CDRs that vary from antibody to antibody, only a limited number of amino acid positions within the CDRs are directly involved in antigen

binding. These positions within the CDRs are called specificity determining residues (SDRs).

References to "V_H" or "VH" refer to the variable region of an immunoglobulin heavy chain, including that of an Fv, scFv, dsFv or Fab.

References to "V_L" or "VL" refer to the variable region of an immunoglobulin light chain, including that of an Fv, scFv, dsFv or Fab.

A "monoclonal antibody" is an antibody produced by a single clone of B-lymphocytes or by a cell into which the light and heavy chain genes of a single antibody have been transfected. Monoclonal antibodies are produced by methods known to those of skill in the art, for instance by making hybrid antibody-forming cells from a fusion of myeloma cells with immune spleen cells. Monoclonal antibodies include humanized monoclonal antibodies.

A "polyclonal antibody" is an antibody that is derived from different B-cell lines. Polyclonal antibodies are a mixture of immunoglobulin molecules secreted against a specific antigen, each recognizing a different epitope. These antibodies are produced by methods known to those of skill in the art, for instance, by injection of an antigen into a suitable mammal (such as a mouse, rabbit or goat) that induces the B-lymphocytes to produce IgG immunoglobulins specific for the antigen which are then purified from the mammal's serum.

A "chimeric antibody" has framework residues from one species, such as human, and CDRs (which generally confer antigen binding) from another species, such as a murine antibody that specifically binds an endothelial marker.

A "humanized" immunoglobulin is an immunoglobulin including a human framework region and one or more CDRs from a non-human (for example a mouse, rat, or synthetic) immunoglobulin. The non-human immunoglobulin providing the CDRs is termed a "donor," and the human immunoglobulin providing the framework is termed an "acceptor." In one embodiment, all the CDRs are from the donor immunoglobulin in a humanized immunoglobulin. Constant regions need not be present, but if they are, they are substantially identical to human immunoglobulin constant regions, *e.g.*, at least about 85-90%, such as about 95% or more identical. Hence, all parts of a humanized immunoglobulin, except possibly the CDRs, are substantially identical to corresponding parts of natural human immunoglobulin

sequences. Humanized immunoglobulins can be constructed by means of genetic engineering (see for example, U.S. Patent No. 5,585,089).

Array: An arrangement of molecules, such as biological macromolecules (such as peptides or nucleic acid molecules), in addressable locations on or in a substrate. A "microarray" is an array that is miniaturized so as to require or be aided by microscopic examination for evaluation or analysis.

The array of molecules ("features") makes it possible to carry out a very large number of analyses on a sample at one time. In certain example arrays, one or more molecules (such as oligonucleotide probes or antibodies) will occur on the array a plurality of times (such as twice), for instance to provide internal controls. The number of addressable locations on the array can vary, for example from at least four, to at least 10, at least 20, at least 30, at least 50, at least 75, at least 100, at least 150, at least 200, at least 300, at least 500, least 550, at least 600, at least 800, at least 1000, at least 10,000, or more. In particular examples, an array includes nucleic acid molecules, such as oligonucleotide sequences that are at least 15 nucleotides in length, such as about 15-40 nucleotides in length.

In particular examples, an array includes oligonucleotide probes or primers which can be used to detect type II diabetes mellitus-associated nucleic acids. Protein-based arrays include probe molecules that are or include proteins, or where the target molecules are or include proteins, and arrays including nucleic acids to which proteins are bound, or vice versa. For example, such arrays can be used to detect any combination of at least 4 of the type II diabetes mellitus-related proteins (or nucleic acids encoding such proteins) listed in any of Tables 1-4, such as at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 15, at least 20, at least 22, at least 25, at least 30, at least 35, at least 40, or at least 45 of the molecules listed in Table 1, 2, 3, or 4.

Within an array, each arrayed sample is addressable, in that its location can be reliably and consistently determined within at least two dimensions of the array. The feature application location on an array can assume different shapes. For example, the array can be regular (such as arranged in uniform rows and columns) or irregular. Thus, in ordered arrays the location of each sample is assigned to the sample at the time when it is applied to the array, and a key may be provided in

order to correlate each location with the appropriate target or feature position. Often, ordered arrays are arranged in a symmetrical grid pattern, but samples could be arranged in other patterns (such as in radially distributed lines, spiral lines, or ordered clusters). Addressable arrays usually are computer readable, in that a computer can be programmed to correlate a particular address on the array with information about the sample at that position (such as hybridization or binding data, including for instance signal intensity). In some examples of computer readable formats, the individual features in the array are arranged regularly, for instance in a Cartesian grid pattern, which can be correlated to address information by a computer.

Binding affinity: Affinity of one molecule for another, such as an antibody for an antigen (for example, the proteins shown in Tables 1-4). In one example, affinity is calculated by a modification of the Scatchard method described by Frankel *et al.*, *Mol. Immunol.*, 16:101-106, 1979. In another example, binding affinity is measured by an antigen/antibody dissociation rate. In yet another example, a high binding affinity is measured by a competition radioimmunoassay. In several examples, a high binding affinity is at least about 1 x 10^{-8} M. In other examples, a high binding affinity is at least about 1.5 x 10^{-8}, at least about 2.0 x 10^{-8}, at least about 2.5 x 10^{-8}, at least about 3.0 x 10^{-8}, at least about 3.5 x 10^{-8}, at least about 4.0 x 10^{-8}, at least about 4.5 x 10^{-8}, or at least about 5.0 x 10^{-8} M.

Diabetes mellitus: A disease caused by a relative or absolute lack of insulin leading to uncontrolled carbohydrate metabolism, commonly simplified to "diabetes," though diabetes mellitus should not be confused with diabetes insipidus. As used herein, "diabetes" refers to diabetes mellitus, unless otherwise indicated. A "diabetic condition" includes pre-diabetes and diabetes. Type 1 diabetes (sometimes referred to as "insulin-dependent diabetes" or "juvenile-onset diabetes") is an auto-immune disease characterized by destruction of the pancreatic β cells that leads to a total or near total lack of insulin. In type 2 diabetes (T2DM; sometimes referred to as "non-insulin-dependent diabetes" or "adult-onset diabetes"), the body does not respond to insulin, though it is present.

Symptoms of diabetes include: excessive thirst (polydipsia); frequent urination (polyuria); extreme hunger or constant eating (polyphagia); unexplained

weight loss; presence of glucose in the urine (glycosuria); tiredness or fatigue; changes in vision; numbness or tingling in the extremities (hands, feet); slow-healing wounds or sores; and abnormally high frequency of infection. Diabetes may be clinically diagnosed by a fasting plasma glucose (FPG) concentration of greater than or equal to 7.0 mmol/L (126 mg/dL), or a plasma glucose concentration of greater than or equal to 11.1 mmol/L (200 mg/dL) at about two hours after an oral glucose tolerance test (OGTT) with a 75 g load. A more detailed description of diabetes may be found in *Cecil Textbook of Medicine*, J.B. Wyngaarden, et al., eds. (W.B. Saunders Co., Philadelphia, 1992, 19th ed.).

The methods disclosed herein provide a means of identifying a subject who has type II diabetes or type II pre-diabetes. A "non-diabetic" or "normal" subject does not have any form of diabetes, such as type II diabetes or pre-diabetes.

Expression: The process by which the coded information of a gene is converted into an operational, non-operational, or structural part of a cell, such as the synthesis of a protein. Gene expression can be influenced by external signals (such as a hormone). Expression of a gene also can be regulated anywhere in the pathway from DNA to RNA to protein. Regulation can include controls on transcription, translation, RNA transport and processing, degradation of intermediary molecules such as mRNA, or through activation, inactivation, compartmentalization or degradation of specific protein molecules after they are produced.

The expression of a type II diabetes mellitus-related nucleic acid molecule or protein can be altered relative to a normal (wild type) nucleic acid molecule or protein (such as in a patient not having type II diabetes mellitus or who is prediabetic). Alterations in gene expression, such as differential expression, include but are not limited to: (1) overexpression (*e.g.*, upregulation); (2) underexpression (*e.g.*, downregulation); or (3) suppression of expression. Alternations in the expression of a nucleic acid molecule can be associated with, and in fact cause, a change in expression of the corresponding protein.

Protein expression can also be altered in some manner to be different from the expression of the protein in a normal (wild type) situation. This includes but is not necessarily limited to: (1) a mutation in the protein such that one or more of the

amino acid residues is different; (2) a short deletion or addition of one or a few (such as no more than 10-20) amino acid residues to the sequence of the protein; (3) a longer deletion or addition of amino acid residues (such as at least 20 residues), such that an entire protein domain or sub-domain is removed or added; (4) expression of an increased amount of the protein compared to a control or standard amount (*e.g.*, upregulation); (5) expression of a decreased amount of the protein compared to a control or standard amount (*e.g.*, downregulation); (6) alteration of the subcellular localization or targeting of the protein; (7) alteration of the temporally regulated expression of the protein (such that the protein is expressed when it normally would not be, or alternatively is not expressed when it normally would be); (8) alteration in stability of a protein through increased longevity in the time that the protein remains localized in a cell; and (9) alteration of the localized (such as organ or tissue specific or subcellular localization) expression of the protein (such that the protein is not expressed where it would normally be expressed or is expressed where it normally would not be expressed), each compared to a control or standard. Controls or standards for comparison to a sample, for the determination of differential expression, include samples believed to be normal (in that they are not altered for the desired characteristic, for example a sample from a subject who does not have type II diabetes) as well as laboratory values, even though possibly arbitrarily set, keeping in mind that such values can vary from laboratory to laboratory.

Laboratory standards and values may be set based on a known or determined population value and can be supplied in the format of a graph or table that permits comparison of measured, experimentally determined values.

Label: A detectable compound. In some examples, a label is conjugated directly or indirectly to another molecule, such as an antibody or a protein, to facilitate detection of that molecule. For example, the label can be capable of detection by ELISA, spectrophotometry, flow cytometry, or microscopy. Specific, non-limiting examples of labels include fluorophores, chemiluminescent agents, enzymatic linkages, and radioactive isotopes. Methods for labeling and guidance in the choice of labels appropriate for various purposes are discussed for example in Sambrook et al. (Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, New York, 1989) and Ausubel et al. (In Current Protocols in Molecular Biology,

John Wiley & Sons, New York, 1998). In a particular example, a label is conjugated to an antibody specific for a protein or peptide disclosed in any of Tables 1-4 to allow for the detection/screening for type II diabetes mellitus.

Lateral flow device: A device that absorbs or adsorbs a liquid sample (such as a serum or plasma sample), routes that liquid sample to a detection zone, and uses antibody- or lectin-based detection methods to generate a visible signal in response to the presence or absence of a specific antigen (such as a protein or glycoprotein) or lectin-binding biomolecule (such as a glycoprotein or glycolipid). The device can be a test strip used in lateral flow chromatography, in which a test sample fluid, suspected of containing an analyte, flows (for example by capillary action) through the strip (which is frequently made of bibulous materials such as paper, nitrocellulose, and cellulose). The test fluid and any suspended analyte can flow along the strip to a detection zone in which the analyte (if present) interacts with a detection agent to indicate a presence, absence and/or quantity of the analyte.

Numerous lateral flow analytical devices have been disclosed, and include those shown in U.S. Patent Nos. 4,313,734; 4,435,504; 4,775,636; 4,703,017; 4,740,468; 4,806,311; 4,806,312; 4,861,711; 4,855,240; 4,857,453; 4,943,522; 4,945,042; 4,496,654; 5,001,049; 5,075,078; 5,126,241; 5,451,504; 5,424,193; 5,712,172; 6,258,548; 6,555,390; 6,699,722; and 6,368,876; EP 0810436; and WO 92/12428; WO 94/01775; WO 95/16207; and WO 97/06439.

Many lateral flow devices are one-step lateral flow assays in which a biological fluid is placed in a sample area on a bibulous strip (though, non-bibulous materials can be used, and rendered bibulous by applying a surfactant to the material), and allowed to migrate along the strip until the liquid comes into contact with a specific binding partner (such as a lectin or antibody) that interacts with an analyte (such as a glycoprotein, glycolipid, or antigen) in the liquid. Once the analyte interacts with the binding partner, a signal (such as a fluorescent or otherwise visible dye) indicates that the interaction has occurred. Multiple discrete binding partners can be placed on the strip (for example in parallel lines) to detect multiple analytes (such as type II diabetes-related proteins) in the liquid. The test strips can also incorporate control indicators, which provide a signal that the test has

adequately been performed, even if a positive signal indicating the presence (or absence) of an analyte is not seen on the strip.

Mammal: This term includes both human and non-human mammals. Examples of mammals include, but are not limited to: humans, pigs, cows, goats, cats, dogs, rabbits and mice.

Reference value: An amount of activity or expression determined to be representative of a given condition. Reference values can include a range of values, real or relative expected to occur under certain conditions. These values can be compared with experimental values to determine if a given molecule is up-regulated or down-regulated in a particular sample for instance.

In one example, a reference value or range of values represents an amount of activity or expression of a type II diabetes mellitus-related nucleic acid molecule or proteins in a sample, such as a sample from a non-type II diabetes mellitus patient. This value can then be used to determine if the subject from whom a test sample was obtained has type II diabetes mellitus or is predisposed to developing type II diabetes mellitus by comparing this reference value of expression to the level of expression detected in the test sample. In a particular example, a change in expression or activity of type II diabetes mellitus-related molecules (such as those in any of Tables 1-4) in a test sample as compared to such a reference value indicates that the subject has type II diabetes mellitus.

Sample: Biological specimens containing genomic DNA, cDNA, RNA, or protein obtained from the cells of a subject, such as those present in peripheral blood, urine, saliva, semen, tissue biopsy, surgical specimen, fine needle aspriates, amniocentesis samples and autopsy material. In one example, a sample includes plasma obtained from a mammalian subject.

Solid support (or substrate): Any material which is insoluble, or can be made insoluble by a subsequent reaction. The arrays and lateral flow devices disclosed herein for detecting type II diabetes mellitus-related molecules can include a solid support. Numerous and varied solid supports are known to those in the art and include, without limitation, nitrocellulose, the walls of wells of a reaction tray, multi-well plates, test tubes, polystyrene beads, magnetic beads, membranes, and microparticles (such as latex particles). Any suitable porous material with sufficient

porosity to allow access by detector reagents and a suitable surface affinity to immobilize capture reagents (*e.g.*, lectins or antibodies) is contemplated by this term. For example, the porous structure of nitrocellulose has excellent absorption and adsorption qualities for a wide variety of reagents, for instance, capture reagents. Nylon possesses similar characteristics and is also suitable. Microporous structures are useful, as are materials with gel structure in the hydrated state.

