



- (51) International Patent Classification:
G01N 33/68 (2006.01) *G01N 33/574* (2006.01)
- (21) International Application Number:
PCT/SG2016/050337
- (22) International Filing Date:
15 July 2016 (15.07.2016)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
62/193,014 15 July 2015 (15.07.2015) US
- (71) Applicant: NATIONAL UNIVERSITY OF SINGAPORE [SG/SG]; 21 Lower Kent Ridge Road, Singapore 119077 (SG).
- (72) Inventors: JEYASEKHARAN, Anand Devaprasath; c/o National University of Singapore, Cancer Science Institute of Singapore, 21 Lower Kent Ridge Road, Singapore 119077 (SG). ROY, Rony K.; c/o National University of Singapore, Cancer Science Institute of Singapore, 21 Lower Kent Ridge Road, Singapore 119077 (SG). HOPPE, Michal Marek; c/o National University of Singapore, Cancer Science Institute of Singapore, 21 Lower Kent Ridge Road, Singapore 119077 (SG).
- (74) Agent: AMICA LAW LLC; 30 Raffles Place, #14-01 Chevron House, Singapore 048622 (SG).
- (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, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.
- (84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).
- Declarations under Rule 4.17:
— of inventorship (Rule 4.17(iv))

[Continued on next page]

(54) Title: BIOMARKERS FOR THE DIAGNOSIS AND TREATMENT OF GASTRIC CANCER

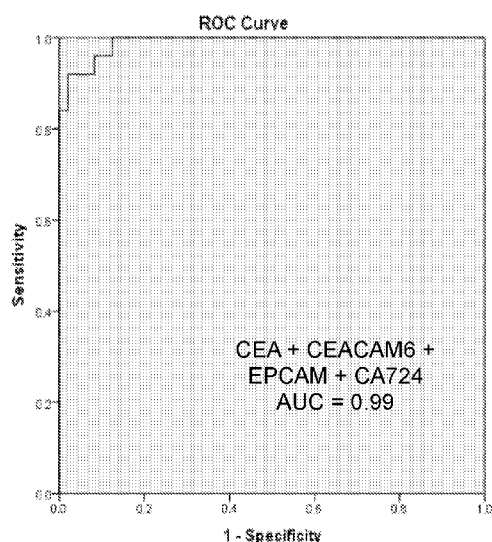


FIG. 3D

(57) Abstract: BIOMARKERS FOR THE DIAGNOSIS AND TREATMENT OF GASTRIC CANCER ABSTRACT OF THE DISCLOSURE
Provided herein are methods of determining whether a subject has gastric cancer. The methods include contacting a biological sample with (i) one or more agents to determine the amount of Carcinoembryonic antigen (CEA) in the biological sample, (ii) one or more agents to determine the amount of carcinoembryonic antigen-related cell adhesion molecule 6 (CEACAM6) in the biological based sample, (iii) one or more agents to determine the amount of epithelial cell adhesion molecule (EPCAM) in the biological sample, (iv) one or more agents to determine the amount of cancer antigen 72-4 (CA72-4) in the biological sample or a combination thereof, and determining the amount of CEA, CEACAM6, EPCAM, CA72-4 or combination thereof in the biological sample. Optionally, the method further comprises determining the amount of CLDN7 and/or CLDN4. Optionally, the method further includes treating the subjects with gastric cancer. Compositions and kits comprising the agents used to determine the amounts of CEA, CEACAM6, EPCAM, CA72-4, CLDN7, CLDN4 and combinations thereof are also provided.



Published:

— *with international search report (Art. 21(3))*

BIOMARKERS FOR THE DIAGNOSIS AND TREATMENT OF GASTRIC CANCER

CROSS REFERENCE TO RELATED APPLICATION

[0001] This application claims priority to U.S. Provisional Application No. 62/193,014, filed July 15, 2015, which is incorporated by reference herein in its entirety.

BACKGROUND OF THE INVENTION

[0002] Early detection of gastric cancers confers a significant mortality benefit, but remains a diagnostic challenge. There are several cell surface markers of gastric cancer which have been reported, but none are in clinical use, partly because it is unclear which ones are the “best” and in what combination.

BRIEF SUMMARY OF THE INVENTION

[0003] Provided herein are methods of determining whether a subject has gastric cancer. The methods include contacting a biological sample with (i) one or more agents to determine the amount of Carcinoembryonic antigen (CEA) in the biological sample, (ii) one or more agents to determine the amount of carcinoembryonic antigen-related cell adhesion molecule 6 (CEACAM6) in the biological based sample, (iii) one or more agents to determine the amount of epithelial cell adhesion molecule (EPCAM) in the biological sample, (iv) one or more agents to determine the amount of cancer antigen 72-4 (CA72-4) in the biological sample, (v) or a combination thereof, and determining the amount of CEA, CEACAM6, EPCAM, CA72-4, or a combination thereof in the biological sample. Optionally, the method further comprises determining the amount of CLDN7 and/or CLDN4. The methods can be performed *in vivo* or *in vitro*. Optionally, the method further includes treating the subjects with gastric cancer. Compositions and kits comprising the agents used to determine the amounts of CEA, CEACAM6, EPCAM, CA72-4, CLDN7, CLDN4 and combinations thereof are also provided.

BRIEF DESCRIPTION OF THE DRAWINGS

[0004] Figure 1 depicts a schematic workflow for the selection of surface markers to detect gastric cancer. Transcript information from the Singapore Gastric Cancer Consortium (SGCC) and The Cancer Genome Atlas (TCGA) public databases were used to identify highly enriched putative membrane transcripts in gastric cancer, which were then refined

using gene expression profile (GEP) datasets and subsequent immunohistochemical experiments on gastric cancer samples.

[0005] Figure 2A are images showing multispectral immunohistochemistry images of stained tissue (chromogenic and fluorescent) unmixed using predefined spectra of each chromogen/fluorophore to obtain pure images for each marker (without interference from auto-fluorescence for fluorescent staining). Figure 2B is an image showing tissue segmentation to recognize tissue architecture to define normal gastric tissue regions from tumor regions. Figure 2C is an image showing cell segmentation used to define cell boundaries on the basis of nuclear counterstaining aided by membrane staining of markers of interest. The tissue and cell segmentations allow us to quantitate the amount of a given marker per cell.

[0006] Figure 3A shows a layout of the HStm-Ade090PG-01 gastric adenocarcinoma tissue microarray (TMA) of 30 cases containing tumor (T) and normal adjacent tissue (NAT). Figure 3B is a graph showing quantitation of all candidate biomarkers in matched normal and advanced gastric adenocarcinoma tissue (compare NAT vs T for each marker). Figure 3C is a graph of ROC curves of all candidate biomarkers to determine whether a subject has gastric cancer (based on quantitative results from Figure 3B). Figure 3D is an ROC curve of a 4-plex combination of the markers (CEA+CEACAM6+EpCAM+CA72-4) in advanced gastric adenocarcinoma, based on combinatorial quantitative results from Figure 3B.

[0007] Figures 4A is a graph showing quantitation of the biomarkers CEA, CEACAM6, EpCAM, and CA72-4 in NAT and early gastric adenocarcinoma tissue (T) (HStm-Ade150CS1-01 TMA) when stained fluorescently in combination on the same slide. Figure 4B are graphs of ROC curves of either single markers or combinations, based on quantitative results from Figure 4A. Figure 4C shows AUC values of single markers and combinations.

[0008] Figure 5 shows combinatorial xenograft staining results. Five (5) gastric cancer patient derived xenografts (PDXs) were harvested and placed in buffered saline. They were treated, without fixation, with a solution of fluorescent dye alone (control), or fluorescently labeled primary antibodies (individually and in combination). Fluorescence and brightfield images were obtained with an inverted epifluorescence microscope. The mixture of 6 antibodies (Mix) yielded a significantly increased signal in this data set, suggesting an additive value of a combination assay in the context of fresh cancer tissue (as may be expected during endoscopy).

DETAILED DESCRIPTION OF THE INVENTION

[0009] As described herein, a combination of novel cell-surface biomarkers was identified and developed to improve the efficiency of gastric cancer detection, e.g., by fluorescent endoscopy. As described herein, a multiple step screening process was performed to identify putative membrane proteins whose expression was significantly higher in gastric cancer when compared to normal tissue. The top 20 hits obtained were tested in a quantitative immunohistochemical platform using a commercially obtained tissue microarray of gastric cancer. The top 6 candidates were then selected on the ability of the antibodies to discriminate gastric cancer from adjacent normal tissue. This combination of antibodies, not reported previously, is backed by quantitative evidence in human material (using the Vectra 2 platform), and provides for both early diagnosis and therapy. A combined reagent using fluorescently labeled versions of antibodies that bind these candidates enables the diagnosis of early gastric cancer on endoscopy.

Definitions

[0010] “Nucleic acid” or “oligonucleotide” or “polynucleotide” or grammatical equivalents used herein means at least two nucleotides covalently linked together. The term “nucleic acid” includes single-, double-, or multiple-stranded DNA, RNA and analogs (derivatives) thereof. Oligonucleotides are typically from about 5, 6, 7, 8, 9, 10, 12, 15, 25, 30, 40, 50 or more nucleotides in length, up to about 100 nucleotides in length. Nucleic acids and polynucleotides are a polymers of any length, including longer lengths, e.g., 200, 300, 500, 1000, 2000, 3000, 5000, 7000, 10,000, etc. In certain embodiments, the nucleic acids herein contain phosphodiester bonds. In other embodiments, nucleic acid analogs are included that may have alternate backbones, comprising, e.g., phosphoramidate, phosphorothioate, phosphorodithioate, or O-methylphosphoroamidite linkages (see Eckstein, *Oligonucleotides and Analogues: A Practical Approach*, Oxford University Press); and peptide nucleic acid backbones and linkages. Other analog nucleic acids include those with positive backbones; non-ionic backbones, and nonribose backbones, including those described in U.S. Patent Nos. 5,235,033 and 5,034,506, and Chapters 6 and 7, *ASC Symposium Series 580, Carbohydrate Modifications in Antisense Research*, Sanghui & Cook, eds. Nucleic acids containing one or more carbocyclic sugars are also included within one definition of nucleic acids. Modifications of the ribose-phosphate backbone may be done for a variety of reasons, e.g., to increase the stability and half-life of such molecules in physiological environments or as probes on a biochip. Mixtures of naturally occurring nucleic acids and analogs can be made;

alternatively, mixtures of different nucleic acid analogs, and mixtures of naturally occurring nucleic acids and analogs may be made.

[0011] Nucleic acids are linear polymers (chains) of nucleotides, which consist of a purine or pyrimidine nucleobase or base, a pentose sugar, and a phosphate group. As used herein, a “polymer backbone” refers to the chain of pentose sugars and phosphate groups lacking the bases normally present in a nucleic acid.

[0012] The terms “identical” or percent sequence “identity,” in the context of two or more nucleic acids or polypeptide sequences, refer to two or more sequences or subsequences that are the same or have a specified percentage of amino acid residues or nucleotides that are the same (i.e., about 60% identity, preferably 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or higher identity over a specified region, when compared and aligned for maximum correspondence over a comparison window or designated region) as measured using a BLAST or BLAST 2.0 sequence comparison algorithms with default parameters described below, or by manual alignment and visual inspection (see, e.g., NCBI web site at ncbi.nlm.nih.gov/BLAST/ or the like). Such sequences are then said to be “substantially identical.” This definition also refers to, or may be applied to, the complement of a test sequence. The definition also includes sequences that have deletions and/or additions, as well as those that have substitutions. Employed algorithms can account for gaps and the like.

[0013] For sequence comparisons, typically one sequence acts as a reference sequence, to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are entered into a computer, subsequence coordinates are designated, if necessary, and sequence algorithm program parameters are designated. Preferably, default program parameters can be used, or alternative parameters can be designated. The sequence comparison algorithm then calculates the percent sequence identities for the test sequences relative to the reference sequence, based on the program parameters.

[0014] A “comparison window”, as used herein, includes reference to a segment of any one of the number of contiguous positions selected from the group consisting of from 20 to 600, usually about 50 to about 200, more usually about 100 to about 150 in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned. Methods of alignment of sequences for comparison are well-known in the art. Optimal alignment of sequences for comparison can be conducted,

e.g., by the local homology algorithm of Smith & Waterman, *Adv. Appl. Math.* 2:482 (1981), by the homology alignment algorithm of Needleman & Wunsch, *J. Mol. Biol.* 48:443 (1970), by the search for similarity method of Pearson & Lipman, *Proc. Nat'l. Acad. Sci. USA* 85:2444 (1988), by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), or by manual alignment and visual inspection (see, e.g., *Current Protocols in Molecular Biology* (Ausubel et al., eds. 1995 supplement)).

[0015] A preferred example of algorithm that is suitable for determining percent sequence identity and sequence similarity are the BLAST and BLAST 2.0 algorithms, which are described in Altschul et al., *Nuc. Acids Res.* 25:3389-3402 (1977) and Altschul et al., *J. Mol. Biol.* 215:403-410 (1990), respectively.