Further examples of useful solid supports include: natural polymeric carbohydrates and their synthetically modified, cross-linked or substituted derivatives, such as agar, agarose, cross-linked alginic acid, substituted and cross-linked guar gums, cellulose esters, especially with nitric acid and carboxylic acids, mixed cellulose esters, and cellulose ethers; natural polymers containing nitrogen, such as proteins and derivatives, including cross-linked or modified gelatins; natural hydrocarbon polymers, such as latex and rubber; synthetic polymers which may be prepared with suitably porous structures, such as vinyl polymers, including polyethylene, polypropylene, polystyrene, polyvinylchloride, polyvinylacetate and its partially hydrolyzed derivatives, polyacrylamides, polymethacrylates, copolymers and terpolymers of the above polycondensates, such as polyesters, polyamides, and other polymers, such as polyurethanes or polyepoxides; porous inorganic materials such as sulfates or carbonates of alkaline earth metals and magnesium, including barium sulfate, calcium sulfate, calcium carbonate, silicates of alkali and alkaline earth metals, aluminum and magnesium; and aluminum or silicon oxides or hydrates, such as clays, alumina, talc, kaolin, zeolite, silica gel, or glass (these materials may be used as filters with the above polymeric materials); and mixtures or copolymers of the above classes, such as graft copolymers obtained by initializing polymerization of synthetic polymers on a pre-existing natural polymer.

It is contemplated that porous solid supports, such as nitrocellulose, described herein are preferably in the form of sheets or strips. The thickness of such sheets or strips may vary within wide limits, for example, from about 0.01 to 0.5 mm, from about 0.02 to 0.45 mm, from about 0.05 to 0.3 mm, from about 0.075 to 0.25 mm, from about 0.1 to 0.2 mm, or from about 0.11 to 0.15 mm. The pore size of such sheets or strips may similarly vary within wide limits, for example from

about 0.025 to 15 microns, or more specifically from about 0.1 to 3 microns; however, pore size is not intended to be a limiting factor in selection of the solid support. The flow rate of a solid support, where applicable, can also vary within wide limits, for example from about 12.5 to 90 sec/cm (*i.e.*, 50 to 300 sec/4 cm), about 22.5 to 62.5 sec/cm (*i.e.*, 90 to 250 sec/4 cm), about 25 to 62.5 sec/cm (*i.e.*, 100 to 250 sec/4 cm), about 37.5 to 62.5 sec/cm (*i.e.*, 150 to 250 sec/4 cm), or about 50 to 62.5 sec/cm (*i.e.*, 200 to 250 sec/4 cm). In specific embodiments of devices described herein, the flow rate is about 62.5 sec/cm (*i.e.*, 250 sec/4 cm). In other specific embodiments of devices described herein, the flow rate is about 37.5 sec/cm (*i.e.*, 150 sec/4 cm).

The surface of a solid support may be activated by chemical processes that cause covalent linkage of an agent (*e.g.*, a capture reagent) to the support. However, any other suitable method may be used for immobilizing an agent (*e.g.*, a capture reagent) to a solid support including, without limitation, ionic interactions, hydrophobic interactions, covalent interactions and the like. The particular forces that result in immobilization of an agent on a solid phase are not important for the methods and devices described herein.

A solid phase can be chosen for its intrinsic ability to attract and immobilize an agent, such as a capture reagent (such as an antibody or oligonucleotide probe). Alternatively, the solid phase can possess a factor that has the ability to attract and immobilize an agent, such as a capture reagent. The factor can include a charged substance that is oppositely charged with respect to, for example, the capture reagent itself or to a charged substance conjugated to the capture reagent. In another embodiment, a specific binding member may be immobilized upon the solid phase to immobilize its binding partner (*e.g.*, a capture reagent). In this example, therefore, the specific binding member enables the indirect binding of the capture reagent to a solid phase material.

Except as otherwise physically constrained, a solid support may be used in any suitable shapes, such as films, sheets, strips, or plates, or it may be coated onto or bonded or laminated to appropriate inert carriers, such as paper, glass, plastic films, or fabrics.

A "**lateral flow substrate**" is any solid support or substrate that is useful in a lateral flow device.

Subject: Living multicellular vertebrate organisms, a category which includes both human and veterinary subjects that are in need of the desired diagnosis, such as diagnosis of type II diabetes mellitus. Examples include, but are not limited to: humans, apes, dogs, cats, mice, rats, rabbits, horses, pigs, and cows.

Type II Diabetes Mellitus-Related Molecules

The inventors have identified at least 45 different proteins whose expression is altered (such as upregulated or downregulated) in patients having type II diabetes mellitus, relative to patients not having type II diabetes mellitus or who are prediabetic (see Tables 1-4). These proteins and corresponding nucleic acid molecules are referred to herein as "type II diabetes mellitus-related molecules" and includes type II diabetes mellitus-related nucleic acid molecules (such as DNA, RNA, for example cDNA or mRNA) and proteins. The term includes those molecules listed in Tables 1-4. The number of proteins identified depended on the method used to analyze the data.

Based on this observation, methods are provided for screening for type II diabetes by detecting these proteins (or peptide fragments thereof), or alternatively detecting DNA or RNA nucleic acid molecules encoding these proteins. The disclosed methods provide a rapid, straightforward, and accurate screening method performed in one assay for diagnosis of type II diabetes (for example to determine if a subject is predisposed to develop type II diabetes). It allows identification of subjects who may require treatment for type II diabetes. For example, by establishing that an individual has type II diabetes, effective therapeutic measures, such as insulin therapy or diet modifications, can be instituted. Arrays and kits that can be used in such methods are also provided.

The results of shown herein using non-diabetic, diabetic and control samples showed the presence of each of these proteins at significantly elevated or decreased levels in persons diagnosed with type II diabetes, while the normal and control samples did not have these altered levels. Thus a method wherein human serum or plasma is tested for at least one protein from this list serves as an effective predictive

or diagnostic screen or test for type II diabetes. In addition, the disclosed type II diabetes mellitus-related proteins can be further explored in targeted proteomic studies utilizing isotopically-labeled peptide internal standards for absolute quantitation, which enable the determination of laboratory-defined sensitivity and specificity with blinded samples. Further embodiments of the disclosure can be made by combining various of these type II diabetes mellitus-related proteins, either alone or in combination with other biomarkers or housekeeping genes.

Various methods for performing the quantitative and qualitative analysis of the type II diabetes mellitus-related proteins (or corresponding nucleic acid molecules) may be utilized. In one embodiment, LC-MS/MS analyses were performed, however it is to be distinctly understood that the disclosure is not limited thereto. Any reliable manner of performing quantitative or qualitative analysis of a sample for the presence and quantity of any of the preselected type II diabetes mellitus-related biomarkers set forth herein and discussed may be utilized.

In particular examples, the following proteins were found to be upregulated: APCS Serum amyloid P-component precursor, APOB Apolipoprotein B-100 precursor, C4A Complement component 4A, C4A;C4B, C4B C4B1, C4B Complement C4-B precursor, C4BPA Uncharacterized protein C4BPA, HP Haptoglobin precursor, HP HP protein, HPR Isoform 1 of Haptoglobin-related protein precursor, HPR Isoform 2 of Haptoglobin-related protein precursor, RBP4 Plasma retinol-binding protein precursor, RBP4 Retinol binding protein 4; plasma, RBP4 Uncharacterized protein RBP4, and VTN Vitronectin precursor, and the following proteins were found to be downregulated: A2M 19 kDa protein, A2M Alpha-2-macroglobulin precursor, PON1 32 kDa protein, PON1 Serum paraoxonase/arylesterase 1, PZP Pregnancy zone protein precursor, PZP and Uncharacterized protein PZP in patients having type II diabetes, as compared to a normal (non-diabetic) subject (see Table 5).

One skilled in the art will appreciate that a full-length protein can be detected, or individual peptides thereof. For example alpha-2-macroglobulin was found to be downregualted in subjects having type II diabetes mellitus relative to subjects not having type II diabetes. Therefore, to determine whether expression of alpha-2-macroglobulin is decreased in a test sample, the full-length alpha-2-

macroglobulin protein can be detected (*e.g.*, using an antibody), or one or more of the individual alpha-2-macroglobulin peptide sequences can be detected (such as SASNMAIVDVK (SEQ ID NO: 6), NQGNTWLTAFVLK (SEQ ID NO: 195), HYDGSYSTFGER (SEQ ID NO: 208), NEDSLVFVQTDK (SEQ ID NO: 5), DTVIKPLLVEPEGLEK (SEQ ID NO: 12), ALLAYAFALAGNQDK (SEQ ID NO: 209), VGFYESDVMGR (SEQ ID NO: 210), DMYSFLEDMGLK (SEQ ID NO: 4), LVHVEEPHTETVRK (SEQ ID NO: 1), MVSGFIPLKPTVK (SEQ ID NO: 2), VGFYESDVMGR (SEQ ID NO: 210), AIGYLNTGYQR (SEQ ID NO: 211), MVSGFIPLKPTVK (SEQ ID NO: 2), TEHPFTVEEFVLPK (SEQ ID NO: 7), LPPNVVEESAR (SEQ ID NO: 205), LSFVKVDSHFR (SEQ ID NO: 206), ATVLNYLPK (SEQ ID NO: 10), TVIKPLLVEPEGLEK (SEQ ID NO: 207), LLIYAVLPTGDVIGDSAK (SEQ ID NO: 79), TGTHGLLVKQEDMK (SEQ ID NO: 80), TEVSSNHVLIYLDK (SEQ ID NO: 81), KDNSVHWERPQKPK (SEQ ID NO: 82), SSSNEEVMFLTVQVK (SEQ ID NO: 76), FEVQVTVPK (SEQ ID NO: 83), SVSGKPQYMVLVPSLLHTETTEK (SEQ ID NO: 77), IAQWQSFQLEGGLK (SEQ ID NO: 84), KDTVIKPLLVEPEGLEK (SEQ ID NO: 85), SGGRTEHPFTVEEFVLPK (SEQ ID NO: 86), LLLQQVSLPELPGEYSMK (SEQ ID NO: 87), TEHPFTVEEFVLPKFEVQVTVPK (SEQ ID NO: 88), VSVQLEASPAFLAVPVEK (SEQ ID NO: 89), VDLSFSPSQSLPASHAHLR (SEQ ID NO: 90), VTAAPQSVCALR (SEQ ID NO: 91), GVPIPNKVIFIR (SEQ ID NO: 92), KPQYMVLVPSLLHTETTEK (SEQ ID NO: 78), ALLAYAFALAGNQDKR (SEQ ID NO: 93), TAQEGDHGSHVYTK (SEQ ID NO: 94), SSSNEEVMFLTVQVK (SEQ ID NO: 76), VVSMDENFHPLNELIPLVYIQDPK (SEQ ID NO: 95), SSGSLLNNAIK (SEQ ID NO: 96), AAQVTIQSSGTFSSK (SEQ ID NO: 97), GGFSSTQDTVVALHALSK (SEQ ID NO: 98), LLLQQVSLPELPGEYSMK (SEQ ID NO: 87), TGTHGLLVKQEDMK, RKEYEMK (SEQ ID NO: 11), FQVDNNNR (SEQ ID NO: 9), NEDSLVFVQTDK (SEQ ID NO: 5), DLKPAIVK (SEQ ID NO: 3), TTVMVK (SEQ ID NO: 8), or combinations thereof), for example using mass spec methods. In addition, nucleic acid molecules encoding such peptides can also be detected.

Diagnosis of Type II Diabetes Mellitus

Provided herein are methods of diagnosing type II diabetes mellitus. Particular examples of diagnosing type II diabetes mellitus include determining whether a subject, such as an otherwise healthy subject, or a subject suspected or at risk of having type II diabetes, or who has had type II diabetes, currently has type II diabetes or is predisposed to developing type II diabetes mellitus. In some examples, a sample (such as a serum or blood plasma sample) obtained from a mammal (such as a human) is analyzed to detect the presence of particular type II diabetes-associated molecules (such as those in Tables 1-4). In some examples, detection includes quantification.

In particular examples, the methods include detecting expression (such as an increase or decrease in gene or protein expression) in any combination of at least 10, at least 15, at least 20, at least 22, or at least 40 of the proteins (or nucleic acids encoding such proteins) listed in any of Tables 1-4. In one example, the method includes detecting all of the proteins (or nucleic acids encoding such proteins) in Table 1, 2, 3, or 4. The amount of protein (or nucleic acid encoding such proteins) detected can be quantified. Detecting or determining the level of expression can involve measuring an amount of the type II diabetes mellitus-related molecules in a sample derived from the subject. In some examples, the amount of protein (or nucleic acid encoding such proteins) detected is compared to an amount of the same protein present or expected in a control sample, such as a sample from a subject not having type II diabetes mellitus. For example, a difference (such as an increase or a decrease reflecting an upregulation or downregulation, respectively) in the level of at least five or at least 22 of the type II diabetes mellitus-related molecules listed in Tables 1-4 (such as the molecules listed in Table 3 or 4) in the subject relative to the control sample indicates that the subject has type II diabetes mellitus. In some example, the amount of protein (or nucleic acid encoding such proteins) detected is compared to a reference value or range of values expected if the subject has, does not have, or is predisposed to developing type II diabetes (or combinations thereof). For example, a reduction of at least 30% (such as at least 33%, at least 35%, at least 40%, or at least 45%) of a type II diabetes mellitus-related protein or nucleic acid in a test sample relative to the amount of the same type II diabetes mellitus-related

protein or nucleic acid in a non-type II diabetes control (or an amount expected in such a sample), indicates that the particular type II diabetes mellitus-related protein or nucleic acid is downregulated in the test sample. For example, it is shown herein that alpha-2-macroglobulin is downregulated in subjects having type II diabetes mellitus relative to subjects not having type II diabetes mellitus (see Table 3). In contrast, an increase of at least 1.5 fold (such as at least 2-fold, at least 2.5 fold. at least 3-fold, at least 4-fold, or at least 10-fold) of a type II diabetes mellitus-related protein or nucleic acid in a test sample relative to the amount of the same type II diabetes mellitus-related protein or nucleic acid in a non-type II diabetes control (or an amount expected in such a sample), indicates that the particular type II diabetes mellitus-related protein or nucleic acid is upregulated in the test sample. For example, it is shown herein that apolipoprotein B-100 is upregulated in subjects having type II diabetes mellitus relative to subjects not having type II diabetes mellitus (see Table 3).

In particular examples, detection of differential expression in at least five type II diabetes mellitus-related nucleic acid molecules (or proteins) listed in Tables 1-4, such as changes in gene (or protein) expression in any combination of at least 10, at least 15, at least 20, at least 22, at least 30, at least 35, at least 40, or at least 44 type II diabetes mellitus-related molecules listed in Tables 1-4 (such as those in Table 3 or peptides thereof), indicates that the subject has type II diabetes mellitus or is predisposed to developing type II diabetes mellitus. An appropriate treatment can then be selected or initiated to treat or prevent type II diabetes mellitus. Differential expression can be represented by increased or decreased expression in type II diabetes mellitus-related molecule (for instance, a nucleic acid or a protein). For example, differential expression includes, but is not limited to, an increase or decrease in an amount of a nucleic acid molecule or protein, the stability of a nucleic acid molecule or protein, the localization of a nucleic acid molecule or protein, or the biological activity of a nucleic acid molecule or protein.

In particular examples, the number of type II diabetes mellitus-related molecules screened is at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, at least 24, at least

25, at least 26, at least 27, at least 28, at least 29, at least 31, at least 31, at least 32, at least 33, at least 34, at least 35, at least 36, at least 37, at least 38, at least 39, at least 40, at least 41, at least 42, at least 43, at least 44, or at least 45 type II diabetes mellitus -related molecules. In other examples, the methods employ screening no more than 45, no more than 44, no more than 35, no more than 30, no more than 25, no more than 22, no more than 20, no more than 15, no more than 14, no more than 13, no more than 12, no more than 11, no more than 10, no more than 9, no more than 8, no more than 7, no more than 6, or no more than 5 type II diabetes mellitus-related molecules.

In particular examples, the disclosed method of diagnosing type II diabetes mellitus is at least 70% sensitive (such as at least 80% sensitive, at least 85% sensitive, at least 90% sensitive, at least 95% sensitive, at least 98% sensitive or at least 99% sensitive) and at least 70% specific (such as at least 80% specific, at least 85% specific, at least 90% specific, at least 95% specific, at least 98% specific, or at least 99% specific) for determining whether a subject has or is predisposed to developing type II diabetes mellitus. For example, using the methods provided herein, blinded samples can be used to determine the sensitivity and specificity for particular combinations of type II diabetes mellitus-related molecules.

Clinical Specimens

Appropriate specimens for use with the current disclosure in diagnosing type II diabetes include any conventional clinical samples, for instance blood or blood-fractions (such as serum or plasma). In a specific example, the sample is a human plasma or serum sample processed for detecting type II diabetes-related biomarkers. Techniques for acquisition of such samples are well known in the art (for example see Schluger *et al. J. Exp. Med.* 176:1327-33, 1992, for the collection of serum samples). Serum or other blood fractions can be prepared in the conventional manner. For example, 5 µL to 1000 µl of serum can be used to screen for the presence of proteins, depending on the detection method used (*e.g.*, immunoaffinity methods or targeted mass spectrometry). For discovery-based proteomics methods, more material is required (*e.g.*, > 100 µl serum), since immunodepletion is typically used to remove the majority of plasma protein mass, and subsequent sample

processing, such as digestion or enrichment of specific peptides, will inherently incur sample losses.

For example, rapid DNA preparation can be performed using a commercially available kit (such as the InstaGene Matrix, BioRad, Hercules, CA; the NucliSens isolation kit, Organon Teknika, Netherlands. In one example, the DNA preparation method yields a nucleotide preparation that is accessible to, and amenable to, nucleic acid amplification. Similarly, RNA can be prepared using a commercially available kit (such as the RNeasy Mini Kit, Qiagen, Valencia, CA).

Once a sample has been obtained, the sample can be used directly, concentrated (for example by centrifugation or filtration), purified, amplified, treated with enzymes (*e.g.*, proteases, such as trypsin), or combinations thereof. In particular examples, the sample is immunodepleted to remove proteins not of interest (*e.g.,* proteins which need not be detected to practice the methods disclosed herein). In particular examples, the sample is processed such that the sample can be analyzed for the presence of type II diabetes-related molecules using mass spectrometry, such as capillary liquid chromatography-Fourier transform ion cyclotron resonance mass spectrometry of protein digests of human plasma and serum samples. In some examples, data is analyzed using the accurate Mass and Time tag approach (AMT tag).

Arrays for Detecting Expression

In particular examples, methods for detecting a change in expression in the disclosed type II diabetes mellitus proteins listed in Tables 1-4 use the arrays disclosed herein. Arrays can be used to detect the presence of molecules whose expression is upregulated or downregulated, for example using specific oligonucleotide probes or antibody probes. The arrays can be used to diagnose type II diabetes mellitus (for example to determine if a person is predisposed to developing the disease). In particular examples, the disclosed arrays can include nucleic acid molecules, such as DNA or RNA molecules, or antibodies.

Nucleic acid arrays

In one example, the array includes nucleic acid oligonucleotide probes that can hybridize to any combination of at least four of the type II diabetes mellitus genes encoding the proteins listed in Tables 1-4, such as at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, at least 24, at least 25, at least 26, at least 27, at least 28, at least 29, at least 31, at least 31, at least 32, at least 33, at least 34, at least 35, at least 36, at least 37, at least 38, at least 39, at least 40, at least 41, at least 42, at least 43, at least 44, at least 45 or even all of the proteins listed in Table 1, 2, 3, or 4. In particular examples, an array includes oligonucleotides that can recognize nucleic acid molecules encoding all 22 type II diabetes mellitus associated proteins listed in Table 3, all 9 proteins listed in Table 2, or all 45 listed in Table 4. Certain of such arrays (as well as the methods described herein) can further include type II diabetes mellitus-related molecules that are not listed in Tables 1-4, such as internal controls (*e.g.*, housekeeping genes such as one or more of β-actin, glyceraldehyde 3-phosphate dehydrogenase (GADPH), succinate dehydrogenase (SDHA), hypoxanthine phosphoribosyl transferase 1 (HRPTI), HBS1-like protein (HBS1L), a cyclophilin family member protein, and alpha haemoglobin stabilizing protein (AHSP)).

In one example, a set of oligonucleotide probes is attached to the surface of a solid support for use in detection of type II diabetes mellitus-associated sequences, such as those nucleic acid sequences (such as cDNA or mRNA) obtained from the subject. Additionally, if an internal control nucleic acid sequence is used (such as nucleic acid molecules obtained from a subject who does not have type II diabetes mellitus or is pre-diabetic) an oligonucleotide probe can be included to detect the presence of this control nucleic acid molecule.

The oligonucleotide probes bound to the array can specifically bind sequences obtained from the subject, or amplified from the subject (such as under high stringency conditions). Thus, sequences of use with the method are oligonucleotide probes that recognize type II diabetes mellitus-related sequences, such as gene sequences (or corresponding proteins) listed in Tables 1-4. Such

sequences can be determined by examining the sequences of the different species, and choosing oligonucleotide sequences that specifically anneal to a particular type II diabetes mellitus-related sequence (such as those listed in Tables 1-4), but not others. One of skill in the art can identify other type II diabetes mellitus-associated oligonucleotide molecules that can be attached to the surface of a solid support for the detection of other type II diabetes mellitus-associated nucleic acid sequences.

The methods and apparatus in accordance with the present disclosure take advantage of the fact that under appropriate conditions oligonucleotides form base-paired duplexes with nucleic acid molecules that have a complementary base sequence. The stability of the duplex is dependent on a number of factors, including the length of the oligonucleotides, the base composition, and the composition of the solution in which hybridization is effected. The effects of base composition on duplex stability can be reduced by carrying out the hybridization in particular solutions, for example in the presence of high concentrations of tertiary or quaternary amines.

The thermal stability of the duplex is also dependent on the degree of sequence similarity between the sequences. By carrying out the hybridization at temperatures close to the anticipated T_m's of the type of duplexes expected to be formed between the target sequences and the oligonucleotides bound to the array, the rate of formation of mis-matched duplexes may be substantially reduced.

The length of each oligonucleotide sequence employed in the array can be selected to optimize binding of target type II diabetes mellitus-associated nucleic acid sequences. An optimum length for use with a particular type II diabetes mellitus-associated nucleic acid sequence under specific screening conditions can be determined empirically. Thus, the length for each individual element of the set of oligonucleotide sequences including in the array can be optimized for screening. In one example, oligonucleotide probes are from about 20 to about 35 nucleotides in length or about 25 to about 40 nucleotides in length.

The oligonucleotide probe sequences forming the array can be directly linked to the support. Alternatively, the oligonucleotide probes can be attached to the support by non-type II diabetes mellitus-associated sequences such as oligonucleotides or other molecules that serve as spacers or linkers to the solid

support. The oligonucleotide probes can further include one or more detectable labels, to permit detection of hybridization signals between the probe and a target sequence.

Protein arrays

In another example, an array includes protein sequences (or a fragment of such proteins, or antibodies specific to such proteins or protein fragments), which include at least four of the type II diabetes mellitus-related proteins listed in any of Tables 1-4, such as at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, at least 24, at least 25, at least 26, at least 27, at least 28, at least 29, at least 31, at least 31, at least 32, at least 33, at least 34, at least 35, at least 36, at least 37, at least 38, at least 39, at least 40, at least 41, at least 42, at least 43, at least 44, at least 45, or even all of the proteins listed in Table 1, 2, 3, or 4. In particular examples, an array includes antibodies that can recognize all 22 type II diabetes mellitus associated proteins listed in Table 3, all 9 proteins listed in Table 2, or all 45 listed in Table 4. In one example, the array includes antibodies specific for all of the peptides in Table 4 that are specific for the proteins in Table 3. Certain of such arrays (as well as the methods described herein) can further include molecules to detect type II diabetes mellitus-related proteins that are not listed in Tables 1-4, such as internal controls (*e.g.*, housekeeping genes such as beta-actin).

The proteins or antibodies forming the array can be directly linked to a solid support. Alternatively, the proteins or antibodies can be attached to the support by spacers or linkers to the solid support. For example, the antibodies specific for the type II diabetes mellitus-related proteins described herein can be part of a lateral flow device.

Changes in expression of type II diabetes mellitus-related proteins can be detected using, for instance, a type II diabetes mellitus protein-specific binding agent, which in some instances is labeled with an agent that can be detected. In certain examples, detecting a change in protein expression includes contacting a protein sample obtained from the serum or plasma of a subject with a type II

diabetes mellitus protein-specific binding agent (which can be for example present on an array); and detecting whether the binding agent is bound by the sample and thereby measuring the levels of the type II diabetes mellitus-related protein present in the sample. A difference in the level of an type II diabetes mellitus-related protein in the sample, relative to the level of a type II diabetes mellitus-related protein found an analogous sample from a subject who does not have type II diabetes mellitus or is pre-diabetic, in particular examples indicates whether the test subject has or is predisposed to developing type II diabetes mellitus.

Array substrate

The solid support can be formed from an organic polymer. Suitable materials for the solid support include, but are not limited to: polypropylene, polyethylene, polybutylene, polyisobutylene, polybutadiene, polyisoprene, polyvinylpyrrolidine, polytetrafluroethylene, polyvinylidene difluroide, polyfluoroethylene-propylene, polyethylenevinyl alcohol, polymethylpentene, polycholorotrifluoroethylene, polysulfornes, hydroxylated biaxially oriented polypropylene, aminated biaxially oriented polypropylene, thiolated biaxially oriented polypropylene, etyleneacrylic acid, thylene methacrylic acid, and blends of copolymers thereof (see U.S. Patent No. 5,985,567).

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

In one example, the solid support surface is polypropylene. Polypropylene is chemically inert and hydrophobic. Non-specific binding is generally avoidable, and detection sensitivity is improved. Polypropylene has good chemical resistance to a variety of organic acids (such as formic acid), organic agents (such as acetone or ethanol), bases (such as sodium hydroxide), salts (such as sodium chloride),

oxidizing agents (such as peracetic acid), and mineral acids (such as hydrochloric acid). Polypropylene also provides a low fluorescence background, which minimizes background interference and increases the sensitivity of the signal of interest.

In another example, a surface activated organic polymer is used as the solid support surface. One example of a surface activated organic polymer is a polypropylene material aminated via radio frequency plasma discharge. Such materials are easily utilized for the attachment of nucleotide molecules. The amine groups on the activated organic polymers are reactive with nucleotide molecules such that the nucleotide molecules can be bound to the polymers. Other reactive groups can also be used, such as carboxylated, hydroxylated, thiolated, or active ester groups.

In one example, the array is formed on a polymer medium, which is a thread, membrane or film. An example of an organic polymer medium is a polypropylene sheet having a thickness on the order of about 1 mil. (0.001 inch) to about 20 mil., although the thickness of the film is not critical and can be varied over a fairly broad range. Particularly disclosed for preparation of arrays are biaxially oriented polypropylene (BOPP) films; in addition to their durability, BOPP films exhibit a low background fluorescence.

Substrates suitable for lateral flow devises can also be used.

Array formats

A wide variety of array formats can be employed in accordance with the present disclosure. One example includes a linear array of oligonucleotide or antibody bands, generally referred to in the art as a dipstick. Another suitable format includes a two-dimensional pattern of discrete cells (such as 4096 squares in a 64 by 64 array). As is appreciated by those skilled in the art, other array formats including, but not limited to slot (rectangular) and circular arrays are equally suitable for use (see U.S. Patent No. 5,981,185).

The array formats of the present disclosure can be included in a variety of different types of formats. A "format" includes any format to which the solid support can be affixed, such as microtiter plates, test tubes, inorganic sheets,

dipsticks, and the like. For example, when the solid support is a polypropylene thread, one or more polypropylene threads can be affixed to a plastic dipstick-type device; polypropylene membranes can be affixed to glass slides. The particular format is, in and of itself, unimportant. All that is necessary is that the solid support can be affixed thereto without affecting the functional behavior of the solid support or any biopolymer absorbed thereon, and that the format (such as the dipstick or slide) is stable to any materials into which the device is introduced (such as clinical samples and hybridization solutions).

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

A suitable array can be produced using automated means to synthesize oligonucleotides in the cells of the array by laying down the precursors for the four bases in a predetermined pattern. Briefly, a multiple-channel automated chemical delivery system is employed to create probe populations in parallel rows (corresponding in number to the number of channels in the delivery system) across the substrate. Following completion of synthesis in a first direction, the substrate can then be rotated by 90° to permit synthesis to proceed within a second (2°) set of rows that are now perpendicular to the first set. This process creates a multiple-channel array whose intersection generates a plurality of discrete cells.

Oligonucleotides can be bound to the polypropylene support by either the 3' end of the oligonucleotide or by the 5' end of the oligonucleotide. In one example, the oligonucleotides are bound to the solid support by the 3' end. However, one of

skill in the art can determine whether the use of the 3' end or the 5' end of the oligonucleotide is suitable for bonding to the solid support. In general, the internal complementarity of an oligonucleotide probe in the region of the 3' end and the 5' end determines binding to the support.

In particular examples, the probes on the array include one or more labels, that permit detection of probe:target sequence complexes.

Detection of Nucleic Acid and Protein Molecules

The samples obtained from the subject (for example a serum or plasma sample) can contain altered levels of one or more nucleic acids or proteins associated with type II diabetes mellitus, such as those listed in Tables 1-4. Changes in expression can be detected to determine if a subject is predisposed to developing type II diabetes, or has type II diabetes. The present disclosure is not limited to particular methods of detecting proteins or nucleic acid molecules. Any method of detecting a nucleic acid molecule or protein can be used, such as physical or functional assays. For example, the level of gene expressing can be quantified utilizing methods well known in the art and those disclosed herein, such as Northern-Blots, RNase protection assays, nucleic acid or antibody probe arrays, quantitative PCR (such as TaqMan assays), dot blot assays, in-situ hybridization, or combinations thereof. In addition, proteins can be detected using antibody probe arrays, quantitative spectroscopic methods (for example mass spectrometry, such as surface-enhanced laser desorption/ionization (SELDI)-based mass spectrometry), or combinations thereof.