[0016] The terms “polypeptide,” “peptide,” and “protein” are used interchangeably herein to refer to a polymer of amino acid residues. The terms apply to amino acid polymers in which one or more amino acid residue is an artificial chemical mimetic of a corresponding naturally occurring amino acid, as well as to naturally occurring amino acid polymers and non-naturally occurring amino acid polymer.

[0017] The term “amino acid” refers to naturally occurring and synthetic amino acids, as well as amino acid analogs and amino acid mimetics that function in a manner similar to the naturally occurring amino acids. Naturally occurring amino acids are those encoded by the genetic code, as well as those amino acids that are later modified, e.g., hydroxyproline, γ -carboxyglutamate, and O-phosphoserine. Amino acid analogs refers to compounds that have the same basic chemical structure as a naturally occurring amino acid, i.e., an α carbon that is bound to a hydrogen, a carboxyl group, an amino group, and an R group, e.g., homoserine, norleucine, methionine sulfoxide, methionine methyl sulfonium. Such analogs have modified R groups (e.g., norleucine) or modified peptide backbones, but retain the same basic chemical structure as a naturally occurring amino acid. Amino acid mimetics refers to chemical compounds that have a structure that is different from the general chemical structure of an amino acid, but that functions in a manner similar to a naturally occurring amino acid.

[0018] Amino acids may be referred to herein by either their commonly known three letter symbols or by the one-letter symbols recommended by the IUPAC-IUB Biochemical Nomenclature Commission. Nucleotides, likewise, may be referred to by their commonly accepted single-letter codes.

[0019] “Conservatively modified variants” applies to both amino acid and nucleic acid sequences. With respect to particular nucleic acid sequences, conservatively modified variants refers to those nucleic acids which encode identical or essentially identical amino acid sequences, or where the nucleic acid does not encode an amino acid sequence, to essentially identical sequences. Because of the degeneracy of the genetic code, a large number of functionally identical nucleic acids encode any given protein. For instance, the codons GCA, GCC, GCG and GCU all encode the amino acid alanine. Thus, at every position where an alanine is specified by a codon, the codon can be altered to any of the corresponding codons described without altering the encoded polypeptide. Such nucleic acid variations are “silent variations,” which are one species of conservatively modified variations. Every nucleic acid sequence herein which encodes a polypeptide also describes every possible silent variation of the nucleic acid. One of skill will recognize that each codon in a nucleic acid (except AUG, which is ordinarily the only codon for methionine, and TGG, which is ordinarily the only codon for tryptophan) can be modified to yield a functionally identical molecule. Accordingly, each silent variation of a nucleic acid which encodes a polypeptide is implicit in each described sequence with respect to the expression product, but not with respect to actual probe sequences.

[0020] As to amino acid sequences, one of skill will recognize that individual substitutions, deletions or additions to a nucleic acid, peptide, polypeptide, or protein sequence which alters, adds or deletes a single amino acid or a small percentage of amino acids in the encoded sequence is a “conservatively modified variant” where the alteration results in the substitution of an amino acid with a chemically similar amino acid. Conservative substitution tables providing functionally similar amino acids are well known in the art. Such conservatively modified variants are in addition to and do not exclude polymorphic variants, interspecies homologs, and alleles.

[0021] The following eight groups each contain amino acids that are conservative substitutions for one another: 1) Alanine (A), Glycine (G); 2) Aspartic acid (D), Glutamic acid (E); 3) Asparagine (N), Glutamine (Q); 4) Arginine (R), Lysine (K); 5) Isoleucine (I), Leucine (L), Methionine (M), Valine (V); 6) Phenylalanine (F), Tyrosine (Y), Tryptophan (W); 7) Serine (S), Threonine (T); and 8) Cysteine (C), Methionine (M) (see, e.g., Creighton, *Proteins* (1984)).

[0022] The phrase “stringent hybridization conditions” refers to conditions under which a probe will hybridize to its target subsequence, typically in a complex mixture of nucleic acids, but to no other sequences. Stringent conditions are sequence-dependent and will be different in different circumstances. Longer sequences hybridize specifically at higher temperatures. An extensive guide to the hybridization of nucleic acids is found in Tijssen, *Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Probes*, “Overview of principles of hybridization and the strategy of nucleic acid assays” (1993). Generally, stringent conditions are selected to be about 5-10°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength pH. The T_m is the temperature (under defined ionic strength, pH, and nucleic concentration) at which 50% of the probes complementary to the target hybridize to the target sequence at equilibrium (as the target sequences are present in excess, at T_m , 50% of the probes are occupied at equilibrium). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide. For selective or specific hybridization, a positive signal is at least two times background, preferably 10 times background hybridization. Exemplary stringent hybridization conditions can be as following: 50% formamide, 5x SSC, and 1% SDS, incubating at 42°C, or, 5x SSC, 1% SDS, incubating at 65°C, with wash in 0.2x SSC, and 0.1% SDS at 65°C.

[0023] Nucleic acids that do not hybridize to each other under stringent conditions are still substantially identical if the polypeptides which they encode are substantially identical. This occurs, for example, when a copy of a nucleic acid is created using the maximum codon degeneracy permitted by the genetic code. In such cases, the nucleic acids typically hybridize under moderately stringent hybridization conditions. Exemplary “moderately stringent hybridization conditions” include a hybridization in a buffer of 40% formamide, 1 M NaCl, 1% SDS at 37°C, and a wash in 1X SSC at 45°C. A positive hybridization is at least twice background. Those of ordinary skill will readily recognize that alternative hybridization and wash conditions can be utilized to provide conditions of similar stringency. Additional guidelines for determining hybridization parameters are provided in numerous reference, e.g., and *Current Protocols in Molecular Biology*, ed. Ausubel, et al., John Wiley & Sons.

[0024] For PCR, a temperature of about 36°C is typical for low stringency amplification, although annealing temperatures may vary between about 32°C and 48°C depending on primer length. For high stringency PCR amplification, a temperature of about 62°C is typical, although high stringency annealing temperatures can range from about 50°C to about 65°C,

depending on the primer length and specificity. Typical cycle conditions for both high and low stringency amplifications include a denaturation phase of 90°C to 95°C for 30 seconds to 2 minutes, an annealing phase lasting 30 seconds to 2 minutes, and an extension phase of about 72°C for 1 to 2 minutes. Protocols and guidelines for low and high stringency amplification reactions are provided, e.g., in Innis et al. (1990) PCR Protocols, A Guide to Methods and Applications, Academic Press, Inc. N.Y.).

[0025] A “control” sample or value refers to a sample that serves as a reference, usually a known reference, for comparison to a test sample. For example, a test sample can be taken from a test condition, e.g., in the presence of a test compound, and compared to samples from known conditions, e.g., in the absence of the test compound (negative control), or in the presence of a known compound (positive control). For example, a test sample can be taken from a patient suspected of having a given disease (e.g. an autoimmune disease, inflammatory autoimmune disease, cancer, infectious disease, immune disease, or other disease) and compared to a known normal (non-diseased) individual (e.g. a standard control subject). A control can also represent an average value gathered from a number of tests or results. A control value can also be obtained from the same individual, e.g. from an earlier-obtained sample from the patient prior to disease onset. One of skill in the art will recognize that controls can be designed for assessment of any number of parameters. For example, a control can be devised to compare therapeutic benefit based on pharmacological data (e.g., half-life) or therapeutic measures (e.g., comparison of side effects).

[0026] One of skill in the art will understand which controls are valuable in a given situation and be able to analyze data based on comparisons to control values. Controls are also valuable for determining the significance of data. For example, if values for a given parameter are widely variant in controls, variation in test samples will not be considered as significant.

[0027] The terms higher, increases, elevates, or elevation refer to increases above a control. The terms low, lower, reduces, or reduction refer to any decrease below control levels. For example, control levels are *in vivo* levels prior to, or in the absence of, addition of an agent. The reduction includes a complete elimination of the invasiveness. Inhibit, inhibiting, and inhibition mean to decrease an activity, response, condition, disease, or other biological parameter. This can include but is not limited to the complete ablation of the activity, response, condition, or disease. This may also include, for example, a 10% reduction in the

activity, response, condition, or disease as compared to the native or control level. Thus, the reduction can be a 10, 20, 30, 40, 50, 60, 70, 80, 90, 100%, or any amount of reduction in between as compared to native or control levels.

[0028] The term “probe” or “primer”, as used herein, is defined to be one or more nucleic acid fragments whose specific hybridization to a sample can be detected. A probe or primer can be of any length depending on the particular technique it will be used for. For example, PCR primers are generally between 10 and 40 nucleotides in length, while nucleic acid probes for, e.g., a Southern blot, can be more than a hundred nucleotides in length. The probe may be unlabeled or labeled as described below so that its binding to the target or sample can be detected. The probe can be produced from a source of nucleic acids from one or more particular (preselected) portions of a chromosome, e.g., one or more clones, an isolated whole chromosome or chromosome fragment, or a collection of polymerase chain reaction (PCR) amplification products. The length and complexity of the nucleic acid fixed onto the target element is not critical to the invention. One of skill can adjust these factors to provide optimum hybridization and signal production for a given hybridization procedure, and to provide the required resolution among different genes or genomic locations.

[0029] The probe may also be isolated nucleic acids immobilized on a solid surface (e.g., nitrocellulose, glass, quartz, fused silica slides), as in an array. In some embodiments, the probe may be a member of an array of nucleic acids as described, for instance, in WO 96/17958. Techniques capable of producing high density arrays can also be used for this purpose (see, e.g., Fodor (1991) *Science* 767-773; Johnston (1998) *Curr. Biol.* 8: R171-R174; Schummer (1997) *Biotechniques* 23: 1087-1092; Kern (1997) *Biotechniques* 23: 120-124; U.S. Patent No. 5,143,854).

[0030] “Antibody” refers to a polypeptide comprising a framework region from an immunoglobulin gene or fragments thereof that specifically binds and recognizes an antigen. The recognized immunoglobulin genes include the kappa, lambda, alpha, gamma, delta, epsilon, and mu constant region genes, as well as the myriad immunoglobulin variable region genes. Light chains are classified as either kappa or lambda. Heavy chains are classified as gamma, mu, alpha, delta, or epsilon, which in turn define the immunoglobulin classes, IgG, IgM, IgA, IgD and IgE, respectively. Typically, the antigen-binding region of an antibody will be most critical in specificity and affinity of binding. In some embodiments, antibodies or fragments of antibodies may be derived from different organisms, including humans, mice,

rats, hamsters, camels, etc. Antibodies of the invention may include antibodies that have been modified or mutated at one or more amino acid positions to improve or modulate a desired function of the antibody (e.g. glycosylation, expression, antigen recognition, effector functions, antigen binding, specificity, etc.).

[0031] An exemplary immunoglobulin (antibody) structural unit comprises a tetramer. Each tetramer is composed of two identical pairs of polypeptide chains, each pair having one “light” (about 25 kD) and one “heavy” chain (about 50-70 kD). The N-terminus of each chain defines a variable region of about 100 to 110 or more amino acids primarily responsible for antigen recognition. The terms variable light chain (VL) and variable heavy chain (VH) refer to these light and heavy chains respectively.

[0032] Antibodies exist, e.g., as intact immunoglobulins or as a number of well-characterized fragments produced by digestion with various peptidases. Thus, for example, pepsin digests an antibody below the disulfide linkages in the hinge region to produce F(ab)², a dimer of Fab which itself is a light chain joined to VH-CH1 by a disulfide bond. The F(ab)² may be reduced under mild conditions to break the disulfide linkage in the hinge region, thereby converting the F(ab)² dimer into an Fab' monomer. The Fab' monomer is essentially Fab with part of the hinge region (see Fundamental Immunology (Paul ed., 3d ed. 1993). While various antibody fragments are defined in terms of the digestion of an intact antibody, one of skill will appreciate that such fragments may be synthesized de novo either chemically or by using recombinant DNA methodology. Thus, the term antibody, as used herein, also includes antibody fragments either produced by the modification of whole antibodies, or those synthesized de novo using recombinant DNA methodologies (e.g., single chain Fv) or those identified using phage display libraries (see, e.g., McCafferty et al., Nature 348:552-554 (1990)).