Methods for labeling nucleic acid molecules and proteins so that they can be detected are well known. Examples of such labels include non-radiolabels and radiolabels. Non-radiolabels include, but are not limited to enzymes, chemiluminescent compounds, fluorophores, metal complexes, haptens, colorimetric agents, dyes, or combinations thereof. Radiolabels include, but are not limited to, ^{125}I and ^{35}S. Radioactive and fluorescent labeling methods, as well as other methods known in the art, are suitable for use with the present disclosure. In one example, the primers used to amplify the subject's nucleic acids are labeled (such as with biotin, a radiolabel, or a fluorophore). In another example, the amplified nucleic

acid samples are end-labeled to form labeled amplified material. For example, amplified nucleic acid molecules can be labeled by including labeled nucleotides in the amplification reactions. In another example, nucleic acid molecules obtained from a subject are labeled, and applied to an array containing oligonucleotides. In a particular example, proteins obtained from a subject are labeled and subsequently analyzed, for example by applying them to an array.

For such procedures, a biological sample of the subject is assayed for an increase or decrease in expression of type II diabetes mellitus-related molecules, such as those listed in Tables 1-4. Suitable biological samples include blood samples that contain DNA or RNA (including mRNA) or proteins. In a particular example, the sample is a serum or plasma sample that has been immunodepleted.

The detection in the biological sample of increased or decreased expression in a plurality of type II diabetes mellitus-related nucleic acid molecules, such as those listed in Table 3 or 4, can be achieved by methods known in the art. For example, Increased or decreased expression of a type II diabetes mellitus-related molecule can be detected by measuring the cellular level of type II diabetes mellitus-related nucleic acid molecule-specific mRNA. mRNA can be measured using techniques well known in the art, including for instance Northern analysis, RT-PCR and mRNA *in situ* hybridization. Details of mRNA analysis procedures can be found, for instance, in provided examples and in Sambrook *et al.* (ed.), *Molecular Cloning: A Laboratory Manual*, 2nd ed., vol. 1-3, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

Oligonucleotides specific to type II diabetes mellitus-related sequences can be chemically synthesized using commercially available machines. These oligonucleotides can then be labeled, for example with radioactive isotopes (such as ^{32}P) or with non-radioactive labels such as biotin (Ward and Langer *et al., Proc. Natl. Acad. Sci. USA* 78:6633-57, 1981) or a fluorophore, and hybridized to individual DNA samples immobilized on membranes or other solid supports by dot-blot or transfer from gels after electrophoresis. These specific sequences are visualized, for example by methods such as autoradiography or fluorometric (Landegren *et al., Science* 242:229-37, 1989) or colorimetric reactions (Gebeyehu *et al., Nucleic Acids Res.* 15:4513-34, 1987).

Nucleic acid molecules isolated from blood samples can be amplified using routine methods to form nucleic acid amplification products. These nucleic acid amplification products can then be contacted with an oligonucleotide probe that will hybridize under stringent conditions with a type II diabetes mellitus-related nucleic acid. The nucleic acid amplification products which hybridize with the probe are then detected and quantified. The oligonucleotide probe can bind specifically to a nucleic acid molecule that encodes a protein listed in any of Tables 1-4.

The nucleic acid molecules obtained from the subject that are associated with type II diabetes mellitus can be applied to a type II diabetes mellitus detection array under suitable hybridization conditions to form a hybridization complex. In particular examples, the nucleic acid molecules include a label. In one example, a pre-treatment solution of organic compounds, solutions that include organic compounds, or hot water, can be applied before hybridization (see U.S. Patent No. 5,985,567, herein incorporated by reference).

Hybridization conditions for a given combination of array and target material can be optimized routinely in an empirical manner close to the T_m of the expected duplexes, thereby maximizing the discriminating power of the method. Identification of the location in the array, such as a cell, in which binding occurs, permits a rapid and accurate identification of sequences associated with type II diabetes mellitus present in the amplified material (see below).

The hybridization conditions are selected to permit discrimination between matched and mismatched oligonucleotides. Hybridization conditions can be chosen to correspond to those known to be suitable in standard procedures for hybridization to filters and then optimized for use with the arrays of the disclosure. For example, conditions suitable for hybridization of one type of target would be adjusted for the use of other targets for the array. In particular, temperature is controlled to substantially eliminate formation of duplexes between sequences other than exactly complementary type II diabetes mellitus-associated sequences. A variety of known hybridization solvents can be employed, the choice being dependent on considerations known to one of skill in the art (see U.S. Patent 5,981,185).

Once the nucleic acid molecules associated with type II diabetes mellitus from the subject have been hybridized with the oligonucleotides present in the type

II diabetes mellitus detection array, the presence of the hybridization complex can be analyzed, for example by detecting the complexes.

Detecting a hybridized complex in an array of oligonucleotide probes has been previously described (see U.S. Patent No. 5,985,567, herein incorporated by reference). In one example, detection includes detecting one or more labels present on the oligonucleotides, the sequences obtained from the subject, or both. In particular examples, developing includes applying a buffer. In one example, the buffer is sodium saline citrate, sodium saline phosphate, tetramethylammonium chloride, sodium saline citrate in ethylenediaminetetra-acetic, sodium saline citrate in sodium dodecyl sulfate, sodium saline phosphate in ethylenediaminetetra-acetic, sodium saline phosphate in sodium dodecyl sulfate, tetramethylammonium chloride in ethylenediaminetetra-acetic, tetramethylammonium chloride in sodium dodecyl sulfate, or combinations thereof. However, other suitable buffer solutions can also be used.

Detection can further include treating the hybridized complex with a conjugating solution to effect conjugation or coupling of the hybridized complex with the detection label, and treating the conjugated, hybridized complex with a detection reagent. In one example, the conjugating solution includes streptavidin alkaline phosphatase, avidin alkaline phosphatase, or horseradish peroxidase. Specific, non-limiting examples of conjugating solutions include streptavidin alkaline phosphatase, avidin alkaline phosphatase, or horseradish peroxidase. The conjugated, hybridized complex can be treated with a detection reagent. In one example, the detection reagent includes enzyme-labeled fluorescence reagents or calorimetric reagents. In one specific non-limiting example, the detection reagent is enzyme-labeled fluorescence reagent (ELF) from Molecular Probes, Inc. (Eugene, OR). The hybridized complex can then be placed on a detection device, such as an ultraviolet (UV) transilluminator (manufactured by UVP, Inc. of Upland, CA). The signal is developed and the increased signal intensity can be recorded with a recording device, such as a charge coupled device (CCD) camera (manufactured by Photometrics, Inc. of Tucson, AZ). In particular examples, these steps are not performed when fluorophores or radiolabels are used.

In particular examples, the method further includes quantification, for instance by determining the amount of hybridization.

Protein expression can be detected using any method known in the art, such as by detecting full-length proteins or portions thereof using antibodies, or using other methods such as mass spectrometry. The determination of increased or decreased type II diabetes mellitus-related protein levels, in comparison to such expression in a control (such as a subject who does not have type II diabetes mellitus or is prediabetic), is an alternative or supplemental approach to the direct determination of the expression level of type II diabetes mellitus-related nucleic acid sequences by the methods outlined above. The availability of antibodies specific to type II diabetes mellitus-related protein(s) will facilitate the detection and quantification of such protein(s) by one of a number of immunoassay methods that are well known in the art, such as those presented in Harlow and Lane (*Antibodies, A Laboratory Manual*, CSHL, New York, 1988). In addition, if such antibodies are not available, methods of constructing antibodies are routine in the art.

Any standard immunoassay format (such as ELISA, Western blot, or RIA assay) can be used to measure type II diabetes mellitus-related protein levels. A comparison to control (*e.g.*, subject who does not have type II diabetes mellitus) and an increase or decrease in type II diabetes mellitus-related peptide levels (such as an increase or decrease in any combination of at least 10 proteins or at least 20 proteins listed in Table 3 or 4) is indicative of type II diabetes mellitus. Immunohistochemical techniques can also be utilized for protein detection and quantification. For example, a tissue sample can be obtained from a subject, and a section stained for the presence of an type II diabetes mellitus-related protein using the appropriate protein specific binding agents and any standard detection system (such as one that includes a secondary antibody conjugated to horseradish peroxidase). General guidance regarding such techniques can be found in Bancroft and Stevens (*Theory and Practice of Histological Techniques*, Churchill Livingstone, 1982) and Ausubel *et al.* (*Current Protocols in Molecular Biology*, John Wiley & Sons, New York, 1998).

For the purposes of quantifying type II diabetes mellitus-related proteins, a biological sample of the subject that includes cellular proteins can be used.

Quantification of a type II diabetes mellitus-related protein can be achieved by immunoassay and the amount compared to levels of the protein found in cells from a subject who does not have type II diabetes mellitus or is prediabetic. A significant increase or decrease in the amount of five or more type II diabetes mellitus-related proteins in the serum or plasma (or other sample) of a subject compared to the amount of the same type II diabetes mellitus-related protein found in a normal sample is usually at least at a 1.5-fold, at least 2-fold, at least 3-fold, at least 4-fold or greater increase, or at least a 20%, at least 30%, at least 35%, or at least 40% decrease. Substantial overexpression or underexpression of five or more type II diabetes mellitus-related protein(s) can be indicative of the presence of type II diabetes mellitus or pre-diabetes.

An alternative method of detecting protein expression is to use spectroscopic method, for example quantitative spectroscopic approaches, such as SELDI. In one example, surface-enhanced laser desorption-ionization time-of-flight (SELDI-TOF) mass spectrometry is used to detect changes in protein expression, for example by using the ProteinChip™ (Ciphergen Biosystems, Palo Alto, CA). Such methods are well known in the art (for example see U.S. Pat. No. 5,719,060; U.S. Pat. No. 6,897,072; and U.S. Pat. No. 6,881,586). SELDI is a solid phase method for desorption in which the analyte is presented to the energy stream on a surface that enhances analyte capture or desorption.

Briefly, one version of SELDI uses a chromatographic surface with a chemistry that selectively captures analytes of interest, such as type II diabetes mellitus-related proteins. Chromatographic surfaces can be composed of hydrophobic, hydrophilic, ion exchange, immobilized metal, or other chemistries. For example, the surface chemistry can include binding functionalities based on oxygen-dependent, carbon-dependent, sulfur-dependent, and/or nitrogen-dependent means of covalent or noncovalent immobilization of analytes. The activated surfaces are used to covalently immobilize specific "bait" molecules such as antibodies, receptors, or oligonucleotides often used for biomolecular interaction studies such as protein-protein and protein-DNA interactions.

The surface chemistry allows the bound analytes to be retained and unbound materials to be washed away. Subsequently, analytes bound to the surface (such as

type II diabetes mellitus-related proteins) can be desorbed and analyzed by any of several means, for example using mass spectrometry. When the analyte is ionized in the process of desorption, such as in laser desorption/ionization mass spectrometry, the detector can be an ion detector. Mass spectrometers generally include means for determining the time-of-flight of desorbed ions. This information is converted to mass. However, one need not determine the mass of desorbed ions to resolve and detect them: the fact that ionized analytes strike the detector at different times provides detection and resolution of them. Alternatively, the analyte can be detectably labeled (for example with a fluorophore or radioactive isotope). In these cases, the detector can be a fluorescence or radioactivity detector. A plurality of detection means can be implemented in series to fully interrogate the analyte components and function associated with retained molecules at each location in the array.

Therefore, in a particular example, the chromatographic surface includes antibodies that recognize type II diabetes mellitus-related proteins. In one example, antibodies are immobilized onto the surface using a bacterial Fc binding support. The chromatographic surface is incubated with a sample from the subject, such as a sample that includes serum or plasma proteins. The antigens present in the sample can recognize the antibodies on the chromatographic surface. The unbound proteins and mass spectrometric interfering compounds are washed away and the proteins that are retained on the chromatographic surface are analyzed and detected by SELDI-TOF. The MS profile from the sample can be then compared using differential protein expression mapping, whereby relative expression levels of proteins at specific molecular weights are compared by a variety of statistical techniques and bioinformatic software systems.

Kits

The present disclosure provides for kits that can be used to diagnose type II diabetes mellitus, for example to determine if a subject has type II diabetes mellitus or has an increased predisposition to developing type II diabetes mellitus. Such kits allow one to determine if a subject has a differential expression in type II diabetes mellitus-related genes, such as any combination of four or more of those listed in

Tables 1-4, such as any combination of 10 or more of those listed in Tables 3 or 4, or any combination of 20 or more of those listed in Table 3 or 4.

The disclosed kits include binding molecules, such as oligonucleotide probes that selectively hybridize to, or antibodies that specifically bind to, type II diabetes mellitus-related molecules that are the target of the kit. In particular examples, the oligonucleotide probes or antibodies are attached to an array, such as a biochip or dipstick. Such an array can include other oligonucleotides or antibodies, for example to serve as negative or positive controls. In one example, the kit includes oligonucleotide probes or primers (or antibodies) that recognize any combination of at least four of the molecules in Tables 1-4, such as at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, at least 24, at least 25, at least 26, at least 27, at least 28, at least 29, at least 31, at least 31, at least 32, at least 33, at least 34, at least 35, at least 36, at least 37, at least 38, at least 39, at least 40, at least 41, at least 42, at least 43, at least 44, or at least 45 of the type II diabetes mellitus-related molecules listed in any of Tables 1-4.

The kit can further include one or more of a buffer solution, a conjugating solution for developing the signal of interest, or a detection reagent for detecting the signal of interest, each in separate packaging, such as a container. Kits can include instructions, for instance instructions that provide calibration curves or charts to compare with the determined (such as experimentally measured) values. For example, instructions can permit the tester to determine whether type II diabetes mellitus-related expression levels are elevated, reduced, or unchanged in comparison to a control sample.

Example 1

This example describes methods used to identify 31 type II diabetes mellitus markers.

Aliquots of plasma (120 μL) from individuals with normal glucose tolerance (NGT; control), impaired glucose tolerance (IGT; pre-diabetic) and type 2 diabetes were subjected to immunodepletion using an Agilent MARS-6 column, in order to

remove the six most abundant plasma proteins. One skilled in the art will appreciate that other immunodepletion methods can be used, for example by using the SuperMix column (GenWay Biotech), which removes the top 50-60 most abundant proteins from plasma. Subsequent LC-MS and data analyses would proceed as described below.

The samples were then digested with trypsin, and aliquots of each individual sample were pooled to create a common reference sample. The reference sample was subsequently labeled with ^{18}O, and each individual sample was labeled with ^{16}O. Aliquots of the reference sample were then mixed with each individual sample (1:1, w/w), and the samples were analyzed by capillary liquid chromatography coupled with a 9.4T Fourier transform ion cyclotron resonance mass spectrometer. Raw data was processed using the PRISM data analysis system, and isotopic pair ratio information was placed into a data matrix. The pair ratios were subjected to Loess normalization, and protein abundances were analyzed by ANOVA.

The 31 proteins identified are shown in Table 1. A subset of those is shown in Table 2, and in some examples, the proteins in Table 2 are analyzed. The individual peptides identified are also listed. One skilled in the art will appreciate that when detecting type II diabetes mellitus-related proteins, either the full-length protein listed in Tables 1 and 2 (*e.g.*, alpha-2-macroglobulin) or the peptides shown (*e.g.*, one or more the 12 peptides listed for alpha-2-macroglobulin), can be detected using routine methods.

Table 1: 31 proteins associated with type II diabetes mellitus

(1) 163 kDa protein; peptide sequences: LVHVEEPHTETVRK (SEQ ID NO: 1), MVSGFIPLKPTVK (SEQ ID NO: 2), DLKPAIVK (SEQ ID NO: 3), DMYSFLEDMGLK (SEQ ID NO: 4), NEDSLVFVQTDK (SEQ ID NO: 5), SASNMAIVDVK (SEQ ID NO: 6), TEHPFTVEEFVLPK (SEQ ID NO: 7), TTVMVK (SEQ ID NO: 8), FQVDNNNR (SEQ ID NO: 9), ATVLNYLPK (SEQ ID NO: 10), RKEYEMK (SEQ ID NO: 11), DTVIKPLLVEPEGLEK (SEQ ID NO: 12)

(2) 45 kDa protein; peptide sequences: DKVNSFFSTFK (SEQ ID NO: 13), LTPYADEFKVK (SEQ ID NO: 14)

(3) afamin; peptide sequences: RHPDLSIPELLR (SEQ ID NO: 15), HFQNLGK (SEQ ID NO: 16)

(4) alpha-2-macroglobulin; peptide sequences: SASNMAIVDVK (SEQ ID NO: 6), RKEYEMK (SEQ ID NO: 11), ATVLNYLPK (SEQ ID NO: 10), FQVDNNNR (SEQ ID NO: 9), DTVIKPLLVEPEGLEK (SEQ ID NO: 12), TEHPFTVEEFVLPK (SEQ ID NO: 7), NEDSLVFVQTDK (SEQ ID NO: 5), DMYSFLEDMGLK (SEQ ID NO: 4), DLKPAIVK (SEQ ID NO: 3), MVSGFIPLKPTVK (SEQ ID NO: 2), LVHVEEPHTETVRK (SEQ ID NO: 1), TTVMVK (SEQ ID NO: 8)

(5) apolipoprotein A-I; peptide sequences: AHVDALR (SEQ ID NO: 17), SGRDYVSQFEGSALGK (SEQ ID NO: 18), PYLDDFQKK (SEQ ID NO: 19), YVSQFEGSALGK (SEQ ID NO: 20), LHELQEK (SEQ ID NO: 21), VSFLSALEEYTK (SEQ ID NO: 22), QKLHELQEK (SEQ ID NO: 23)

(6) apolipoprotein A-IV; peptide sequences: DKVNSFFSTFK (SEQ ID NO: 13), LTPYADEFKVK (SEQ ID NO: 14)

(7) apolipoprotein B-100; peptide sequences: MLETVR (SEQ ID NO: 24), AVSMPSFSILGSDVR (SEQ ID NO: 25), TEVIPPLIENR (SEQ ID NO: 26), SVGFHLPSR (SEQ ID NO: 27)

(8) complement C4B1; peptide sequences: GQIVFMNREPK (SEQ ID NO: 28), GPEVQLVAHSPWLK (SEQ ID NO: 29), SHALQLNNR (SEQ ID NO: 30), DFALLSLQVPLKDAK (SEQ ID NO: 31), YIYGKPVQGVAYVR (SEQ ID NO: 32)

(9) ceruloplasmin; peptide sequences: AETGDKVYVHLK (SEQ ID NO: 33), GPEEEHLGILGPVIWAEVGDTIR (SEQ ID NO: 34)

(10) clusterin isoform 1; peptide sequences: TLLSNLEEAK (SEQ ID NO: 35), FMETVAEK (SEQ ID NO: 36)

(11) clusterin; peptide sequences: TLLSNLEEAK (SEQ ID NO: 35), FMETVAEK (SEQ ID NO: 36)

(12) complement C3; peptide sequences: FYYIYNEK (SEQ ID NO: 37), LMNIFLK (SEQ ID NO: 38), IPIEDGSGEVVLSR (SEQ ID NO: 39), IWDVVEK (SEQ ID NO: 40), TIYTPGSTVLYR (SEQ ID NO: 41), KGYTQQLAFR (SEQ ID NO: 42), RIPIEDGSGEVVLSR (SEQ ID NO: 43), VQLSNDFDEYIMAIEQTIK (SEQ ID NO: 44), RQGALELIKK (SEQ ID NO: 45), AAVYHHFISDGVRK (SEQ ID NO: 46), YYTYLIMNK (SEQ ID NO: 47)

(13) complement C4; peptide sequences: GPEVQLVAHSPWLK (SEQ ID NO: 29), SHALQLNNR (SEQ ID NO: 30), DFALLSLQVPLKDAK (SEQ ID NO: 31), YIYGKPVQGVAYVR (SEQ ID NO: 32), GQIVFMNREPK (SEQ ID NO: 28)

(14) complement C4A; peptide sequences: GPEVQLVAHSPWLK (SEQ ID NO: 29), SHALQLNNR (SEQ ID NO: 30), DFALLSLQVPLKDAK (SEQ ID NO: 31), YIYGKPVQGVAYVR (SEQ ID NO: 32), GQIVFMNREPK (SEQ ID NO: 28)

(15) complement C4B; peptide sequences: YIYGKPVQGVAYVR (SEQ ID NO: 32), GQIVFMNREPK (SEQ ID NO: 28), SHALQLNNR (SEQ ID NO: 30), GPEVQLVAHSPWLK (SEQ ID NO: 29), DFALLSLQVPLKDAK (SEQ ID NO: 31)

(16) FGA protein; peptide sequences: DSHSLTTNIMEILR (SEQ ID NO: 48), GLIDEVNQDFTNR (SEQ ID NO: 49), LKNSLFEYQK (SEQ ID NO: 50)

(17) gelsolin isoform b; peptide sequences: HVVPNEVVVQR (SEQ ID NO: 51), TASDFITK (SEQ ID NO: 52)

(18) gelsolin; peptide sequences: HVVPNEVVVQR (SEQ ID NO: 51), TASDFITK (SEQ ID NO: 52)

(19) hemopexin; peptide sequences: LWWLDLK (SEQ ID NO: 53), RLWWLDLK (SEQ ID NO: 54), GDKVWVYPPEKK (SEQ ID NO: 55)

(20) histidine-rich glycoprotein; peptide sequences: DSPVLIDFFEDTER (SEQ ID NO: 56), ADLFYDVEALDLESPK (SEQ ID NO: 57)

(21) Hypothetical protein DKFZp779N0926; peptide sequences: RLDGSVDFK (SEQ ID NO: 58), TSTADYAMFK (SEQ ID NO: 59)

(22) Inter-alpha-trypsin inhibitor heavy chain H2 precursor; peptide sequences: FYNQVSTPLLR (SEQ ID NO: 60), SLAPTAAAK (SEQ ID NO: 61), TILDDLR (SEQ ID NO: 62)

(23) pregnancy zone protein; peptide sequences: MVSGFIPLKPTVK (SEQ ID NO: 2), DLKPAIVK (SEQ ID NO: 3), ATVLNYLPK (SEQ ID NO: 10)

(24) Similar to fibrinogen, A alpha polypeptide; peptide sequences: LKNSLFEYQK (SEQ ID NO: 50), DSHSLTTNIMEILR (SEQ ID NO: 48), GLIDEVNQDFTNR (SEQ ID NO: 49)

(25) Splice Isoform 1 of Inter-alpha-trypsin inhibitor heavy chain H4; peptide sequences: NVVFVIDK (SEQ ID NO: 63), LGVYELLLK (SEQ ID NO: 64), ETLFSVMPGLK (SEQ ID NO: 65)

(26) Splice Isoform 2 of Inter-alpha-trypsin inhibitor heavy chain H4; peptide sequences: LGVYELLLK (SEQ ID NO: 64), ETLFSVMPGLK (SEQ ID NO: 65), NVVFVIDK (SEQ ID NO: 63)

(27) Splice Isoform Alpha of Fibrinogen alpha/alpha-E chain; peptide sequences: MELERPGGNEITR (SEQ ID NO: 66), LKNSLFEYQK (SEQ ID NO: 50), DSHSLTTNIMEILR (SEQ ID NO: 48), TVIGPDGHKEVTK (SEQ ID NO: 67), GLIDEVNQDFTNR (SEQ ID NO: 49)

(28) Splice Isoform Alpha-E of Fibrinogen alpha/alpha-E chain; peptide sequences: LKNSLFEYQK (SEQ ID NO: 50), DSHSLTTNIMEILR (SEQ ID NO: 48), MELERPGGNEITR (SEQ ID NO: 66), GLIDEVNQDFTNR (SEQ ID NO: 49), TVIGPDGHKEVTK (SEQ ID NO: 67)

(29) Splice Isoform Gamma-A of Fibrinogen gamma chain; peptide sequences: RLDGSVDFK (SEQ ID NO: 58), TSTADYAMFK (SEQ ID NO: 59)

(30) Splice Isoform Gamma-B of Fibrinogen gamma chain; peptide sequences: RLDGSVDFK (SEQ ID NO: 58), TSTADYAMFK (SEQ ID NO: 59)

(31) Vitamin D-binding protein; peptide sequences: ELPEHTVK (SEQ ID NO: 68), KFPSGTFEQVSQLVK (SEQ ID NO: 69), THLPEVFLSK (SEQ ID NO: 70), HLSLLTTLSNR (SEQ ID NO: 71)

Table 2: 9 proteins associated with type II diabetes mellitus

(1) 163 kDa protein; peptide sequences: LVHVEEPHTETVRK (SEQ ID NO: 1), MVSGFIPLKPTVK (SEQ ID NO: 2), DLKPAIVK (SEQ ID NO: 3), DMYSFLEDMGLK (SEQ ID NO: 4), NEDSLVFVQTDK (SEQ ID NO: 5), SASNMAIVDVK (SEQ ID NO: 6), TEHPFTVEEFVLPK (SEQ ID NO: 7), TTVMVK (SEQ ID NO: 8), FQVDNNNR (SEQ ID NO: 9), ATVLNYLPK (SEQ ID NO: 10), RKEYEMK (SEQ ID NO: 11), DTVIKPLLVEPEGLEK (SEQ ID NO: 12)

(2) 45 kDa protein; peptide sequences: DKVNSFFSTFK (SEQ ID NO: 13), LTPYADEFKVK (SEQ ID NO: 14)

(3) alpha-2-macroglobulin; peptide sequences: SASNMAIVDVK (SEQ ID NO: 6), RKEYEMK (SEQ ID NO: 11), ATVLNYLPK (SEQ ID NO: 10), FQVDNNNR (SEQ ID NO: 9), DTVIKPLLVEPEGLEK (SEQ ID NO: 12), TEHPFTVEEFVLPK (SEQ ID NO: 7), NEDSLVFVQTDK (SEQ ID NO: 5), DMYSFLEDMGLK (SEQ ID NO: 4), DLKPAIVK (SEQ ID NO: 3), MVSGFIPLKPTVK (SEQ ID NO: 2), LVHVEEPHTETVRK (SEQ ID NO: 1), TTVMVK (SEQ ID NO: 8)

(4) apolipoprotein B-100; peptide sequences: MLETVR (SEQ ID NO: 24), AVSMPSFSILGSDVR (SEQ ID NO: 25), TEVIPPLIENR (SEQ ID NO: 26), SVGFHLPSR (SEQ ID NO: 27)

(5) complement C4B1; peptide sequences: GQIVFMNREPK (SEQ ID NO: 28), GPEVQLVAHSPWLK (SEQ ID NO: 29), SHALQLNNR (SEQ ID NO: 30), DFALLSLQVPLKDAK (SEQ ID NO: 31), YIYGKPVQGVAYVR (SEQ ID NO: 32)

(6) ceruloplasmin; peptide sequences: AETGDKVYVHLK (SEQ ID NO: 33), GPEEEHLGILGPVIWAEVGDTIR (SEQ ID NO: 34)

(7) complement C3; peptide sequences: FYYIYNEK (SEQ ID NO: 37), LMNIFLK (SEQ ID NO: 38), IPIEDGSGEVVLSR (SEQ ID NO: 39), IWDVVEK (SEQ ID NO: 40), TIYTPGSTVLYR (SEQ ID NO: 41), KGYTQQLAFR (SEQ ID NO: 42), RIPIEDGSGEVVLSR (SEQ ID NO: 43), VQLSNDFDEYIMAIEQTIK (SEQ ID NO: 44), RQGALELIKK (SEQ ID NO: 45), AAVYHHFISDGVRK (SEQ ID NO: 46), YYTYLIMNK (SEQ ID NO: 47)

(8) complement C4; peptide sequences: GPEVQLVAHSPWLK (SEQ ID NO: 29), SHALQLNNR (SEQ ID NO: 30), DFALLSLQVPLKDAK (SEQ ID NO: 31), YIYGKPVQGVAYVR (SEQ ID NO: 32), GQIVFMNREPK (SEQ ID NO: 28)

(9) complement C4B; peptide sequences: YIYGKPVQGVAYVR (SEQ ID NO: 32), GQIVFMNREPK (SEQ ID NO: 28), SHALQLNNR (SEQ ID NO: 30), GPEVQLVAHSPWLK (SEQ ID NO: 29), DFALLSLQVPLKDAK (SEQ ID NO: 31)

EXAMPLE 2

This example describes methods used to identify 31 type II diabetes mellitus markers.

Aliquots of plasma (120 μL) from individuals with normal glucose tolerance (NGT; control), impaired glucose tolerance (IGT; pre-diabetic) and type 2 diabetes were subjected to immunodepletion using an Agilent MARS-6 column, in order to remove the 6 most abundant plasma proteins. One skilled in the art will appreciate that other immunodepletion methods can be used, for example by using the SuperMix column (GenWay Biotech), which removes the top 50-60 most abundant

proteins from plasma. Subsequent LC-MS and data analyses would proceed as described below.