[0033] For preparation of suitable antibodies of the invention and for use according to the invention, e.g., recombinant, monoclonal, or polyclonal antibodies, many techniques known in the art can be used (see, e.g., Kohler & Milstein, Nature 256:495-497 (1975); Kozbor et al., Immunology Today 4: 72 (1983); Cole et al., pp. 77-96 in Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, Inc. (1985); Coligan, Current Protocols in Immunology (1991); Harlow & Lane, Antibodies, A Laboratory Manual (1988); and Goding, Monoclonal Antibodies: Principles and Practice (2d ed. 1986)). The genes encoding the heavy and light chains of an antibody of interest can be cloned from a cell, e.g., the genes encoding a

monoclonal antibody can be cloned from a hybridoma and used to produce a recombinant monoclonal antibody. Gene libraries encoding heavy and light chains of monoclonal antibodies can also be made from hybridoma or plasma cells. Random combinations of the heavy and light chain gene products generate a large pool of antibodies with different antigenic specificity (see, e.g., Kuby, *Immunology* (3rd ed. 1997)). Techniques for the production of single chain antibodies or recombinant antibodies (U.S. Patent 4,946,778, U.S. Patent No. 4,816,567) can be adapted to produce antibodies to polypeptides of this invention. Also, transgenic mice, or other organisms such as other mammals, may be used to express humanized or human antibodies (see, e.g., U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,016, Marks et al., *Bio/Technology* 10:779-783 (1992); Lonberg et al., *Nature* 368:856-859 (1994); Morrison, *Nature* 368:812-13 (1994); Fishwild et al., *Nature Biotechnology* 14:845-51 (1996); Neuberger, *Nature Biotechnology* 14:826 (1996); and Lonberg & Huszar, *Intern. Rev. Immunol.* 13:65-93 (1995)). Alternatively, phage display technology can be used to identify antibodies and heteromeric Fab fragments that specifically bind to selected antigens (see, e.g., McCafferty et al., *Nature* 348:552-554 (1990); Marks et al., *Biotechnology* 10:779-783 (1992)). Antibodies can also be made bispecific, i.e., able to recognize two different antigens (see, e.g., WO 93/08829, Traunecker et al., *EMBO J.* 10:3655-3659 (1991); and Suresh et al., *Methods in Enzymology* 121:210 (1986)). Antibodies can also be heteroconjugates, e.g., two covalently joined antibodies, or immunotoxins (see, e.g., U.S. Patent No. 4,676,980, WO 91/00360; WO 92/200373; and EP 03089).

[0034] Methods for humanizing or primatizing non-human antibodies are well known in the art (e.g., U.S. Patent Nos. 4,816,567; 5,530,101; 5,859,205; 5,585,089; 5,693,761; 5,693,762; 5,777,085; 6,180,370; 6,210,671; and 6,329,511; WO 87/02671; EP Patent Application 0173494; Jones et al. (1986) *Nature* 321:522; and Verhoyen et al. (1988) *Science* 239:1534). Humanized antibodies are further described in, e.g., Winter and Milstein (1991) *Nature* 349:293. Generally, a humanized antibody has one or more amino acid residues introduced into it from a source which is non-human. These non-human amino acid residues are often referred to as import residues, which are typically taken from an import variable domain. Humanization can be essentially performed following the method of Winter and co-workers (see, e.g., Morrison et al., *PNAS USA*, 81:6851-6855 (1984), Jones et al., *Nature* 321:522-525 (1986); Riechmann et al., *Nature* 332:323-327 (1988); Morrison and Oi, *Adv. Immunol.*, 44:65-92 (1988), Verhoeven et al., *Science* 239:1534-1536 (1988) and Presta,

Curr. Op. Struct. Biol. 2:593-596 (1992), Padlan, Molec. Immun., 28:489-498 (1991); Padlan, Molec. Immun., 31(3):169-217 (1994)), by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. Accordingly, such humanized antibodies are chimeric antibodies (U.S. Patent No. 4,816,567), wherein substantially less than an intact human variable domain has been substituted by the corresponding sequence from a non-human species. In practice, humanized antibodies are typically human antibodies in which some CDR residues and possibly some FR residues are substituted by residues from analogous sites in rodent antibodies. For example, polynucleotides comprising a first sequence coding for humanized immunoglobulin framework regions and a second sequence set coding for the desired immunoglobulin complementarity determining regions can be produced synthetically or by combining appropriate cDNA and genomic DNA segments. Human constant region DNA sequences can be isolated in accordance with well known procedures from a variety of human cells.

[0035] A “chimeric antibody” is an antibody molecule in which (a) the constant region, or a portion thereof, is altered, replaced or exchanged so that the antigen binding site (variable region) is linked to a constant region of a different or altered class, effector function and/or species, or an entirely different molecule which confers new properties to the chimeric antibody, e.g., an enzyme, toxin, hormone, growth factor, drug, etc.; or (b) the variable region, or a portion thereof, is altered, replaced or exchanged with a variable region having a different or altered antigen specificity. The preferred antibodies of, and for use according to the invention include humanized and/or chimeric monoclonal antibodies.

[0036] Techniques for conjugating agents to antibodies are well known (see, e.g., Freise and Wu, In vivo imaging with antibodies and engineered fragments, Mol. Immunol. pii: S0161-5890(15)00360-0 (2015); Azhdarinia, et al., Mol Imaging Biol. Jun;14(3):261-76 (2012); Arnon et al., “Monoclonal Antibodies For Immunotargeting Of Drugs In Cancer Therapy”, in Monoclonal Antibodies And Cancer Therapy, Reisfeld et al. (eds.), pp. 243-56 (Alan R. Liss, Inc. 1985); Hellstrom et al., “Antibodies For Drug Delivery” in Controlled Drug Delivery (2nd Ed.), Robinson et al. (eds.), pp. 623-53 (Marcel Dekker, Inc. 1987); Thorpe, “Antibody Carriers Of Cytotoxic Agents In Cancer Therapy: A Review” in Monoclonal Antibodies ‘84: Biological And Clinical Applications, Pinchera et al. (eds.), pp. 475-506 (1985); and Thorpe et al., “The Preparation And Cytotoxic Properties Of Antibody-Toxin Conjugates”, Immunol. Rev., 62:119-58 (1982)).

[0037] The phrase “specifically (or selectively) binds” to an antibody or “specifically (or selectively) immunoreactive with,” when referring to a protein or peptide, refers to a binding reaction that is determinative of the presence of the protein, often in a heterogeneous population of proteins and other biologics. Thus, under designated immunoassay conditions, the specified antibodies bind to a particular protein at least two times the background and more typically more than 10 to 100 times background. Specific binding to an antibody under such conditions requires an antibody that is selected for its specificity for a particular protein. For example, polyclonal antibodies can be selected to obtain only those polyclonal antibodies that are specifically immunoreactive with the selected antigen and not with other proteins. This selection may be achieved by subtracting out antibodies that cross-react with other molecules. A variety of immunoassay formats may be used to select antibodies specifically immunoreactive with a particular protein. For example, solid-phase ELISA immunoassays are routinely used to select antibodies specifically immunoreactive with a protein (see, e.g., Harlow & Lane, *Using Antibodies, A Laboratory Manual* (1998) for a description of immunoassay formats and conditions that can be used to determine specific immunoreactivity).

[0038] A “labeled probe” or “labeled antibody” is one that is bound, either covalently, through a linker or a chemical bond, or noncovalently, through ionic, van der Waals, electrostatic, or hydrogen bonds to a label such that the presence of the probe or antibody may be detected by detecting the presence of the label bound to the probe or antibody. Alternatively, a method using high affinity interactions may achieve the same results where one of a pair of binding partners binds to the other, e.g., biotin, streptavidin.

[0039] A “label” or a “detectable moiety” is a composition detectable by spectroscopic, photochemical, biochemical, immunochemical, chemical, or other physical means. For example, useful labels include ³²P, fluorescent dyes, electron-dense reagents, enzymes (e.g., as commonly used in an ELISA), biotin, digoxigenin, or haptens and proteins or other entities which can be made detectable, e.g., by incorporating a radiolabel into a peptide or antibody specifically reactive with a target peptide. Any method known in the art for conjugating an antibody to the label may be employed, e.g., using methods described in Hermanson, *Bioconjugate Techniques* 1996, Academic Press, Inc., San Diego.

[0040] As used herein, the term “cancer” refers to all types of cancer, neoplasm, or malignant tumors found in mammals, including leukemia, carcinomas and sarcomas.

Exemplary cancers include cancer of the brain, breast, cervix, colon, head & neck, liver, kidney, lung, non-small cell lung, melanoma, mesothelioma, ovary, sarcoma, stomach, uterus and Medulloblastoma. Additional examples include, Hodgkin's Disease, Non-Hodgkin's Lymphoma, multiple myeloma, neuroblastoma, ovarian cancer, rhabdomyosarcoma, primary thrombocytosis, primary macroglobulinemia, primary brain tumors, cancer, malignant pancreatic insulanoma, malignant carcinoid, urinary bladder cancer, premalignant skin lesions, testicular cancer, lymphomas, thyroid cancer, neuroblastoma, esophageal cancer, genitourinary tract cancer, malignant hypercalcemia, endometrial cancer, adrenal cortical cancer, neoplasms of the endocrine and exocrine pancreas, and prostate cancer.

[0041] As used herein, the term “gastric cancer,” refers to a cancer developing from the lining of the stomach. Most cases of stomach cancers are gastric carcinomas; however, lymphomas and mesenchymal tumors may also develop within the stomach. Diagnosis is usually by biopsy done during endoscopy.

[0042] As used herein, “treating” or “treatment of” a condition, disease or disorder or symptoms associated with a condition, disease or disorder refers to an approach for obtaining beneficial or desired results, including clinical results. Beneficial or desired clinical results can include, but are not limited to, alleviation or amelioration of one or more symptoms or conditions, diminishment of extent of condition, disorder or disease, stabilization of the state of condition, disorder or disease, prevention of development of condition, disorder or disease, prevention of spread of condition, disorder or disease, delay or slowing of condition, disorder or disease progression, delay or slowing of condition, disorder or disease onset, amelioration or palliation of the condition, disorder or disease state, and remission, whether partial or total. “Treating” can also mean prolonging survival of a subject beyond that expected in the absence of treatment. “Treating” can also mean inhibiting the progression of the condition, disorder or disease, slowing the progression of the condition, disorder or disease temporarily, although in some instances, it involves halting the progression of the condition, disorder or disease permanently. As used herein the terms treatment, treat, or treating refers to a method of reducing the effects of one or more symptoms of a disease or condition characterized by expression of the protease or symptom of the disease or condition characterized by expression of the protease. Thus in the disclosed method, treatment can refer to a 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 100% reduction in the severity of an established disease, condition, or symptom of the disease or condition. For example, a method for treating a disease is considered to be a treatment if there is a 10% reduction in one or more

symptoms of the disease in a subject as compared to a control. Thus the reduction can be a 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, or any percent reduction in between 10% and 100% as compared to native or control levels. It is understood that treatment does not necessarily refer to a cure or complete ablation of the disease, condition, or symptoms of the disease or condition. Further, as used herein, references to decreasing, reducing, or inhibiting include a change of 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% or greater as compared to a control level and such terms can include but do not necessarily include complete elimination.

[0043] Provided herein are methods of determining or detecting gastric cancer in a subject. One of skill will appreciate that each of the detection methods and agents recited below can be used in a variety of combinations. Specifically, provided is a method of determining whether a subject has gastric cancer. The method includes contacting a biological sample with (i) one or more agents to determine the amount of Carcinoembryonic antigen (CEA) in the biological sample, (ii) one or more agents to determine the amount of carcinoembryonic antigen-related cell adhesion molecule 6 (CEACAM6) in the biological based sample, (iii) one or more agents to determine the amount of epithelial cell adhesion molecule (EPCAM) in the biological sample, (iv) one or more agents to determine the amount of cancer antigen 72-4 (CA72-4) in the biological sample or any combination thereof. The method also includes determining the amount of CEA, CEACAM6, EPCAM, CA72-4, or a combination thereof in the biological sample. In the method, increased levels of each of CEA, CEACAM6, EPCAM and CA72-4 in the biological sample as compared to a control indicates the subject has gastric cancer. Optionally, the method further comprises contacting the biological sample with an agent to determine the amount of claudin-7 (CLDN7) in the biological sample. Optionally, the method further comprises contacting the biological sample with an agent to determine the amount of claudin-4 (CLDN4) in the biological sample. Optionally, the method includes detecting two, three, four, five or more of CEA, CEACAM6, EPCAM, CA72-4, CLDN4 and CLDN7 in any combination. For example, the method includes detecting three biomarkers selected from CEA, CEACAM6, EPCAM, and CA72-4, and one biomarker selected from CLDN4 and CLDN7. By way of another example, the method includes detecting four biomarkers selected from CEA, CEACAM6, EPCAM, CA72-4, CLDN4 and CLDN7. Thus, provided is a method of contacting a biological sample with two, three, four, five or six of the following (i) one or more agents to determine the amount of Carcinoembryonic antigen (CEA) in the biological sample, (ii) one or more agents to

determine the amount of carcinoembryonic antigen-related cell adhesion molecule 6 (CEACAM6) in the biological based sample, (iii) one or more agents to determine the amount of epithelial cell adhesion molecule (EPCAM) in the biological sample, (iv) one or more agents to determine the amount of cancer antigen 72-4 (CA72-4) in the biological sample, (v) one or more agents to determine the amount of CLDN4 in the biological sample, and (vi) one or more agents to determine the amount of CLDN7 in the biological sample. The provided method includes determining the amounts of the two or more of CEA, CEACAM6, EPCAM, CA72-4, CLDN4 and CLDN7 in any combination, an increase in the amounts indicating the subject has gastric cancer. The methods can be carried out *in vivo* or *in vitro*. Thus, when the methods are performed *in vitro*, the methods optionally include obtaining a biological sample from a subject. Optionally, the subject is a non-human animal or a human.