The samples were then digested with trypsin, and aliquots of each individual sample were pooled to create a common reference sample. The reference sample was subsequently labeled with ^{18}O, and each individual sample was labeled with ^{16}O. Aliquots of the reference sample were then mixed with each individual sample (1:1, w/w), and the samples were analyzed by capillary liquid chromatography coupled with a 9.4T Fourier transform ion cyclotron resonance mass spectrometer. Raw data was processed using the PRISM data analysis system, and isotopic pair ratio information was placed into a data matrix. The pair ratios were subjected to linear regression normalization, and protein abundances were analyzed by ANOVA.

This proteins identified are shown in Table 3. The individual peptides identified are listed in Table 4 (Table 4 also includes the peptides detected in Example 1). One skilled in the art will appreciate that when detecting type II diabetes mellitus-related proteins, either the full-length protein listed in Table 3 (*e.g.*, apolipoprotein B-100) or the peptides shown in Table 4 (*e.g.*, one or more the 55 peptides listed for apolipoprotein B-100), can be detected using routine methods. In Table 3, the IPI number is the International Protein Index number for each particular protein listed. The IPI numbers can be referenced on the EMBL-EBI website (for example at www.ebi.ac.uk/IPI/IPIhelp.html).

In Table 3, abundance refers to protein abundance and is in arbitrary units. The number of non-missing refers to in how many individuals (control (C), pre-diabetic (PD), or type 2 diabetic (D)) the marker was observed out of a possible 10 control, 10 pre-diabetic, and 9 type 2 diabetic individuals. A negative abundance does not necessarily mean down regulation. Since this is on the log2 scale, a negative value indicates that the original abundance was less than zero. Whether a protein was up or down regulated is indicated in the fold changes column.

Table 3: 22 proteins associated with type II diabetes mellitus

IPI	p value	q value	Protein Description
IPI:IPI00789547.1	0.004336644	0.018635889	A2M 19 kDa protein
IPI:IPI00878729.1	0.000100305	0.002885668	A2M 19 kDa protein
IPI:IPI00478003.1	0.010079138	0.024244412	A2M Alpha-2-macroglobulin precursor
IPI:IPI00022391.1	0.002451639	0.017632734	APCS Serum amyloid P-component precursor
IPI:IPI00022229.1	0.000382972	0.005508841	APOB Apolipoprotein B-100 precursor
IPI:IPI00643525.1	0.003339336	0.018635889	C4A Complement component 4A
IPI:IPI00032258.4	0.004534454	0.018635889	C4A;C4B
IPI:IPI00418163.3	0.006757072	0.022909885	C4B C4B1
IPI:IPI00654875.1	0.004534454	0.018635889	C4B Complement C4-B precursor
IPI:IPI00872510.1	0.000256426	0.004918055	C4BPA Uncharacterized protein C4BPA
IPI:IPI00641737.1	0.022091814	0.035308729	HP Haptoglobin precursor
IPI:IPI00478493.3	0.014137715	0.024633302	HP HP protein
IPI:IPI00477597.1	0.006784291	0.022909885	HPR Isoform 1 of Haptoglobin-related protein precursor
IPI:IPI00607707.1	0.006784291	0.022909885	HPR Isoform 2 of Haptoglobin-related protein precursor
IPI:IPI00798167.1	0.008361592	0.022909885	PON1 32 kDa protein
IPI:IPI00218732.3	0.008361592	0.022909885	PON1 Serum paraoxonase/arylesterase 1
IPI:IPI00025426.2	0.042916747	0.052539041	PZP Pregnancy zone protein precursor
IPI:IPI00748437.2	0.029265308	0.040091935	PZP Uncharacterized protein PZP
IPI:IPI00022420.3	0.01179822	0.024244412	RBP4 Plasma retinol-binding protein precursor
IPI:IPI00480192.1	0.01179822	0.024244412	RBP4 Retinol binding protein 4; plasma
IPI:IPI00844536.2	0.01179822	0.024244412	RBP4 Uncharacterized protein RBP4
IPI:IPI00298971.1	0.000492196	0.005663973	VTN Vitronectin precursor

IPI	Control abundance (log 2)	Control abundance	Control (#non-missing)	Pre-Diabetic abundance (log 2)
IPI:IPI00789547.1	0.20824898	1.155285146	10	-0.335318801
IPI:IPI00878729.1	1.03602219	2.050566015	10	0.455304577
IPI:IPI00478003.1	2.485177849	5.599033645	10	2.068434497
	-			
IPI:IPI00022391.1	0.930461873	0.524690337	10	-0.887633417
IPI:IPI00022229.1	0.543333761	1.457336224	10	1.107507174
IPI:IPI00643525.1	0.272086051	1.207552615	10	0.928486083
IPI:IPI00032258.4	0.265645666	1.202173952	10	0.946084173
IPI:IPI00418163.3	0.269797038	1.205638203	10	0.946648284
IPI:IPI00654875.1	0.265645666	1.202173952	10	0.946084173
IPI:IPI00872510.1	1.548911716	2.925963386	10	1.715333326
	-			
IPI:IPI00641737.1	1.529573021	0.346379866	10	-0.51483747
IPI:IPI00478493.3	-1.72901697	0.301657432	10	-0.401835904
	-			
IPI:IPI00477597.1	1.660908682	0.316239902	10	-0.681677014
	-			
IPI:IPI00607707.1	1.660908682	0.316239902	10	-0.681677014
IPI:IPI00798167.1	5.76E-02	1.04072629	10	1.06E-02
IPI:IPI00218732.3	5.76E-02	1.04072629	10	1.06E-02
IPI:IPI00025426.2	0.228928422	1.171964137	10	-0.201879344
IPI:IPI00748437.2	0.284621764	1.218090871	10	-0.229710368
IPI:IPI00022420.3	1.000643365	2.000892092	10	1.694869809
IPI:IPI00480192.1	1.000643365	2.000892092	10	1.694869809
IPI:IPI00844536.2	1.000643365	2.000892092	10	1.694869809
IPI:IPI00298971.1	0.14065272	1.102403765	10	0.868307887

IPI	Pre-Diabetic abundance	Pre-Diabetic (#non-missing)	Diabetic abundance (log 2)	Diabetic abundance	Diabetic(#non-missing)
IPI:IPI00789547.1	0.79260897	10	-0.541226968	0.68718623	9
IPI:IPI00878729.1	1.371072231	10	9.77E-02	1.070073396	9
IPI:IPI00478003.1	4.194312914	10	1.875230313	3.668601784	9
IPI:IPI00022391.1	0.540500022	10	-0.283711488	0.821474964	9
IPI:IPI00022229.1	2.154730106	10	1.206731871	2.308141821	9
IPI:IPI00643525.1	1.903277711	10	1.095248036	2.136498102	9
IPI:IPI00032258.4	1.926636192	10	1.084065693	2.120002099	9
IPI:IPI00418163.3	1.927389678	10	1.081077098	2.115614984	9
IPI:IPI00654875.1	1.926636192	10	1.084065693	2.120002099	9
IPI:IPI00872510.1	3.283725033	10	2.147178821	4.42960735	9
IPI:IPI00641737.1	0.699871774	10	9.82E-03	1.006828825	9
IPI:IPI00478493.3	0.756894483	10	5.77E-03	1.004005469	9
IPI:IPI00477597.1	0.623440155	10	3.68E-02	1.025814147	9
IPI:IPI00607707.1	0.623440155	10	3.68E-02	1.025814147	9
IPI:IPI00798167.1	1.007388573	10	-0.534050354	0.690613119	9
IPI:IPI00218732.3	1.007388573	10	-0.534050354	0.690613119	9
IPI:IPI00025426.2	0.869417268	10	-0.393467588	0.761297588	9
IPI:IPI00748437.2	0.852806082	10	-0.409525909	0.752870738	9
IPI:IPI00022420.3	3.237476698	9	2.267218423	4.813940878	9
IPI:IPI00480192.1	3.237476698	9	2.267218423	4.813940878	9
IPI:IPI00844536.2	3.237476698	9	2.267218423	4.813940878	9
IPI:IPI00298971.1	1.825520521	10	0.852296318	1.805372226	9

IPI	FoldChanges_C_vs_PD (log 2)	FoldChanges_C_vs_PD_AbsoluteValue (log2)	FoldChanges_C_vs_PD_AbsoluteValue	FoldChanges_PD_vs_D (log 2)	FoldChanges_PD_vs_D_AbsoluteValue (log 2)
IPI:IPI00789547.1	0.543567781	0.5435678	1.4575726	0.2059082	0.2059082
IPI:IPI00878729.1	0.580717613	0.5807176	1.495593	0.3575948	0.3575948
IPI:IPI00478003.1	0.416743353	0.4167434	1.3349108	0.1932042	0.1932042
IPI:IPI00022391.1	-0.042828456	0.0428285	1.0301315	-0.6039219	0.6039219
IPI:IPI00022229.1	-0.564173412	0.5641734	1.4785401	-9.92E-02	9.92E-02
IPI:IPI00643525.1	-0.656400033	0.6564	1.5761447	-0.166762	0.166762
IPI:IPI00032258.4	-0.680438506	0.6804385	1.6026268	-0.1379815	0.1379815
IPI:IPI00418163.3	-0.676851247	0.6768512	1.5986468	-0.1344288	0.1344288
IPI:IPI00654875.1	-0.680438506	0.6804385	1.6026268	-0.1379815	0.1379815
IPI:IPI00872510.1	-0.16642161	0.1664216	1.1222714	-0.4318455	0.4318455
IPI:IPI00641737.1	-1.014735551	1.0147356	2.0205325	-0.5246559	0.5246559
IPI:IPI00478493.3	-1.327181067	1.3271811	2.5091193	-0.407603	0.407603
IPI:IPI00477597.1	-0.979231667	0.9792317	1.9714152	-0.7184464	0.7184464
IPI:IPI00607707.1	-0.979231667	0.9792317	1.9714152	-0.7184464	0.7184464
IPI:IPI00798167.1	4.70E-02	4.70E-02	1.0330932	0.5446706	0.5446706
IPI:IPI00218732.3	4.70E-02	4.70E-02	1.0330932	0.5446706	0.5446706
IPI:IPI00025426.2	0.430807767	0.4308078	1.3479881	0.1915882	0.1915882
IPI:IPI00748437.2	0.514332131	0.5143321	1.4283328	0.1798155	0.1798155
IPI:IPI00022420.3	-0.694226444	0.6942264	1.6180166	-0.5723486	0.5723486
IPI:IPI00480192.1	-0.694226444	0.6942264	1.6180166	-0.5723486	0.5723486
IPI:IPI00844536.2	-0.694226444	0.6942264	1.6180166	-0.5723486	0.5723486
IPI:IPI00298971.1	-0.727655167	0.7276552	1.6559455	1.60E-02	1.60E-02

5

IPI	FoldChanges_PD_ vs_D_AbsoluteValue	FoldChanges _C_ vs_D (log 2)	FoldChanges _C_ vs_D_AbsoluteValue (log 2)	FoldChanges _C_ vs_D_AbsoluteValue
IPI:IPI00789547.1	1.1534122	0.7494759	0.7494759	1.681182
IPI:IPI00878729.1	1.281288	0.9383124	0.9383124	1.9162854
IPI:IPI00478003.1	1.1433001	0.6099475	0.6099475	1.5262037
IPI:IPI00022391.1	1.5198426	-0.6467504	0.6467504	1.5656377
IPI:IPI00022229.1	1.0711976	-0.6633981	0.6633981	1.5838087
IPI:IPI00643525.1	1.1225362	-0.823162	0.823162	1.7692795
IPI:IPI00032258.4	1.1003645	-0.81842	0.81842	1.7634737
IPI:IPI00418163.3	1.0976581	-0.8112801	0.8112801	1.7547677
IPI:IPI00654875.1	1.1003645	-0.81842	0.81842	1.7634737
IPI:IPI00872510.1	1.3489581	-0.5982671	0.5982671	1.5138971
IPI:IPI00641737.1	1.4385904	-1.5393914	1.5393914	2.9067187
IPI:IPI00478493.3	1.3264801	-1.7347841	1.7347841	3.3282968
IPI:IPI00477597.1	1.6454092	-1.6976781	1.6976781	3.2437847
IPI:IPI00607707.1	1.6454092	-1.6976781	1.6976781	3.2437847
IPI:IPI00798167.1	1.4586873	0.591641	0.591641	1.5069599
IPI:IPI00218732.3	1.4586873	0.591641	0.591641	1.5069599
IPI:IPI00025426.2	1.1420203	0.622396	0.622396	1.5394297
IPI:IPI00748437.2	1.132739	0.6941477	0.6941477	1.6179283
IPI:IPI00022420.3	1.4869422	-1.2665751	1.2665751	2.4058973
IPI:IPI00480192.1	1.4869422	-1.2665751	1.2665751	2.4058973
IPI:IPI00844536.2	1.4869422	-1.2665751	1.2665751	2.4058973
IPI:IPI00298971.1	1.0111602	-0.7116436	0.7116436	1.6376688

Thus, using the disclosed methods, the 45 proteins shown in Table 4 (a combination of the proteins identified in Examples 1 and 2) were found to be associated with type II diabetes mellitus. The individual peptides identified are also listed.