[0044] In the provided methods and as discussed throughout, a variety of agents can be used to determine the amounts of CEA, CEACAM6, EPCAM, CA72-4, CLDN4, and CLDN7. For example, the agents can be proteins, e.g., antibodies, or nucleic acids, e.g., primers or probes. Optionally, the agents are primers or probes for determining the amount of the biomarker. Optionally, the agent is a labeled antibody. Suitable antibodies to CEA, CEACAM6, EPCAM, CA72-4, CLDN4 and CLDN7 include, but are not limited to, those that can be obtained from commercial sources. See, for example, CEA antibodies from Dako Denmark A/S, catalog number M7072, Clone II-7 (Glostrup Denmark), CEACAM6 antibodies from Abcam, catalog number ab78029, Clone 9A6 (Cambridge, United Kingdom), EPCAM antibodies from Santa Cruz Biotechnology, Inc., catalog number sc-25308, clone C-10 (Dallas, TX), CA72-4 antibodies from Santa Cruz Biotechnology, Inc., catalog number sc-20043, clone CC49 (Dallas, TX), CLDN4 antibodies from Thermo Fisher, catalog number 32-9400, clone 3E2C1 (Rockford, IL), and CLDN7 antibodies from Invitrogen, catalog number 37-4800, clone 5D10F3 (Camarillo, CA). Optionally, the provided methods include contacting the sample with a labeled antibody that binds CEA, a labeled antibody that binds CEACAM6, a labeled antibody that binds EPCAM, and a labeled antibody that binds CA72-4. Optionally, the provided methods include contacting the sample with a labeled antibody that binds CEA, a labeled antibody that binds CEACAM6, a labeled antibody that binds EPCAM, a labeled antibody that binds CA72-4, a labeled antibody that binds CLDN7 and a labeled antibody that binds CLDN4. Optionally, the provided method is an immunofluorescence based assay. Optionally, the immunofluorescence based assay is flow

cytometry, fluorescence activated cell sorting (FACS), or an enzyme-linked immune assay (ELISA).

[0045] The provided methods can be carried out *in vivo* or *in vitro*. Thus, optionally, the step of contacting the biological sample occurs *in vivo*. Optionally, the *in vivo* methods include administering to the subject the agents that bind CEA, CEACAM6, EPCAM and CA72-4. Optionally, the methods further include administering to the subject an agent that binds CLDN7 and/or an agent that binds CLDN4. The agents can be administered simultaneously or sequentially. Optionally, the agents are administered sequentially. The agents can be administered sequentially in any order. By way of example, the agents can be administered sequentially in order of CEA, EPCAM, CA72-4 and CEACAM6 or any other desired order. Optionally, the agents are labeled antibodies. Optionally, the label is a fluorescent molecule. Optionally, the antibodies are detected visually using an endoscope. Optionally, the biological sample is the stomach or digestive track. As discussed, the step of contacting the biological sample optionally, occurs *in vitro*. Optionally, *in vitro*, the biological sample is a tissue sample, e.g., of the stomach.

[0046] As used herein, biological samples include, but are not limited to, cells, tissues and bodily fluids. Optionally, the biological sample is the stomach or stomach tissue. Bodily fluids that used to evaluate the presence or absence of the herein disclosed biomarkers include without limitation blood, urine, serum, tears, lymph, bile, cerebrospinal fluid, interstitial fluid, aqueous or vitreous humor, colostrum, sputum, amniotic fluid, saliva, perspiration, transudate, exudate, and synovial fluid. By way of example, levels of biomarker are measured during endoscopy before and/or after treatment in a subject. Optionally, the biomarkers are visualized within the stomach, e.g., to determine the location of for biopsy. Optionally, a biopsy of stomach tissue is obtained for analysis. Biopsy refers to the removal of a sample of tissue for purposes of diagnosis. For example, a biopsy is from a cancer or tumor, including a sample of tissue from an abnormal area or an entire tumor. Optionally, the biological sample is a stomach-derived biological sample.

[0047] By biomarker is meant any assayable characteristics or compositions that are used to identify or monitor a condition (e.g., a tumor or other cancer, or lack thereof) or a therapy for said condition in a subject or sample. A biomarker is, for example, CA72-4, CEA, CEACAM6, EPCAM, CLDN7, or CLDN4, whose presence, absence, or relative amount is used to identify a condition or status of a condition in a subject or sample. Biomarkers

identified herein are measured to determine levels, expression, activity, or to detect variants. Thus, disclosed herein are biomarkers and methods for identifying and using the biomarkers. A biomarker is, for example, a protein or combination of proteins whose presence, absence, or relative amount is used to identify a condition or status of a condition in a subject or sample.

[0048] Methods for detecting and identifying nucleic acids and proteins and interactions between such molecules involve conventional molecular biology, microbiology, and recombinant DNA techniques within the skill of the art. Such techniques are explained fully in the literature (see, e.g., Sambrook, Fritsch & Maniatis, *Molecular Cloning: A Laboratory Manual*, Second Edition 1989, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.; *Animal Cell Culture*, R. I. Freshney, ed., 1986).

[0049] Methods for detecting nucleic acids are largely cumulative with the nucleic acid detection assays and include, for example, Northern blots, RT-PCR, arrays including microarrays and sequencing including high-throughput sequencing methods. In some embodiments, a reverse transcriptase reaction is carried out and the targeted sequence is then amplified using standard PCR. Quantitative PCR (qPCR) or real time PCR (RT-PCR) is useful for determining relative expression levels, when compared to a control. Quantitative PCR techniques and platforms are known in the art, and commercially available (see, e.g., the qPCR Symposium website, available at qpersymposium.com). Nucleic acid arrays are also useful for detecting nucleic acid expression. Customizable arrays are available from, e.g., Affymatrix.

[0050] Optionally, methods for detecting nucleic acids include sequencing methods. Sequencing methods are known and can be performed with a variety of platforms including, but not limited to, platforms provided by Illumina, Inc., (La Jolla, CA) or Life Technologies (Carlsbad, CA). See, e.g., Wang, et al., *Nat Rev Genet.* 10(1):57-63 (2009); and Martin, *Nat Rev Genet.* 12(10):671-82 (2011). Optionally, methods for detecting nucleic acids include microarray methods, which are known and can be performed with a variety of platforms including, but not limited to, platforms provided by Ambion, Inc., (Austin, TX) and Life Technologies (Carlsbad, CA).

[0051] Optionally, the provided methods occur *in vivo*. However, certain methods disclosed herein involve collecting a biological sample from a subject. The collection of biological samples is performed by standard methods. Typically, once a sample is collected,

the biomarkers are detected and measured. The disclosed biomarkers are detected using any suitable technique. Further, molecules that interact with or bind to the disclosed biomarkers, such as antibodies to a biomarker, are detected using known techniques. Many suitable techniques--such as techniques generally known for the detection of proteins, peptides and other analytes and antigens--are known, some of which are described below. In general, these techniques involve direct imaging (e.g., microscopy), immunoassays, or functional determination. By functional determination is meant that a given biomarker, such as a protein that has a function are detected by the detection of said function. For example, an enzyme is detected by evaluating its activity on its substrate.

[0052] Immunodetection methods are used for detecting, binding, purifying, removing and quantifying various molecules including the disclosed biomarkers. Further, antibodies and ligands to the disclosed biomarkers are detected. For example, the disclosed biomarkers are employed to detect antibodies having reactivity therewith. Standard immunological techniques are described, e.g., in Hertenberg, et al., Weir's Handbook of Experimental Immunology, vols. 1-4 (1996); Coligan, Current Protocols in Immunology (1991); Methods in Enzymology, vols. 70, 73, 74, 84, 92, 93, 108, 116, 121, 132, 150, 162, and 163; and Paul, Fundamental Immunology (3d ed. 1993), each of which is incorporated herein by reference in its entirety and specifically for teachings regarding immunodetection methods.

[0053] The steps of various useful immunodetection methods have been described in the scientific literature, such as, e.g., Maggio et al., Enzyme-Immunoassay, (1987) and Nakamura, et al., Enzyme Immunoassays: Heterogeneous and Homogeneous Systems, Handbook of Experimental Immunology, Vol. 1: Immunochemistry, 27.1-27.20 (1986), each of which is incorporated herein by reference in its entirety and specifically for its teaching regarding immunodetection methods. Immunoassays, in their most simple and direct sense, are binding assays involving binding between antibodies and antigen. Many types and formats of immunoassays are known and all are suitable for detecting the disclosed biomarkers. Examples of immunoassays are enzyme linked immunosorbent assays (ELISAs), radioimmunoassays (RIA), radioimmune precipitation assays (RIPA), immunobead capture assays, Western blotting, dot blotting, gel-shift assays, Flow cytometry, protein arrays, multiplexed bead arrays, magnetic capture, in vivo imaging, fluorescence resonance energy transfer (FRET), and fluorescence recovery/localization after photobleaching (FRAP/ FLAP).

[0054] In general, immunoassays involve contacting a sample suspected of containing a molecule of interest (such as the disclosed biomarkers) with an antibody to the molecule of interest or contacting an antibody to a molecule of interest (such as antibodies to the disclosed biomarkers) with a molecule that is bound by the antibody, as the case may be, under conditions effective to allow the formation of immunocomplexes. Contacting a sample with the antibody to the molecule of interest or with the molecule that is bound by an antibody to the molecule of interest under conditions effective and for a period of time sufficient to allow the formation of immune complexes (primary immune complexes) is generally a matter of simply bringing into contact the molecule or antibody and the sample and incubating the mixture for a period of time long enough for the antibodies to form immune complexes with, i.e., to bind to, any molecules (e.g., antigens) present to which the antibodies can bind. In many forms of immunoassay, the sample-antibody composition, such as a tissue section, ELISA plate, dot blot or Western blot, is washed to remove any non-specifically bound antibody species, allowing only those antibodies specifically bound within the primary immune complexes to be detected.

[0055] Immunoassays include methods for detecting or quantifying the amount of a molecule of interest (such as the disclosed biomarkers or their antibodies) in a sample, which methods generally involve the detection or quantitation of any immune complexes formed during the binding process. In general, the detection of immunocomplex formation is achieved through the application of numerous approaches. These methods are generally based upon the detection of a label or marker, such as any radioactive, fluorescent, biological or enzymatic tags or any other known label. See, for example, U.S. Patents 3,817,837; 3,850,752; 3,939,350; 3,996,345; 4,277,437; 4,275,149 and 4,366,241, each of which is incorporated herein by reference in its entirety and specifically for teachings regarding immunodetection methods and labels.

[0056] As used herein, a label includes a fluorescent dye, a member of a binding pair, such as biotin/streptavidin, a metal (e.g., gold), or an epitope tag that specifically interacts with a molecule to be detected, such as by producing a colored substrate or fluorescence. Substances suitable for detectably labeling proteins include fluorescent dyes (also known herein as fluorochromes and fluorophores) and enzymes that react with colorimetric substrates (e.g., horseradish peroxidase). The use of fluorescent dyes is generally used in the practice of the invention as they are detected at very low amounts. Fluorophores are compounds or molecules that luminesce. Typically fluorophores absorb electromagnetic

energy at one wavelength and emit electromagnetic energy at a second wavelength. In the case where multiple antigens are reacted with a single array, each antigen is labeled with a distinct fluorescent compound for simultaneous detection. Labeled spots on the array are detected using a fluorimeter, the presence of a signal indicating an antigen bound to a specific antibody.

[0057] Labeling is either direct or indirect. In direct labeling, the detecting antibody (the antibody for the molecule of interest) or detecting molecule (the molecule that is bound by an antibody to the molecule of interest) includes a label. Detection of the label indicates the presence of the detecting antibody or detecting molecule, which in turn indicates the presence of the molecule of interest or of an antibody to the molecule of interest, respectively. In indirect labeling, an additional molecule or moiety is brought into contact with, or generated at the site of, the immunocomplex. For example, a signal-generating molecule or moiety such as an enzyme is attached to or associated with the detecting antibody or detecting molecule. The signal-generating molecule then generates a detectable signal at the site of the immunocomplex. For example, an enzyme, when supplied with suitable substrate, produces a visible or detectable product at the site of the immunocomplex. ELISAs use this type of indirect labeling.

[0058] As another example of indirect labeling, an additional molecule (which is referred to as a binding agent) that can bind to either the molecule of interest or to the antibody (primary antibody) to the molecule of interest, such as a second antibody to the primary antibody, is contacted with the immunocomplex. The additional molecule optionally has a label or signal-generating molecule or moiety. The additional molecule is, for example, an antibody, which is termed a secondary antibody. Binding of a secondary antibody to the primary antibody forms a so-called sandwich with the first (or primary) antibody and the molecule of interest. The immune complexes contacted with the labeled, secondary antibody under conditions effective and for a period of time sufficient to allow the formation of secondary immune complexes. The secondary immune complexes are generally washed to remove any non-specifically bound labeled secondary antibodies, and the remaining label in the secondary immune complexes can then be detected. The additional molecule includes one of a pair of molecules or moieties that can bind to each other, such as the biotin/avidin pair. In this mode, the detecting antibody or detecting molecule includes the other member of the pair.