Table 4: 45 proteins associated with type II diabetes mellitus

1. 163 kDa protein; peptide sequences: LVHVEEPHTETVRK (SEQ ID NO: 1), MVSGFIPLKPTVK (SEQ ID NO: 2), DLKPAIVK (SEQ ID NO: 3), DMYSFLEDMGLK (SEQ ID NO: 4), NEDSLVFVQTDK (SEQ ID NO: 5), SASNMAIVDVK (SEQ ID NO: 6), TEHPFTVEEFVLPK (SEQ ID NO: 7), TTVMVK (SEQ ID NO: 8), FQVDNNNR (SEQ ID NO: 9), ATVLNYLPK (SEQ ID NO: 10), RKEYEMK (SEQ ID NO: 11), DTVIKPLLVEPEGLEK (SEQ ID NO: 12);
2. 32 kDa protein; peptide sequences: IQNILTEEPK (SEQ ID NO: 72), YVYIAELLAHK (SEQ ID NO: 73), ILLMDLNEEDPTVLELGITGSK (SEQ ID NO: 74), PTVLELGITGSK (SEQ ID NO: 75),
3. 45 kDa protein; peptide sequences: DKVNSFFSTFK (SEQ ID NO: 13), LTPYADEFKVK (SEQ ID NO: 14);
4. A2M 19 kDa protein (IPI:IPI00878729.1); peptide sequences: NEDSLVFVQTDK (SEQ ID NO: 5), SSSNEEVMFLTVQVK (SEQ ID NO: 76), SVSGKPQYMVLVPSLLHTETTEK (SEQ ID NO: 77), KPQYMVLVPSLLHTETTEK (SEQ ID NO: 78), SSSNEEVMFLTVQVK (SEQ ID NO: 76)
5. A2M 19 kDa protein (IPI:IPI00789547.1); peptide sequences: LLIYAVLPTGDVIGDSAK (SEQ ID NO: 79), TGTHGLLVKQEDMK (SEQ ID NO: 80), DLSFSPSQSLPASHAHLR (SEQ ID NO: 212), VTAAPQSVCALR (SEQ ID NO: 91), TGTHGLLVKQEDMK (SEQ ID NO: 80)
6. alpha-2-macroglobulin; peptide sequences: SASNMAIVDVK (SEQ ID NO: 6), NQGNTWLTAFVLK (SEQ ID NO: 195), HYDGSYSTFGER (SEQ ID NO: 208), NEDSLVFVQTDK (SEQ ID NO: 5), DTVIKPLLVEPEGLEK

(SEQ ID NO: 12), ALLAYAFALAGNQDK (SEQ ID NO: 209), VGFYESDVMGR (SEQ ID NO: 210), DMYSFLEDMGLK (SEQ ID NO: 4), LVHVEEPHTETVRK (SEQ ID NO: 1), MVSGFIPLKPTVK (SEQ ID NO: 2), VGFYESDVMGR (SEQ ID NO: 210), AIGYLNTGYQR (SEQ ID NO: 211), MVSGFIPLKPTVK (SEQ ID NO: 2), TEHPFTVEEFVLPK (SEQ ID NO: 7), LPPNVVEESAR (SEQ ID NO: 205), LSFVKVDSHFR (SEQ ID NO: 206), ATVLNYLPK (SEQ ID NO: 10), TVIKPLLVEPEGLEK (SEQ ID NO: 207), LLIYAVLPTGDVIGDSAK (SEQ ID NO: 79), TGTHGLLVKQEDMK (SEQ ID NO: 80), TEVSSNHVLIYLDK (SEQ ID NO: 81), KDNSVHWERPQKPK (SEQ ID NO: 82), SSSNEEVMFLTVQVK (SEQ ID NO: 76), FEVQVTVPK (SEQ ID NO: 83), SVSGKPQYMVLVPSLLHTETTEK (SEQ ID NO: 77), IAQWQSFQLEGGLK (SEQ ID NO: 84), KDTVIKPLLVEPEGLEK (SEQ ID NO: 85), SGGRTEHPFTVEEFVLPK (SEQ ID NO: 86), LLLQQVSLPELPGEYSMK (SEQ ID NO: 87), TEHPFTVEEFVLPKFEVQVTVPK (SEQ ID NO: 88), VSVQLEASPAFLAVPVEK (SEQ ID NO: 89), VDLSFSPSQSLPASHAHLR (SEQ ID NO: 90), VTAAPQSVCALR (SEQ ID NO: 91), GVPIPNKVIFIR (SEQ ID NO: 92), KPQYMVLVPSLLHTETTEK (SEQ ID NO: 78), ALLAYAFALAGNQDKR (SEQ ID NO: 93), TAQEGDHGSHVYTK (SEQ ID NO: 94), SSSNEEVMFLTVQVK (SEQ ID NO: 76), VVSMDENFHPLNELIPLVYIQDPK (SEQ ID NO: 95), SSGSLLNNAIK (SEQ ID NO: 96), AAQVTIQSSGTFSSK (SEQ ID NO: 97), GGFSSTQDTVVALHALSK (SEQ ID NO: 98), LLLQQVSLPELPGEYSMK (SEQ ID NO: 87), TGTHGLLVKQEDMK (SEQ ID NO: 80), RKEYEMK (SEQ ID NO: 11), FQVDNNNR (SEQ ID NO: 9), NEDSLVFVQTDK (SEQ ID NO: 5), DLKPAIVK (SEQ ID NO: 3), TTVMVK (SEQ ID NO: 8)

7. Serum amyloid P-component; peptide sequences: IVLGQEQDSYGGK (SEQ ID NO: 99), IVLGQEQDSYGGKFDR (SEQ ID NO: 100), AYSLFSYNTQGR (SEQ ID NO: 101), DNELLVYK (SEQ ID NO: 102)

8. Apolipoprotein B-100, peptide sequences: EKLTALTK (SEQ ID NO: 103), IYSLWEHSTK (SEQ ID NO: 104), SEILAHWSPAK (SEQ ID NO: 105), YGMVAQVTQTLK (SEQ ID NO: 106), EFQVPTFTIPK (SEQ ID NO: 107), IDDIWNLEVK (SEQ ID NO: 108), GFEPTLEALFGK (SEQ ID NO: 109), ATGVLYDYVNK (SEQ ID NO: 110), TSSFALNLPTLPEVK (SEQ ID NO: 111), IEGNLIFDPNNYLPK (SEQ ID NO: 112), STSPPKQAEAVLK (SEQ ID NO: 113), VNWEEEAASGLLTSLKDNVPK (SEQ ID NO: 114), AVSMPSFSILGSDVR (SEQ ID NO: 25), SVSLPSLDPASAK (SEQ ID NO: 115), NRNNALDFVTK (SEQ ID NO: 116), VLVDHFGYTK (SEQ ID NO: 117), SKPTVSSSMEFK (SEQ ID NO: 118), GIISALLVPPETEEAK (SEQ ID NO: 119), KLTISEQNIQR (SEQ ID NO: 120), SVGFHLPSR (SEQ ID NO: 27), LLSGGNTLHLVSTTK (SEQ ID NO: 121), LTISEQNIQR (SEQ ID NO: 122), SPAFTDLHLR (SEQ ID NO: 123), EIQIYKK (SEQ ID NO: 124), IADFELPTIIVPEQTIEIPSIK (SEQ ID NO: 125), YHWEHTGLTLR (SEQ ID NO: 126), TGISPLALIK (SEQ ID NO: 127), KSISAALEHK (SEQ ID NO: 128), KMGLAFESTK (SEQ ID NO: 129), VQGVEFSHR (SEQ ID NO: 130), SLHMYANR (SEQ ID NO: 131), LSLESLTSYFSIESSTKGDVK (SEQ ID NO: 132), MTSNFPVDLSDYPK (SEQ ID NO: 133), DAVEKPQEFTIVAFVK (SEQ ID NO: 134), AHLDIAGSLEGHLR (SEQ ID NO: 135), ATLYALSHAVNNYHK (SEQ ID NO: 136), TLQGIPQMIGEVIR (SEQ ID NO: 137), IAELSATAQEIIK (SEQ ID NO: 138), EVYGFNPEGK (SEQ ID NO: 139), DLKVEDIPLAR (SEQ ID NO: 140), TSSFALNLPTLPEVKFPEVDVLTK (SEQ ID NO: 141), KGNVATEISTER (SEQ ID NO: 142), NNALDFVTK (SEQ ID NO: 143), VSALLTPAEQTGTWK (SEQ ID NO: 144), FVTQAEGAK (SEQ ID NO: 145), NLQNNAEWVYQGAIR (SEQ ID NO: 146), ILGEELGFASLHDLQLLGK (SEQ ID NO: 147), NHLQLEGLFFTNGEHTSK (SEQ ID NO: 148), SGGNTLHLVSTTK (SEQ ID NO: 149), DKIGVELTGR (SEQ ID NO: 150), FPEVDVLTK (SEQ ID NO: 151), SISAALEHK (SEQ ID NO: 152), MLETVR (SEQ ID NO: 24), TEVIPPLIENR (SEQ ID NO: 26), SVGFHLPSR (SEQ ID NO: 27)

9. apolipoprotein A-I; peptide sequences: AHVDALR (SEQ ID NO: 17), SGRDYVSQFEGSALGK (SEQ ID NO: 18), PYLDDFQKK (SEQ ID NO: 19), YVSQFEGSALGK (SEQ ID NO: 20), LHELQEK (SEQ ID NO: 21), VSFLSALEEYTK (SEQ ID NO: 22), QKLHELQEK (SEQ ID NO: 23)
10. apolipoprotein A-IV; peptide sequences: DKVNSFFSTFK (SEQ ID NO: 13), LTPYADEFKVK (SEQ ID NO: 14)
11. complement C4B1; peptide sequences: GSFEFPVGDAVSK (SEQ ID NO: 153), YVLPNFEVK (SEQ ID NO: 154), GQIVFMNREPK (SEQ ID NO: 28), ITQVLHFTK (SEQ ID NO: 155), GLEEELQFSLGSK (SEQ ID NO: 156), GPEVQLVAHSPWLK (SEQ ID NO: 29), DFALLSLQVPLKDAK (SEQ ID NO: 31), KYVLPNFEVK (SEQ ID NO: 157), LVNGQSHISLSK (SEQ ID NO: 158), LNMGITDLQGLR (SEQ ID NO: 159), YIYGKPVQGVAYVR (SEQ ID NO: 32), DHAVDLIQK (SEQ ID NO: 160), VGDTLNLNLR (SEQ ID NO: 161), LGQYASPTAKR (SEQ ID NO: 164), SHALQLNNR (SEQ ID NO: 30), SHKPLNMGK (SEQ ID NO: 162), LGQYASPTAK (SEQ ID NO: 163), EAPKVVEEQESR (SEQ ID NO: 165), EMSGSPASGIPVK (SEQ ID NO: 166), GSSTWLTAFVLK (SEQ ID NO: 167), FGLLDEDGKK (SEQ ID NO: 168), VFALDQK (SEQ ID NO: 169), TTNIQGINLLFSSR (SEQ ID NO: 170), PNMIPDGDFNSYVR (SEQ ID NO: 171), YLDKTEQWSTLPPETK (SEQ ID NO: 172), HLVPGAPFLLQALVR (SEQ ID NO: 173), AEMADQAAAWLTR (SEQ ID NO: 174), YRVFALDQK (SEQ ID NO: 175), TLGSEGALSPGGVASLLR (SEQ ID NO: 176), AEFQDALEKLNMGITDLQGLR (SEQ ID NO: 177), VGLSGMAIADVTLLSGFHALR (SEQ ID NO: 178), VDFTLSSER (SEQ ID NO: 179), ASAGLLGAHAAAITAYALTLTK (SEQ ID NO: 180), EVYMPSSIFQDDFVIPDISEPGTWK (SEQ ID NO: 181), EFHLHLR (SEQ ID NO: 182), SCGLHQLLR (SEQ ID NO: 183)
12. ceruloplasmin; peptide sequences: AETGDKVYVHLK (SEQ ID NO: 33), GPEEEHLGILGPVIWAEVGDTIR (SEQ ID NO: 34)
13. clusterin isoform 1; peptide sequences: TLLSNLEEAK (SEQ ID NO: 35), FMETVAEK (SEQ ID NO: 36)

14. clusterin; peptide sequences: TLLSNLEEAK (SEQ ID NO: 35), FMETVAEK (SEQ ID NO: 36)
15. complement C3; peptide sequences: FYYIYNEK (SEQ ID NO: 37), LMNIFLK (SEQ ID NO: 38), IPIEDGSGEVVLSR (SEQ ID NO: 39), IWDVVEK (SEQ ID NO: 40), TIYTPGSTVLYR (SEQ ID NO: 41), KGYTQQLAFR (SEQ ID NO: 42), RIPIEDGSGEVVLSR (SEQ ID NO: 43), VQLSNDFDEYIMAIEQTIK (SEQ ID NO: 44), RQGALELIKK (SEQ ID NO: 45), AAVYHHFISDGVRK (SEQ ID NO: 46), YYTYLIMNK (SEQ ID NO: 47)
16. complement C4; peptide sequences: GPEVQLVAHSPWLK (SEQ ID NO: 29), SHALQLNNR (SEQ ID NO: 30), DFALLSLQVPLKDAK (SEQ ID NO: 31), YIYGKPVQGVAYVR (SEQ ID NO: 32), GQIVFMNREPK (SEQ ID NO: 28)
17. complement C4A; peptide sequences: GSFEFPVGDAVSK (SEQ ID NO: 153), YVLPNFEVK (SEQ ID NO: 154), GQIVFMNREPK (SEQ ID NO: 28), DSSTWLTAFVLK, ITQVLHFTK (SEQ ID NO: 155), GLEEELQFSLGSK (SEQ ID NO: 156), GPEVQLVAHSPWLK (SEQ ID NO: 29), DFALLSLQVPLKDAK (SEQ ID NO: 31), KYVLPNFEVK (SEQ ID NO: 157), LVNGQSHISLSK (SEQ ID NO: 158), LNMGITDLQGLR (SEQ ID NO: 159), YIYGKPVQGVAYVR (SEQ ID NO: 32), DHAVDLIQK (SEQ ID NO: 160), VGDTLNLNLR (SEQ ID NO: 161), LGQYASPTAKR (SEQ ID NO: 164), SHALQLNNR (SEQ ID NO: 30), SHKPLNMGK (SEQ ID NO: 162), LGQYASPTAK (SEQ ID NO: 163), PVAFSVVPTAAAAVSLK (SEQ ID NO: 213), EAPKVVEEQESR (SEQ ID NO: 165), EMSGSPASGIPVK (SEQ ID NO: 166), FGLLDEDGKK (SEQ ID NO: 168), VFALDQK (SEQ ID NO: 169), TTNIQGINLLFSSR (SEQ ID NO: 170), PNMIPDGDFNSYVR (SEQ ID NO: 171), YLDKTEQWSTLPPETK (SEQ ID NO: 172), HLVPGAPFLLQALVR (SEQ ID NO: 173), AEMADQAAAWLTR (SEQ ID NO: 174), YRVFALDQK (SEQ ID NO: 175), TLGSEGALSPGGVASLLR (SEQ ID NO: 176), AEFQDALEKLNMGITDLQGLR (SEQ ID NO: 177), VGLSGMAIADVTLLSGFHALR (SEQ ID NO: 178), VDFTLSSER (SEQ

ID NO: 179), ASAGLLGAHAAAITAYALTLTK (SEQ ID NO: 180), EVYMPSSIFQDDFVIPDISEPGTWK (SEQ ID NO: 181), EFHLHLR (SEQ ID NO: 182), SCGLHQLLR (SEQ ID NO: 183),

18. complement C4B; peptide sequences: GSFEFPVGDAVSK (SEQ ID NO: 153), YVLPNFEVK (SEQ ID NO: 154), GQIVFMNREPK (SEQ ID NO: 28), DSSTWLTAFVLK, ITQVLHFTK (SEQ ID NO: 155), GLEEELQFSLGSK (SEQ ID NO: 156), GPEVQLVAHSPWLK (SEQ ID NO: 29), DFALLSLQVPLKDAK (SEQ ID NO: 31), KYVLPNFEVK (SEQ ID NO: 157), LVNGQSHISLSK (SEQ ID NO: 158), LNMGITDLQGLR (SEQ ID NO: 159), YIYGKPVQGVAYVR (SEQ ID NO: 32), DHAVDLIQK (SEQ ID NO: 160), VGDTLNLNLR (SEQ ID NO: 161), LGQYASPTAKR (SEQ ID NO: 164), SHALQLNNR (SEQ ID NO: 30), SHKPLNMGK (SEQ ID NO: 162), LGQYASPTAK (SEQ ID NO: 163), PVAFSVVPTAAAAVSLK (SEQ ID NO: 213), EAPKVVEEQESR (SEQ ID NO: 165), EMSGSPASGIPVK (SEQ ID NO: 166), FGLLDEDGKK (SEQ ID NO: 168), VFALDQK (SEQ ID NO: 169), TTNIQGINLLFSSR (SEQ ID NO: 170), PNMIPDGDFNSYVR (SEQ ID NO: 171), YLDKTEQWSTLPPETK (SEQ ID NO: 172), HLVPGAPFLLQALVR (SEQ ID NO: 173), YRVFALDQK (SEQ ID NO: 175), TLGSEGALSPGGVASLLR (SEQ ID NO: 176), AEFQDALEKLNMGITDLQGLR (SEQ ID NO: 177), VGLSGMAIADVTLLSGFHALR (SEQ ID NO: 178), VDFTLSSER (SEQ ID NO: 179), EVYMPSSIFQDDFVIPDISEPGTWK (SEQ ID NO: 181), EFHLHLR (SEQ ID NO: 182), SCGLHQLLR (SEQ ID NO: 183)