[0059] Other modes of indirect labeling include the detection of primary immune complexes by a two-step approach. For example, a molecule (which is referred to as a first binding agent), such as an antibody, that has binding affinity for the molecule of interest or corresponding antibody is used to form secondary immune complexes, as described above. After washing, the secondary immune complexes is contacted with another molecule (which is referred to as a second binding agent) that has binding affinity for the first binding agent, again under conditions effective and for a period of time sufficient to allow the formation of immune complexes (thus forming tertiary immune complexes). The second binding agent is linked to a detectable label or signal-generating molecule or moiety, allowing detection of the tertiary immune complexes thus formed. This system provides for signal amplification.

[0060] Immunoassays that involve the detection of a substance, such as a protein or an antibody to a specific protein, include label-free assays, protein separation methods (i.e., electrophoresis), solid support capture assays, or *in vivo* detection. Label-free assays are generally diagnostic means of determining the presence or absence of a specific protein, or an antibody to a specific protein, in a sample. Protein separation methods are additionally useful for evaluating physical properties of the protein, such as size or net charge. Capture assays are generally more useful for quantitatively evaluating the concentration of a specific protein, or antibody to a specific protein, in a sample. Finally, *in vivo* detection is useful for evaluating the spatial expression patterns of the substance, i.e., where the substance is found in a subject, tissue or cell. For example, endoscopy can be used in combination with labeled antibodies to detect the biomarkers, i.e., proteins of interest, *in vivo*.

[0061] Also contemplated are immunoassays wherein the protein or antibody specific for the protein is bound to a solid support (e.g., tube, well, bead, or cell) to capture the antibody or protein of interest, respectively, from a sample, combined with a method of detecting the protein or antibody specific for the protein on the support. Examples of such immunoassays include Radioimmunoassay (RIA), Enzyme-Linked Immunosorbent Assay (ELISA), Flow cytometry, protein array, multiplexed bead assay, and magnetic capture.

[0062] In some examples of the disclosed methods, when the level of expression of a biomarker(s) is assessed, the level is compared with the level of expression of the biomarker(s) in a reference standard. By reference standard is meant the level of expression of a particular biomarker(s) from a sample or subject lacking a cancer, at a selected stage of cancer, or in the absence of a particular variable such as a therapeutic agent. Alternatively,

the reference standard comprises a known amount of biomarker. Such a known amount correlates with an average level of subjects lacking a cancer, at a selected stage of cancer, or in the absence of a particular variable such as a therapeutic agent. A reference standard also includes the expression level of one or more biomarkers from one or more selected samples or subjects as described herein. For example, a reference standard includes an assessment of the expression level of one or more biomarkers in a sample from a subject that does not have a cancer, is at a selected stage of progression of a cancer, or has not received treatment for a cancer. Another exemplary reference standard includes an assessment of the expression level of one or more biomarkers in samples taken from multiple subjects that do not have a cancer, are at a selected stage of progression of a cancer, or have not received treatment for a cancer.

[0063] When the reference standard includes the level of expression of one or more biomarkers in a sample or subject in the absence of a therapeutic agent, the control sample or subject is optionally the same sample or subject to be tested before or after treatment with a therapeutic agent or is a selected sample or subject in the absence of the therapeutic agent. Alternatively, a reference standard is an average expression level calculated from a number of subjects without a particular cancer. A reference standard also includes a known control level or value known in the art. In one aspect of the methods disclosed herein, it is desirable to age-match a reference standard with the subject diagnosed with a cancer.

[0064] In one technique to compare protein levels of expression from two different samples (e.g., a sample from a subject diagnosed with a cancer and a reference standard), each sample is separately subjected to 2D gel electrophoresis. Alternatively, each sample is differently labeled and both samples are loaded onto the same 2D gel. See, e.g., Unlu et al. Electrophoresis, 1997;18:2071-2077, which is incorporated by reference herein for at least its teachings of methods to assess and compare levels of protein expression. The same protein or group of proteins in each sample is identified by the relative position within the pattern of proteins resolved by 2D electrophoresis. The expression levels of one or more proteins in a first sample is then compared to the expression level of the same protein(s) in the second sample, thereby allowing the identification of a protein or group of proteins that is expressed differently between the two samples (e.g., a biomarker). This comparison is made for subjects before and after they are suspected of having a cancer, before and after they begin a therapeutic regimen, and over the course of that regimen.

[0065] In another technique, the expression level of one or more proteins is in a single sample as a percentage of total expressed proteins. This assessed level of expression is compared to a preexisting reference standard, thereby allowing for the identification of proteins that are differentially expressed in the sample relative to the reference standard.

[0066] Provided herein are also methods of treating gastric cancer in the subject. The methods include contacting the biological sample with (i) an agent to determine the amount of Carcinoembryonic antigen (CEA) in the biological sample, (ii) an agent to determine the amount of carcinoembryonic antigen-related cell adhesion molecule 6 (CEACAM6) in the biological based sample, (iii) an agent to determine the amount of epithelial cell adhesion molecule (EPCAM) in the biological sample, and (iv) an agent to determine the amount of cancer antigen 72-4 (CA72-4) in the biological sample; and determining the amount of CEA, CEACAM6, EPCAM and CA72-4 in the biological sample, wherein increased levels of each of CEA, CEACAM6, EPCAM and CA72-4 in the biological sample as compared to a control indicating the subject has gastric cancer. Optionally, the methods further include determining the amount of CLDN7 and/or CLDN4 in the biological sample. The steps of contacting the biological sample and determining the amounts of the protein biomarkers can occur *in vivo* or *in vitro*. The methods include treating the subjects with gastric cancer. Suitable methods for treatment of gastric cancer are known and include, for example, radiation, surgery, and agents for treatment of the gastric cancer or a combination thereof. Optionally, one or more agents are administered to the subject with gastric cancer. Optionally, the agent is a chemotherapeutic agent. Optionally, the chemotherapeutic agent is selected from the group consisting of carboquone, nimustine, 5-fluorouracil analogs, doxorubicin, mitomycin C, aclarubicin, pirarubicin, epirubicin, irinotecan, and paclitaxel analogs. Optionally, the chemotherapeutic agent is 5-FU (fluorouracil) or its analog capecitabine, BCNU, methyl-CCNU, doxorubicin, mitomycin C, cisplatin, oxaliplatin, titanium silicate (TS1) or taxotere.

[0067] Provided are compositions comprising one or more agents to determine the amounts of CEA, CEACAM6, EPCAM, CA72-4, CLDN4, and CLDN7. Also provided are compositions comprising one or more agents for treatment of gastric cancer, e.g., chemotherapeutic agents. Optionally, the agents to determine the amounts of CEA, CEACAM6, EPCAM, CA72-4, CLDN4, and CLDN7 are therapeutic agents, e.g., therapeutic antibodies. The provided compositions are, optionally, suitable for formulation and administration *in vitro* or *in vivo*. Suitable carriers and excipients and their formulations are described in Remington: The Science and Practice of Pharmacy, 21st Edition, David B. Troy,

ed., Lippicott Williams & Wilkins (2005). By pharmaceutically acceptable carrier is meant a material that is not biologically or otherwise undesirable, i.e., the material is administered to a subject without causing undesirable biological effects or interacting in a deleterious manner with the other components of the pharmaceutical composition in which it is contained. If administered to a subject, the carrier is optionally selected to minimize degradation of the active ingredient and to minimize adverse side effects in the subject.

[0068] The compositions disclosed herein can be administered by any means known in the art. For example, compositions may include administration to a subject intravenously, intradermally, intraarterially, intraperitoneally, intralesionally, intracranially, intraarticularly, intraprostatically, intrapleurally, intratracheally, intranasally, intravitreally, intravaginally, intrarectally, topically, intratumorally, intramuscularly, intrathecally, subcutaneously, subconjunctival, intravesicularly, mucosally, intrapericardially, intraumbilically, intraocularly, orally, locally, by inhalation, by injection, by infusion, by continuous infusion, by localized perfusion, via a catheter, via a lavage, in a creme, or in a lipid composition. Administration can be local, e.g., at or near the site of a tumor, or systemic.

[0069] The compositions for administration will commonly comprise an agent as described herein dissolved in a pharmaceutically acceptable carrier, preferably an aqueous carrier. A variety of aqueous carriers can be used, e.g., buffered saline and the like. These solutions are sterile and generally free of undesirable matter. These compositions may be sterilized by conventional, well known sterilization techniques. The compositions may contain pharmaceutically acceptable auxiliary substances as required to approximate physiological conditions such as pH adjusting and buffering agents, toxicity adjusting agents and the like, for example, sodium acetate, sodium chloride, potassium chloride, calcium chloride, sodium lactate and the like. The concentration of active agent in these formulations can vary widely, and will be selected primarily based on fluid volumes, viscosities, body weight and the like in accordance with the particular mode of administration selected and the subject's needs.

[0070] The formulations of compounds can be presented in unit-dose or multi-dose sealed containers, such as ampules and vials. Thus, the composition can be in unit dosage form. In such form the preparation is subdivided into unit doses containing appropriate quantities of the active component. Thus, the compositions can be administered in a variety of unit dosage forms depending upon the method of administration. For example, unit dosage forms suitable

for oral administration include, but are not limited to, powder, tablets, pills, capsules and lozenges.

[0071] Compositions can be formulated to provide quick, sustained or delayed release after administration by employing procedures known in the art. Certain carriers may be more preferable depending upon, for instance, the route of administration and concentration of composition being administered. Suitable formulations for use in the provided compositions can be found in Remington: The Science and Practice of Pharmacy, 21st Edition, David B. Troy, ed., Lippicott Williams & Wilkins (2005).

[0072] Combinations of agents, e.g., detection agents or therapeutic agents, may be administered either concomitantly (e.g., as an admixture), separately but simultaneously (e.g., via separate intravenous lines into the same subject), or sequentially (e.g., one of the compounds or agents is given first followed by the second). Thus, the term combination is used to refer to either concomitant, simultaneous, or sequential administration of two or more agents. The course of treatment is best determined on an individual basis depending on the particular characteristics of the subject and the type of treatment selected. The treatment, such as those disclosed herein, can be administered to the subject on a daily, twice daily, bi-weekly, monthly or any applicable basis that is therapeutically effective. The treatment can be administered alone or in combination with any other treatment disclosed herein or known in the art. The additional treatment can be administered simultaneously with the first treatment, at a different time, or on an entirely different therapeutic schedule (e.g., the first treatment can be daily, while the additional treatment is weekly).

[0073] Provided herein are kits for detecting biomarkers of gastric cancer, optionally, along with instructions for use. The kit can be a kit for diagnosing or prognosing a gastric cancer, or for monitoring the progression of disease or the efficacy of treatment. Optionally, the kit comprises one or more agents for determining the amount of a biomarker of gastric cancer in a suitable container. Optionally, the kit comprises an agent that selectively binds the biomarker of gastric cancer. Thus, provided is a kit comprising an agent selectively binding CEA, an agent selectively binding CEACAM6, an agent selectively binding EPCAM and an agent selectively binding CA72-4. Optionally, the kit further includes an agent selectively binding CLDN7. Optionally, the kit further includes an agent selectively binding CLDN4. Optionally, the agents are antibodies. Optionally, the antibodies are labeled antibodies. Optionally, the kit includes in a separate container labels for the antibodies. Kits can include

a detecting reagent or a detecting apparatus capable of indicating binding of the marker protein binding agent to the substance. The kit can further include assay containers (tubes), buffers, or enzymes necessary for carrying out the detection assay. Optionally, the kit further includes a sample collection device for collecting a sample from a subject. Optionally, the human subject has gastric cancer. The kit will generally include at least one vial, test tube, flask, bottle, syringe or other container means, into which the testing agent, can be suitably reacted or aliquoted. Kits can also include components for comparing results such as a suitable control sample, for example a positive and/or negative control. The kit can also include a collection device for collecting and/ or holding the sample from the subject. The collection device can include a sterile swab or needle (for collecting blood), and/or a sterile tube (*e.g.*, for holding the swab or a bodily fluid sample).

[0074] Disclosed are materials, compositions, and components that can be used for, can be used in conjunction with, can be used in preparation for, or are products of the disclosed methods and compositions. These and other materials are disclosed herein, and it is understood that when combinations, subsets, interactions, groups, etc. of these materials are disclosed that while specific reference of each various individual and collective combinations and permutations of these compounds may not be explicitly disclosed, each is specifically contemplated and described herein. For example, if a method is disclosed and discussed and a number of modifications that can be made to a number of molecules including the method are discussed, each and every combination and permutation of the method, and the modifications that are possible are specifically contemplated unless specifically indicated to the contrary. Likewise, any subset or combination of these is also specifically contemplated and disclosed. This concept applies to all aspects of this disclosure including, but not limited to, steps in methods using the disclosed compositions. Thus, if there are a variety of additional steps that can be performed, it is understood that each of these additional steps can be performed with any specific method steps or combination of method steps of the disclosed methods, and that each such combination or subset of combinations is specifically contemplated and should be considered disclosed.