19. uncharacterized protein C4BPA; peptide sequences: LSLEIEQLELQR (SEQ ID NO: 184), TWYPEVPK (SEQ ID NO: 185), KPDVSHGEMVSGFGPIYNYK (SEQ ID NO: 186), KPDVSHGEMVSGFGPIYNYKDTIVFK (SEQ ID NO: 187)

20. FGA protein; peptide sequences: DSHSLTTNIMEILR (SEQ ID NO: 48), GLIDEVNQDFTNR (SEQ ID NO: 49), LKNSLFEYQK (SEQ ID NO: 50)

21. gelsolin isoform b; peptide sequences: HVVPNEVVVQR (SEQ ID NO: 51), TASDFITK (SEQ ID NO: 52)

22. gelsolin; peptide sequences: HVVPNEVVVQR (SEQ ID NO: 51), TASDFITK (SEQ ID NO: 52)
23. haptoglobin; peptide sequences: VTSIQDWVQK (SEQ ID NO: 188), ILGGHLDAK (SEQ ID NO: 189), TEGDGVYTLNDKK (SEQ ID NO: 190), LRTEGDGVYTLNDKK (SEQ ID NO: 191), TEGDGVYTLNNEK (SEQ ID NO: 192), HYEGSTVPEKK (SEQ ID NO: 193), LRTEGDGVYTLNNEK (SEQ ID NO: 194)
24. HP protein; peptide sequences: VTSIQDWVQK (SEQ ID NO: 188), ILGGHLDAK (SEQ ID NO: 189), TEGDGVYTLNNEK (SEQ ID NO: 192), HYEGSTVPEKK (SEQ ID NO: 193), LRTEGDGVYTLNNEK (SEQ ID NO: 194)
25. isoform 1 of Haptoglobin-related protein; peptide sequences: VTSIQDWVQK (SEQ ID NO: 188), ILGGHLDAK (SEQ ID NO: 189), TEGDGVYTLNDKK (SEQ ID NO: 190), LRTEGDGVYTLNDKK (SEQ ID NO: 191)
26. isoform 2 of Haptoglobin-related protein; peptide sequences: VTSIQDWVQK (SEQ ID NO: 188), ILGGHLDAK (SEQ ID NO: 189), TEGDGVYTLNDKK (SEQ ID NO: 190), LRTEGDGVYTLNDKK (SEQ ID NO: 191)
27. hemopexin; peptide sequences: LWWLDLK (SEQ ID NO: 53), RLWWLDLK (SEQ ID NO: 54), GDKVWVYPPEKK (SEQ ID NO: 55)
28. histidine-rich glycoprotein; peptide sequences: DSPVLIDFFEDTER (SEQ ID NO: 56), ADLFYDVEALDLESPK (SEQ ID NO: 57)
29. hypothetical protein DKFZp779N0926; peptide sequences: RLDGSVDFK (SEQ ID NO: 58), TSTADYAMFK (SEQ ID NO: 59)
30. inter-alpha-trypsin inhibitor heavy chain H2 precursor; peptide sequences: FYNQVSTPLLR (SEQ ID NO: 60), SLAPTAAAK (SEQ ID NO: 61), TILDDLR (SEQ ID NO: 62)
31. serum paraoxonase/arylesterase 1; peptide sequences: IQNILTEEPK (SEQ ID NO: 72), YVYIAELLAHK (SEQ ID NO: 73), ILLMDLNEEDPTVLELGITGSK (SEQ ID NO: 74), PTVLELGITGSK (SEQ ID NO: 75)

32. pregnancy zone protein; peptide sequences: NQGNTWLTAFVLK (SEQ ID NO: 195), MVSGFIPLKPTVK (SEQ ID NO: 2), MVSGFIPLKPTVK (SEQ ID NO: 2), ATVLNYLPK (SEQ ID NO: 10), TEPQYMVLVPSLLHTEAPK (SEQ ID NO: 196), SSGSLLNNAIK (SEQ ID NO: 96)
33. uncharacterized protein PZP; peptide sequences: NQGNTWLTAFVLK (SEQ ID NO: 195), MVSGFIPLKPTVK (SEQ ID NO: 2), MVSGFIPLKPTVK (SEQ ID NO: 2), ATVLNYLPK (SEQ ID NO: 10), SSGSLLNNAIK (SEQ ID NO: 96)
34. plasma retinol-binding protein; peptide sequences: YWGVASFLQK (SEQ ID NO: 197), FSGTWYAMAK (SEQ ID NO: 198), DPNGLPPEAQK (SEQ ID NO: 199), MKYWGVASFLQK (SEQ ID NO: 200)
35. Retinol binding protein 4 (plasma); peptide sequences: YWGVASFLQK (SEQ ID NO: 197), FSGTWYAMAK (SEQ ID NO: 198), DPNGLPPEAQK (SEQ ID NO: 199), MKYWGVASFLQK (SEQ ID NO: 200)
36. similar to fibrinogen, A alpha polypeptide; peptide sequences: LKNSLFEYQK (SEQ ID NO: 50), DSHSLTTNIMEILR (SEQ ID NO: 48), GLIDEVNQDFTNR (SEQ ID NO: 49)
37. splice Isoform 1 of Inter-alpha-trypsin inhibitor heavy chain H4; peptide sequences: NVVFVIDK (SEQ ID NO: 63), LGVYELLLK (SEQ ID NO: 64), ETLFSVMPGLK (SEQ ID NO: 65)
38. splice Isoform 2 of Inter-alpha-trypsin inhibitor heavy chain H4; peptide sequences: LGVYELLLK (SEQ ID NO: 64), ETLFSVMPGLK (SEQ ID NO: 65), NVVFVIDK (SEQ ID NO: 63)
39. splice Isoform Alpha of Fibrinogen alpha/alpha-E chain; peptide sequences: MELERPGGNEITR (SEQ ID NO: 66), LKNSLFEYQK (SEQ ID NO: 50), DSHSLTTNIMEILR (SEQ ID NO: 48), TVIGPDGHKEVTK (SEQ ID NO: 67), GLIDEVNQDFTNR (SEQ ID NO: 49)
40. splice Isoform Alpha-E of Fibrinogen alpha/alpha-E chain; peptide sequences: LKNSLFEYQK (SEQ ID NO: 50), DSHSLTTNIMEILR (SEQ ID NO: 48), MELERPGGNEITR (SEQ ID NO: 66), GLIDEVNQDFTNR (SEQ ID NO: 49), TVIGPDGHKEVTK (SEQ ID NO: 67)

41. splice Isoform Gamma-A of Fibrinogen gamma chain; peptide sequences: RLDGSVDFK (SEQ ID NO: 58), TSTADYAMFK (SEQ ID NO: 59)
42. splice Isoform Gamma-B of Fibrinogen gamma chain; peptide sequences: RLDGSVDFK (SEQ ID NO: 58), TSTADYAMFK (SEQ ID NO: 59)
43. uncharacterized protein RBP4; peptide sequences: YWGVASFLQK (SEQ ID NO: 197), FSGTWYAMAK (SEQ ID NO: 198), DPNGLPPEAQK (SEQ ID NO: 199), MKYWGVASFLQK (SEQ ID NO: 200)
44. vitamin D-binding protein; peptide sequences: ELPEHTVK (SEQ ID NO: 68), KFPSGTFEQVSQLVK (SEQ ID NO: 69), THLPEVFLSK (SEQ ID NO: 70), HLSLLTTLSNR (SEQ ID NO: 71)
45. vitronectin; peptide sequences: DVWGIEGPIDAAFTR (SEQ ID NO: 201), IYISGMAPRPSLAK (SEQ ID NO: 202), DWHGVPGQVDAAMAGR (SEQ ID NO: 203), FEDGVLDPDYPR (SEQ ID NO: 204).

Example 3

This example describes particular changes in expression, such as gene or protein expression, that are associated with predisposition to developing type II diabetes mellitus. Although particular type II diabetes mellitus-related molecules are listed in this example, one skilled in the art will appreciate that other molecules can be used based on the teachings in this disclosure.

In particular examples, detecting nucleic acid or protein expression includes detecting differences in expression (such as an increase, decrease, or both). The method can further include determining the magnitude of the difference in expression, for example relative to a normal sample. Particular examples of type II diabetes mellitus-related molecules that are differentially expressed in association with the diagnosis of a type II diabetes mellitus, and their direction of change (upregulated or downregulated), and the magnitude of the change (as expressed as a percent, and fold change) are provided in Table 5. One skilled in the art will appreciate that other ranges can be determined from Table 3 (fold changes columns, for example control or prediabetic samples versus the prediabetic or diabetic samples).

Table 5: Exemplary patterns of expression associated with type II diabetes

Type II Diabetes Molecule	Change in Expression relative to control (upregulated or downregulated in type II diabetic patients)	Magnitude of the change (increase or decrease relative to control)
IPI:IPI00789547.1	downregulated	at least 40.5% decrease
IPI:IPI00878729.1	downregulated	at least 47.8% decrease
IPI:IPI00478003.1	downregulated	at least 34.5% decrease
IPI:IPI00022391.1	upregulated	at least 1.56 fold increase
IPI:IPI00022229.1	upregulated	at least 1.58 fold increase
IPI:IPI00643525.1	upregulated	at least 1.77 fold increase
IPI:IPI00032258.4	upregulated	at least 1.76 fold increase
IPI:IPI00418163.3	upregulated	at least 1.75 fold increase
IPI:IPI00654875.1	upregulated	at least 1.76 fold increase
IPI:IPI00872510.1	upregulated	at least 1.51 fold increase
IPI:IPI00641737.1	upregulated	at least 2.91 fold increase
IPI:IPI00478493.3	upregulated	at least 3.33 fold increase
IPI:IPI00477597.1	upregulated	at least 3.24 fold increase
IPI:IPI00607707.1	upregulated	at least 3.24 fold increase
IPI:IPI00798167.1	downregulated	at least 33.6% decrease
IPI:IPI00218732.3	downregulated	at least 33.6% decrease a
IPI:IPI00025426.2	downregulated	at least 35.0% decrease
IPI:IPI00748437.2	downregulated	at least 38.2% decrease
IPI:IPI00022420.3	upregulated	at least 2.40 fold increase
IPI:IPI00480192.1	upregulated	at least 2.40 fold increase
IPI:IPI00844536.2	upregulated	at least 2.40 fold increase
IPI:IPI00298971.1	upregulated	at least 1.64 fold increase

The purpose of the foregoing is to enable the United States Patent and Trademark Office and the public generally, especially the scientists, engineers, and practitioners in the art who are not familiar with patent or legal terms or phraseology, to determine quickly from a cursory inspection, the nature and essence of the technical disclosure of the application. The abstract is neither intended to define the invention of the application, which is measured by the claims, nor is it intended to be limiting as to the scope of the invention in any way.

What is claimed is:

1. A method of identifying a human predisposed for the development of type II diabetes mellitus, comprising

analyzing a serum or blood plasma sample of said human to determine the quantity of at least one constituent selected from the group consisting of those proteins and peptides listed in any of Tables 1-4, and combinations thereof.

2. The method of claim 1 wherein said at least one constituent is selected from the group consisting of those proteins and peptides listed in Table 1.

3. The method of claim 1 wherein said at least one constituent is selected from the group consisting of those proteins and peptides listed in Table 2.

4. The method of claim 1 wherein said at least one constituent is selected from the group consisting of those proteins and peptides listed in Table 3.

5. The method of claim 1 wherein said at least one constituent is selected from the group consisting of those proteins and peptides listed in Table 4.

6. The method of claim 1 wherein said at least one constituent comprises all of the peptides listed in Table 4.

7. The method of claim 1 wherein said at least one constituent comprises all of the proteins listed in Table 3.

8. The method of any of claims 1-7, wherein the method has a specificity of at least 90% and a sensitivity of at least 90%.

9. An array, comprising:

oligonucleotides complementary to type II diabetes mellitus-related gene sequences, or

antibodies specific for type II diabetes mellitus-related protein sequences,

wherein the type II diabetes mellitus-related gene and protein sequences comprise any combination of at least 10 of those listed in Tables 1-4.

10. The array of claim 9, wherein the array comprises antibodies specific for at least 10 or all 22 of the type II diabetes mellitus-related protein sequences listed in Table 3.

11. The array of claim 9, wherein the array consists of antibodies specific for all 22 of the type II diabetes mellitus-related protein sequences listed in Table 3.

12. A kit for diagnosing type II diabetes mellitus, comprising:

the array of any of claims 9-11; and

a buffer solution in separate packaging.

专利名称(译)	II型糖尿病的血清标志物		
公开(公告)号	EP2286226A2	公开(公告)日	2011-02-23
申请号	EP2009747474	申请日	2009-05-13
[标]申请(专利权)人(译)	巴特勒记忆研究所		
申请(专利权)人(译)	巴特尔纪念研究所		
当前申请(专利权)人(译)	巴特尔纪念研究所		
[标]发明人	METZ THOMAS O QIAN WEI JUN JACOBS JON M POLPITIYA ASHOKA D CAMP DAVID G SMITH RICHARD D		
发明人	METZ, THOMAS, O. QIAN, WEI-JUN JACOBS, JON, M. POLPITIYA, ASHOKA, D. CAMP, DAVID, G. SMITH, RICHARD, D.		
IPC分类号	G01N33/53 C40B40/10 C07K7/08 G01N33/68		
CPC分类号	G01N33/6893 G01N2800/042		
优先权	12/120043 2008-05-13 US		
其他公开文献	EP2286226A4		
外部链接	Espacenet		

摘要(译)

一种鉴定患有2型糖尿病或患有II型糖尿病的风险增加的人的方法，其使用本文所述的选定生物标志物单独或组合。本公开允许广泛，可靠，筛选大群体基础并提供其他优点，包括制定有效策略以在不同条件下表征，存档和对比来自多种样品类型的数据。还提供了可用于执行此类方法的阵列和套件。