[0075] Publications cited herein and the material for which they are cited are hereby specifically incorporated by reference in their entireties.

[0076] A number of embodiments have been described. Nevertheless, it will be understood that various modifications may be made. Accordingly, other embodiments are within the scope of the claims below.

Example

Example 1. Biomarkers for Diagnosis of Gastric Cancer.

[0077] Matched tumor and normal tissue samples from 2 RNA-seq datasets (TCGA and SGCC) were analyzed to identify the top 25 putative membrane protein mRNA's overexpressed in tumors when compared to normal tissue. A similar analysis was also performed using 3 gene expression signature databases of gastric cancers and normal stomach tissue (Table 2). For TCGA RNA Seq, the dataset contains 305 tumor samples and 33 normal samples, of which 29 have corresponding matched tumor samples. For SGCC RNA Seq, the dataset contains 15 matched tumor versus normal samples. For SGCC Microarray, the dataset contains 185 tumor and 89 normal samples. For GSE2685 Microarray, the dataset contains 22 primary human advanced gastric cancer tissues and 8 noncancerous gastric tissues. For the GSE13861 Microarray, the dataset contains 65 primary gastric adenocarcinoma, 6 gastrointestinal stromal tumor (GIST) and 19 surrounding normal fresh frozen tissue samples. The hits from all the analyses were then scored on whether they appeared in RNA seq databases, gene expression databases or both, and a final rank sheet was prepared.

[0078] Antibodies to the top hits in the rank list were then obtained from commercial sources. These were chosen on the basis of their previous reports in immunohistochemistry. These antibodies were then tested in sample of gastric cancer to show immunoreactivity and optimize staining conditions. Finally, they were used to stain a single tissue microarray of 30 gastric cancer cases with normal matched control tissue from the stomach. The Vectra 2 automated image analysis platform was used to image and quantify intensity of staining within tumor tissue.

[0079] The final candidates were then chosen on the basis of two factors: a) extent of signal in tumor tissue when compared to normal tissue (quantitated SNR- signal to noise ratio) and b) pattern and intensity of staining in normal tissue. Antibodies with a high SNR and low staining in normal tissue were selected. See Figure 1 for a schematic workflow.

[0080] The results are shown in Figure 3. Specifically, Figure 3A shows a layout of the HStm-Ade090PG-01 gastric adenocarcinoma tissue microarray (TMA) of 30 cases containing tumor (T) and normal adjacent tissue (NAT). Figure 3B is a graph showing quantitation of all candidate biomarkers in matched normal and advanced gastric adenocarcinoma tissue (compare NAT vs T for each marker). Figure 3C is a graph of ROC curves of all candidate biomarkers to determine whether a subject has gastric cancer (based on quantitative results from Figure 3B).

[0081] An Area Under the Curve (AUC) analysis was performed for all our top hits. See Table 1. In order of preference, the top hits are 1. CA72-4 (0.975), 2. CEA (0.967), 3. CEACAM6 (0.966), 4. EpCAM (0.921), 5. CLDN4 (0.907), and 6. CLDN7 (0.858). This calculation of sensitivity and specificity is based on performance of these antibodies in a 30 sample data set by immunohistochemistry.

Table 1. Area Under the Curve

Test Result Variable	Area	Std. Error ^a	Asymptotic Sig. ^b	Asymptotic 95% Confidence Interval	
				Lower Bound	Upper Bound
CEA	.967	.021	.000	.927	1.008
CEACAM6	.966	.022	.000	.923	1.008
CLDN7	.858	.058	.000	.745	.972
EpCAM	.921	.050	.000	.824	1.018
CA724	.975	.018	.000	.939	1.010
CA199	.762	.071	.001	.623	.900
CD98	.569	.092	.368	.389	.750
CLDN3	.819	.061	.000	.700	.939
CLDN4	.907	.038	.000	.832	.982
GDF15	.703	.069	.008	.568	.837
ILDR1	.516	.077	.840	.365	.666
MESO	.692	.077	.012	.541	.844
MMP3	.376	.071	.108	.237	.515
PLA2G7	.779	.071	.000	.640	.918

SPP1	.264	.073	.002	.121	.408
------	------	------	------	------	------

The test result variable(s): CEA, CEACAM6, CLDN7, EpCAM, CD98, CLDN3, CLDN4, GDF15, ILDR1, MESO, MMP3, PLA2G7, SPP1 have at least one tie between positive actual state group and the negative actual state group. Statistics may be biased. ^aUnder the nonparametric assumption. ^bNull hypothesis: true area = 0.5.

[0082] Using a black box model, an ROC analysis was performed for using a combination of CEA, CEACAM6, CA72-4, and EpCAM. The results are shown in Figure 3D. The area under the curve (AUC) for the four marker black box model is 0.990. The following equation was used: probability = $1/(1+\exp(-z))$, where $z = -20 + 23 * \text{CEA} - 2 * \text{CEACAM6} + 63 * \text{CA724} + 45 * \text{EPCAM}$. Optimal cutoff at $p=0.5$.

Table 2. p0.5 * group-Crosstabulation

			group		Total
			.00	1.00	
p0.5	.00	Count	47	2	49
		% within p0.5	95.9%	4.1%	100.0%
		% within group	97.9%	8.0%	67.1%
	1.00	Count	1	23	24
		% within p0.5	4.2%	95.8%	100.0%
		% within group	2.1%	92.0%	32.9%
Total	Count		48	25	73
	% within p0.5		65.8%	34.2%	100.0%
	% within group		100.0%	100.0%	100.0%

Sen=92, spec=97.9, ppv=98.5, npv=95.9

[0083] Using a risk score model, an analysis was performed for using a combination of CEA, CEACAM6, CA72-4, and EpCAM to determine whether a subject has gastric cancer. Taking B as weight, weights are derived proportionately from the “B” estimate in the 4-categorical logistic model as follows: score = $1 * \text{CEA} \geq 0.188 + 1 * \text{CEACAM6} \geq 0.183 + 2 * \text{CA724} \geq 0.190 + 3 * \text{EpCAM} \geq 0.170$. See Table 3.

Table 3.

parameter	points
CEA ≥ 0.188	1
CEACAM6 ≥ 0.183	1
CA724 ≥ 0.190	2
EpCAM ≥ 0.170	3

[0084] The AUC for the risk score model was 0.999. See Tables 4 and 5.

Table 4.

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
No	48	.0833	.34723	.05012	-.0175	.1842	.00	2.00
Yes	25	5.8000	1.80278	.36056	5.0559	6.5441	2.00	7.00
Total	73	2.0411	2.93654	.34370	1.3559	2.7262	.00	7.00

P<0.001

[0085] The cutoff score = 2.

Table 5. score_2 * group-Crosstabulation

			group		Total
			.00	1.00	
score_2	.00	Count	47	0	47
		% within score_2	100.0%	0.0%	100.0%
		% within group	97.9%	0.0%	64.4%
	1.00	Count	1	25	26
		% within score_2	3.8%	96.2%	100.0%
		% within group	2.1%	100.0%	35.6%
Total	Count	48	25	73	
	% within score_2	65.8%	34.2%	100.0%	
	% within group	100.0%	100.0%	100.0%	

Sen=100, spec=97.9, ppv=96.2, npv=100

Example 2. Detection in Early Gastric Cancer

[0086] Sixty-five (65) cases of early gastric cancer were stained with adjacent normal tissue for CEACAM6, CEA, CA72-4, and EpCAM using immunofluorescence in fixed tissue. The fold change in staining between normal and cancer tissue was calculated for all cases, using antibodies binding to CEACAM6, CEA, CA72-4, and EpCAM. Staining was done in fixed tissue (immunofluorescence). Specific combinations of reagents to these surface epitopes may be synergistic towards the discrimination of gastric cancer from normal adjacent tissue.

[0087] Tissue microarrays of 65 early gastric adenocarcinoma (US Biomax, HStm-Ade150CS1-01; T1N0M0=26, T1N1M0=4, T2N0M0=35 and matched normal adjacent tissue, 2 cores/case) were immunostained with antibodies against CEA (II-7, Dako),

CEACAM6 (9A6, Abcam), EpCAM (C-10, Santa Cruz) and CA72-4 (CC49, Santa Cruz) following a multiplex fluorescent immunohistochemistry protocol according to manufactures instruction (OPAL 5-color fIHC Kit, PerkinElmer).

[0088] Acquisition and image analysis was done with the Vectra 2 multispectral automated imaging system (PerkinElmer) and in Form 2.0 image analysis software. Briefly, multispectral images of stained tissue were unmixed using predefined spectra of each chromogen/fluorophore used, to obtain pure images for each marker without interference from auto-fluorescence (Fig. 2A). Tissue segmentation was next employed to recognize tissue architecture to define normal gastric tissue regions from tumor regions (Fig. 2B). Subsequently, cell segmentation was used to define cell boundaries on the basis of nuclear counterstaining, aided by the membrane staining of markers of interest, where possible (Fig. 2C). Intensities of each marker was quantitated in the membrane-cytoplasm region on a cellular basis, and the mean value from all segmented cells was used as a descriptive value for each normal or tumor region.

[0089] Biomarkers showing increased staining in cancers in comparison to normal controls include CEA, CEACAM6, EpCAM, CA72-2, CLDN-7, and CLDN-4. However, the extent was predictably lower than that seen with the advanced cancers. In early gastric cancer, certain combinations: CEACAM6 and EpCAM, CEA and EpCAM and CEA and CA72-4 show non-overlapping staining (i.e., the highest potential for a combinatorial application to improve sensitivity of detection). See Figure 4A, 4B, and 4C. This is more clearly evident from the early gastric cancer dataset due to the lower staining intensity. Therefore, the data indicate these markers can be used to improve detection of cancer including early stage gastric cancer.

[0090] To further investigate these markers, five (5) gastric cancer patient derived xenografts (PDXs) were harvested and placed in buffered saline. They were treated, without fixation, with a solution of fluorescent dye alone (control), or fluorescently labeled primary antibodies to CEA, CEACAM6, CLDN4, CLDN7, EpCAM, and TAG-72 (individually and in combination). Fluorescence and brightfield images were obtained with an inverted epifluorescence microscope. The mixture of 6 antibodies (Mix) yielded a significantly increased signal in this data set, suggesting an additive value of a combination assay in the context of fresh cancer tissue (as may be expected during endoscopy). See Figure 5.

WHAT IS CLAIMED IS:

1. A method of determining whether a subject has gastric cancer, the method comprising the steps of
 - (a) contacting a biological sample with (i) one or more agents to determine the amount of Carcinoembryonic antigen (CEA) in the biological sample, (ii) one or more agents to determine the amount of carcinoembryonic antigen-related cell adhesion molecule 6 (CEACAM6) in the biological based sample, (iii) one or more agents to determine the amount of epithelial cell adhesion molecule (EPCAM) in the biological sample, and (iv) one or more agents to determine the amount of cancer antigen 72-4 (CA72-4) in the biological sample; and
 - (b) determining the amount of CEA, CEACAM6, EPCAM and CA72-4 in the biological sample, wherein increased levels of each of CEA, CEACAM6, EPCAM and CA72-4 in the biological sample as compared to a control indicates the subject has gastric cancer.
2. The method of claim 1, wherein the method further comprises contacting the biological sample with an agent to determine the amount of claudin-7 (CLDN7) in the biological sample.
3. The method of claim 1 or 2, wherein the method further comprises contacting the biological sample with an agent to determine the amount of claudin-4 (CLDN4) in the biological sample.
4. The method of any one of claims 1-3, wherein the agent is a labeled antibody.
5. The method of claim 3, wherein the method comprises contacting the sample with a labeled antibody that binds CEA, a labeled antibody that binds CEACAM6, a labeled antibody that binds EPCAM, and a labeled antibody that binds CA72-4.
6. The method of claim 3, wherein the method comprises contacting the sample with a labeled antibody that binds CEA, a labeled antibody that binds CEACAM6, a labeled antibody that binds EPCAM, a labeled antibody that binds CA72-4, a labeled antibody that binds CLDN7 and a labeled antibody that binds CLDN4.

7. The method of claim 1, wherein said method is an immunofluorescence based assay.
8. The method of claim 1, wherein said immunofluorescence based assay is flow cytometry.
9. The method of claim 1, wherein said immunofluorescence based assay is fluorescence activated cell sorting (FACS).
10. The method of claim 1, wherein said immunofluorescence based assay is an enzyme-linked immune assay (ELISA).
11. The method of claim 1, wherein the subject is a non-human animal.
12. The method of claim 1, wherein the subject is a human.
13. The method of claim 1, wherein the biological sample is a tissue sample.
14. The method of claim 1, wherein said sample is selected from the group consisting of: mucosal sample, blood, blood plasma or serum.
15. The method of claim 1, wherein the step of contacting the biological sample occurs *in vivo*.
16. The method of claim 15, wherein the contacting comprises administering to the subject the agents that bind CEA, CEACAM6, EPCAM and CA72-4.
17. The method of claim 16, further comprising administering to the subject an agent that binds CLDN7 and/or an agent that binds CLDN4.
18. The method of claim 15 or 16, wherein the agents are administered simultaneously or sequentially.
19. The method of claim 15 or 16, wherein the agents are labeled antibodies.

20. The method of claim 18, wherein the label is a fluorescent molecule.
21. The method of claim 19, wherein the antibodies are detected visually using an endoscope.
22. The method of any one of claims 15-21, wherein the biological sample is the stomach or digestive track.
23. The method of claim 1, wherein the step of contacting the biological sample occurs *in vitro*.
24. A method of treating gastric cancer in a subject, the method comprising the steps of
(a) contacting the biological sample with (i) an agent to determine the amount of Carcinoembryonic antigen (CEA) in the biological sample, (ii) an agent to determine the amount of carcinoembryonic antigen-related cell adhesion molecule 6 (CEACAM6) in the biological based sample, (iii) an agent to determine the amount of epithelial cell adhesion molecule (EPCAM) in the biological sample, and (iv) an agent to determine the amount of cancer antigen 72-4 (CA72-4) in the biological sample; and
(b) determining the amount of CEA, CEACAM6, EPCAM and CA72-4 in the biological sample, wherein increased levels of each of CEA, CEACAM6, EPCAM and CA72-4 in the biological sample as compared to a control indicating the subject has gastric cancer;
(c) treating the subjects with gastric cancer with radiation, surgery, one or more agents for treatment of the gastric cancer or a combination thereof.
25. The method of claim 24, wherein the step of contacting the biological sample occurs *in vivo*.
26. The method of claim 24, wherein the step of contacting the biological sample occurs *in vitro*.
27. The method of claim 24, wherein the agent is a chemotherapeutic agent.

28. The method of claim 27, wherein the chemotherapeutic agent is selected from the group consisting of carboquone, nimustine, 5-fluorouracil analogs, doxorubicin, mitomycin C, aclarubicin, pirarubicin, epirubicin, irinotecan, and paclitaxel analogs.
29. A kit comprising an agent selectively binding CEA, an agent selectively binding CEACAM6, an agent selectively binding EPCAM and an agent selectively binding CA72-4.
30. The kit of claim 29, further comprising an agent selectively binding CLDN7.
31. The kit of claim 29 or 30, further comprising an agent selectively binding CLDN4.
32. The kit of any one of claims 29-31, wherein the agent is an antibody.
33. The kit of claim 32, further comprising one or more labels for the antibody.
34. The kit of claim 33, wherein the antibodies and labels are in separate containers.

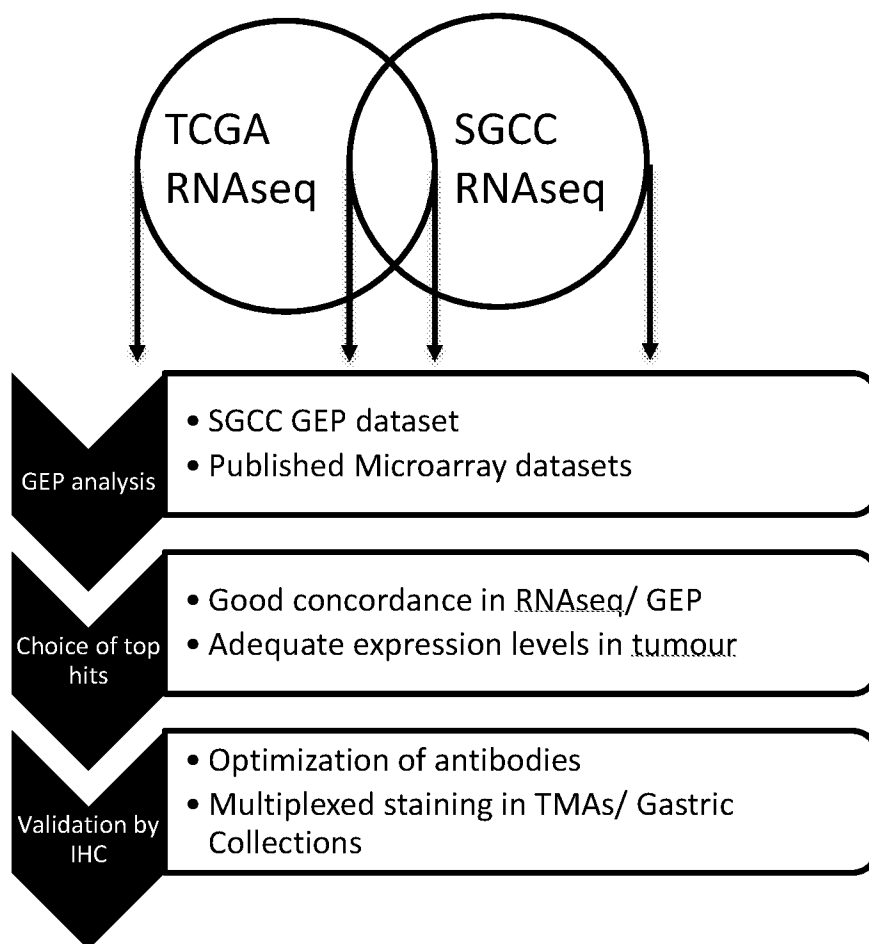


FIG. 1

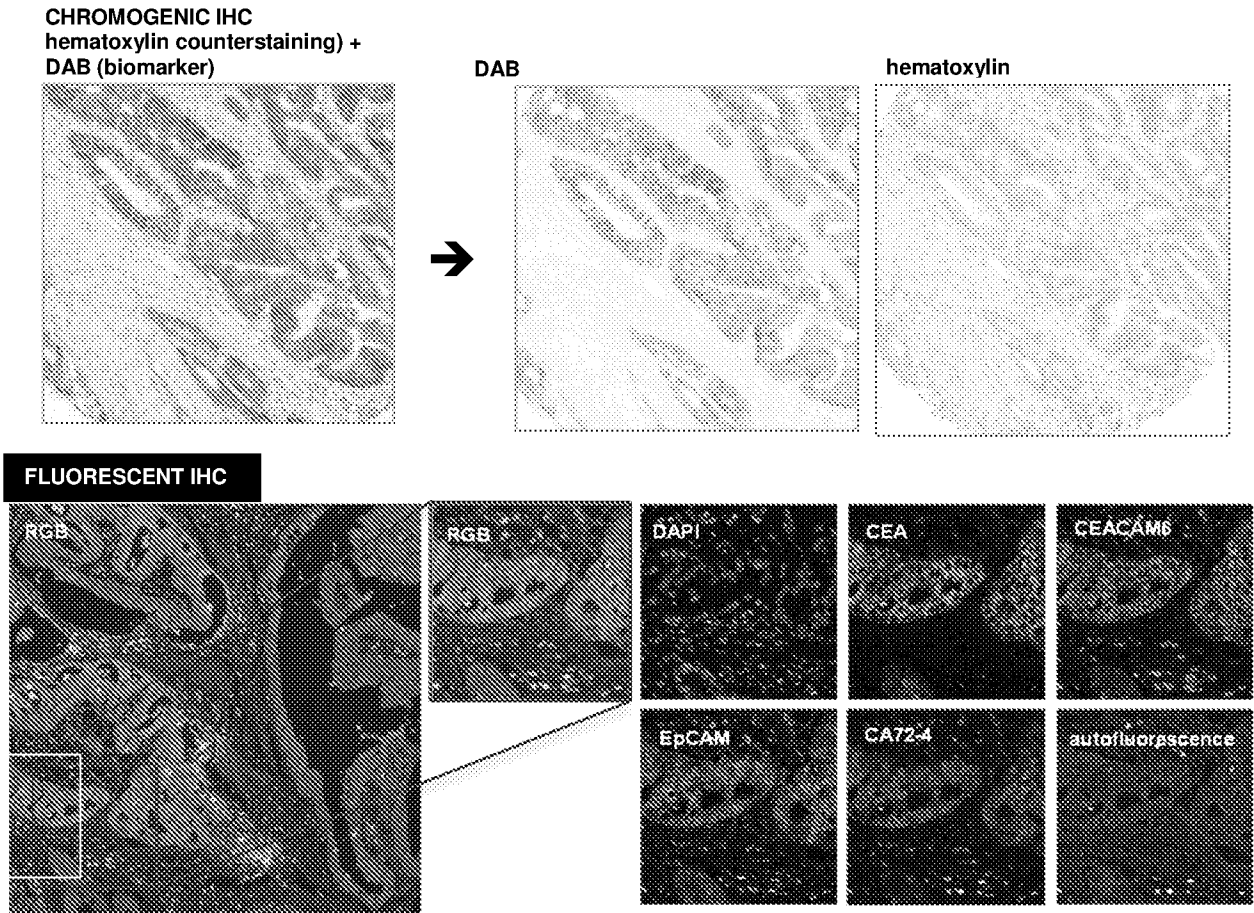


FIG. 2A

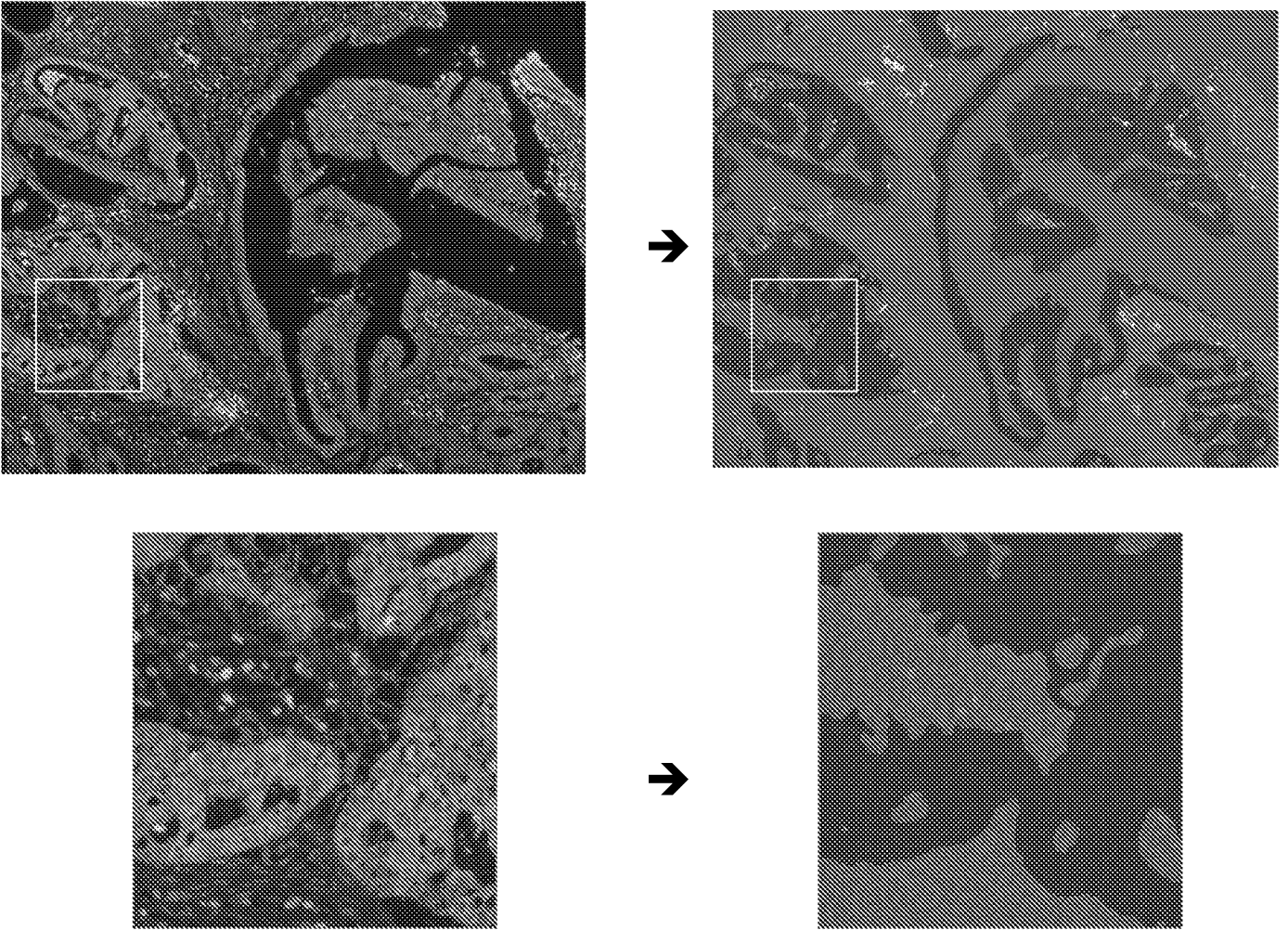


FIG. 2B

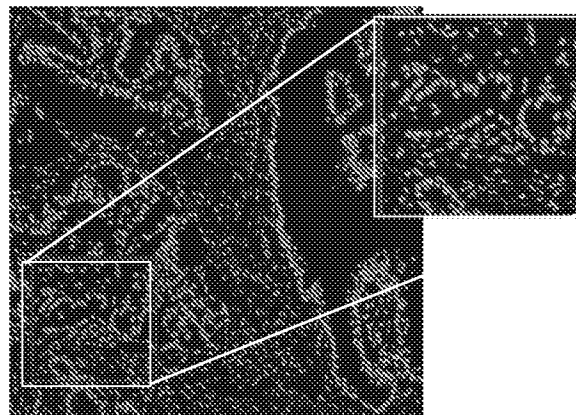


FIG. 2C

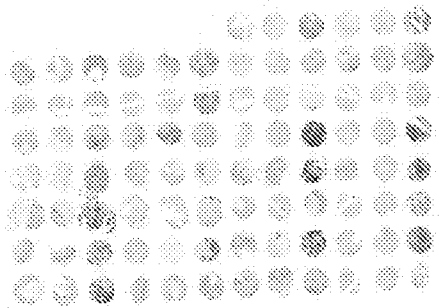


FIG. 3A

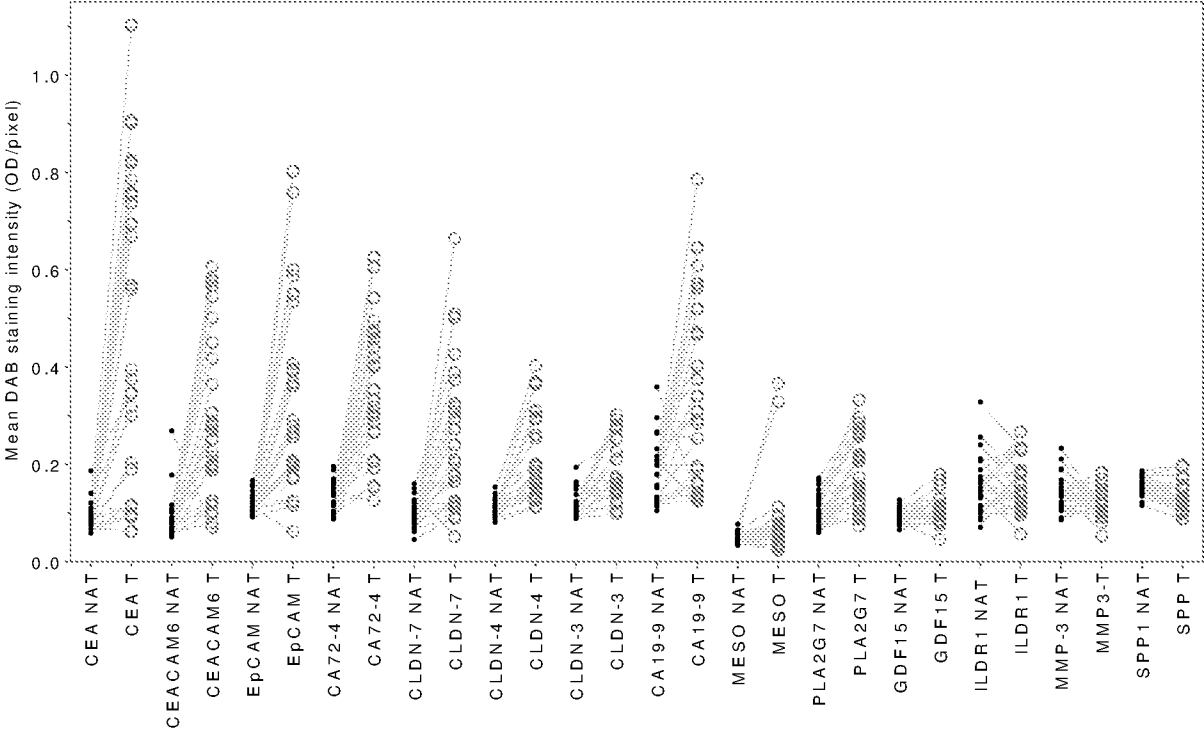


FIG. 3B

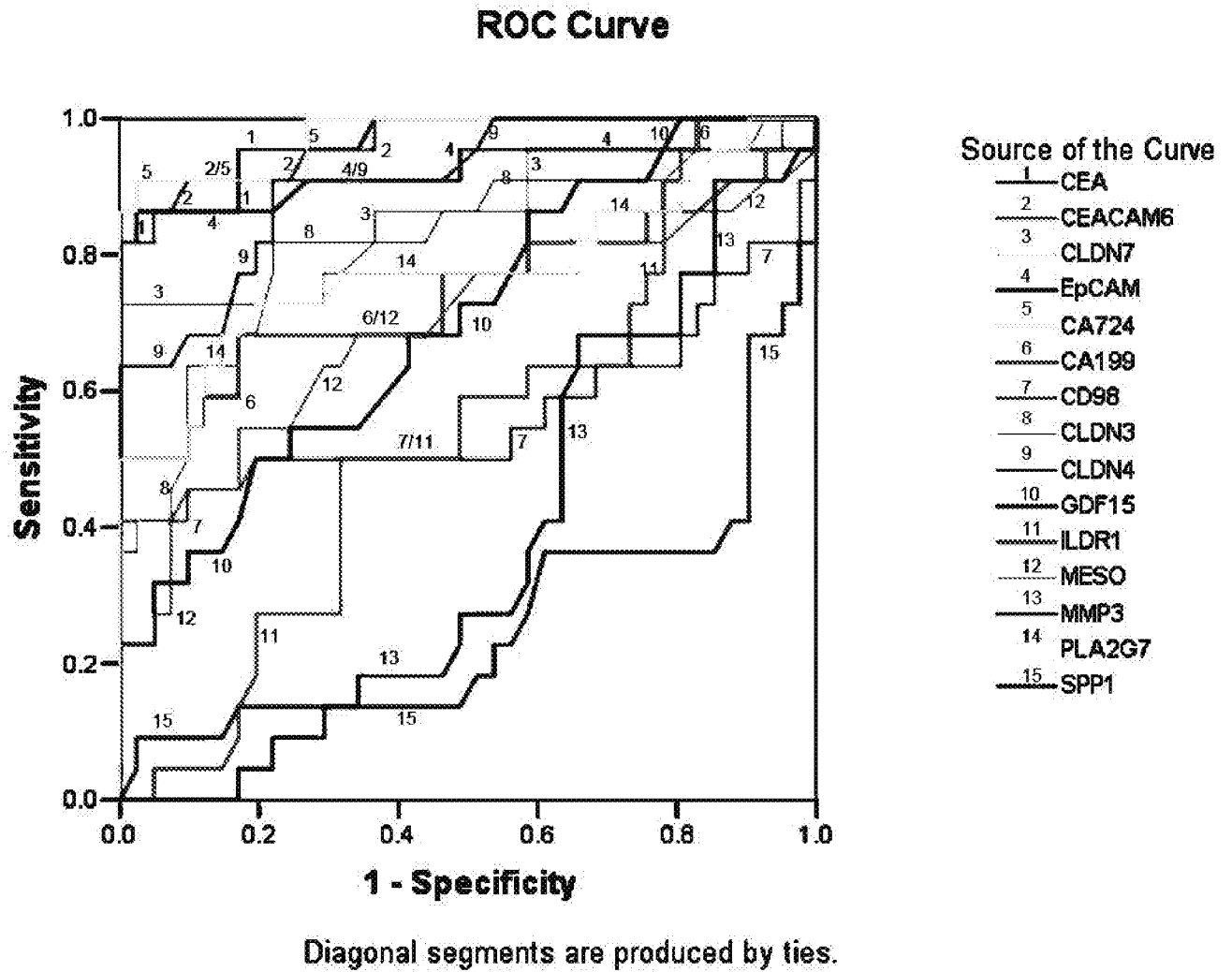


FIG. 3C

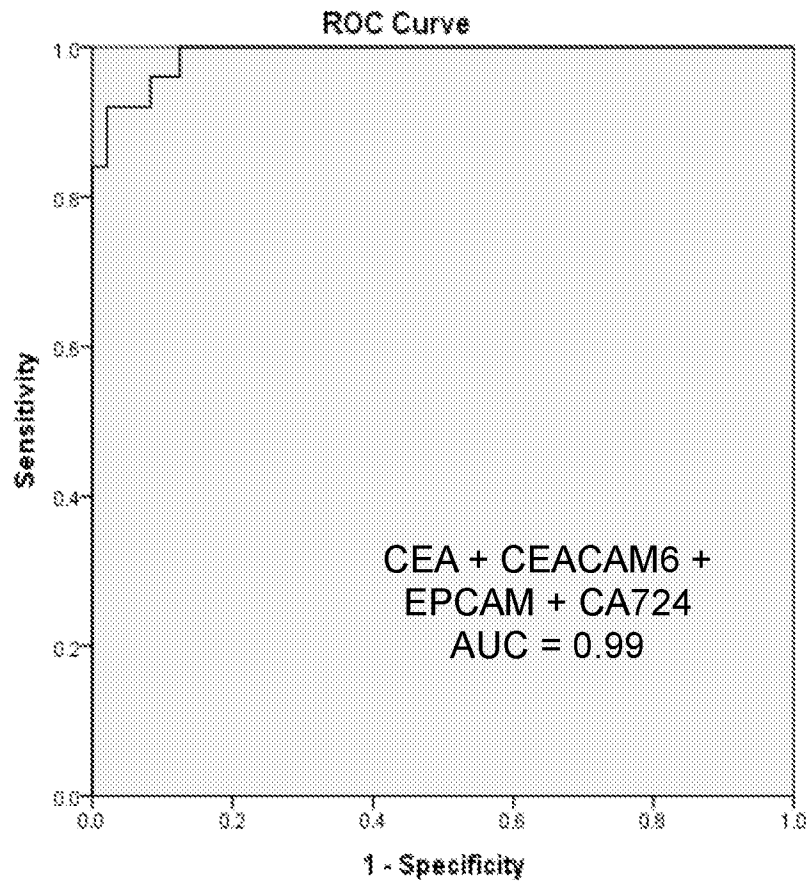


FIG. 3D

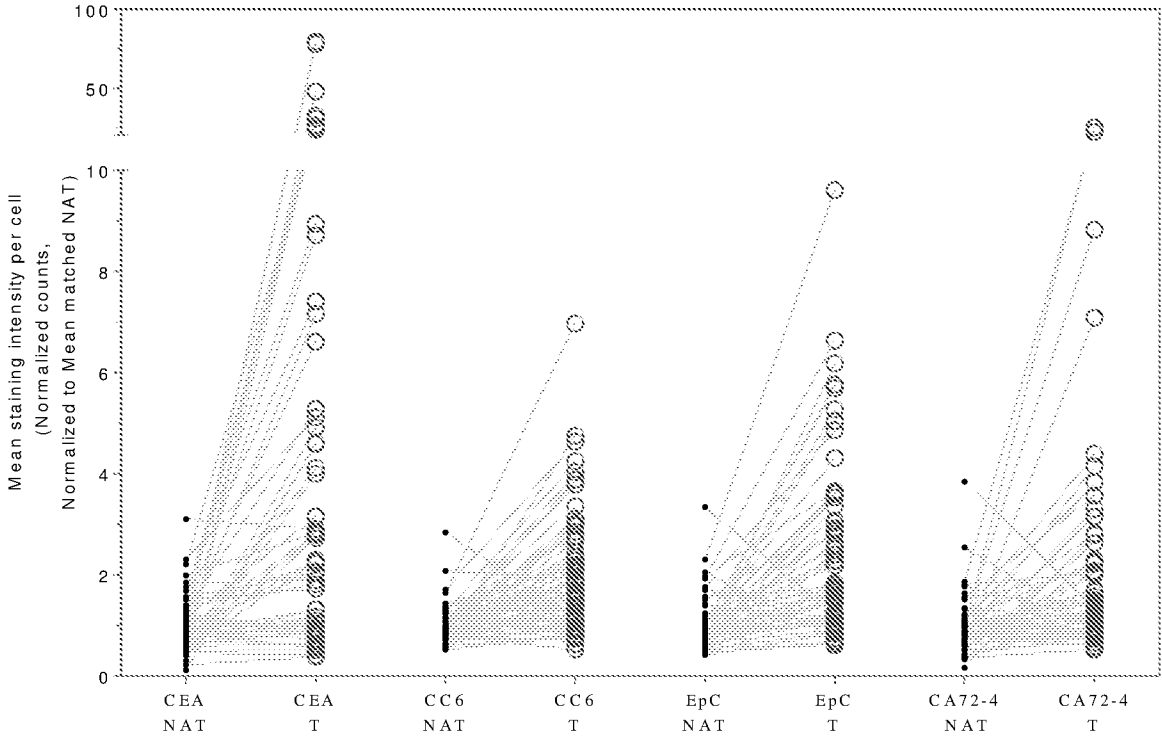


FIG. 4A

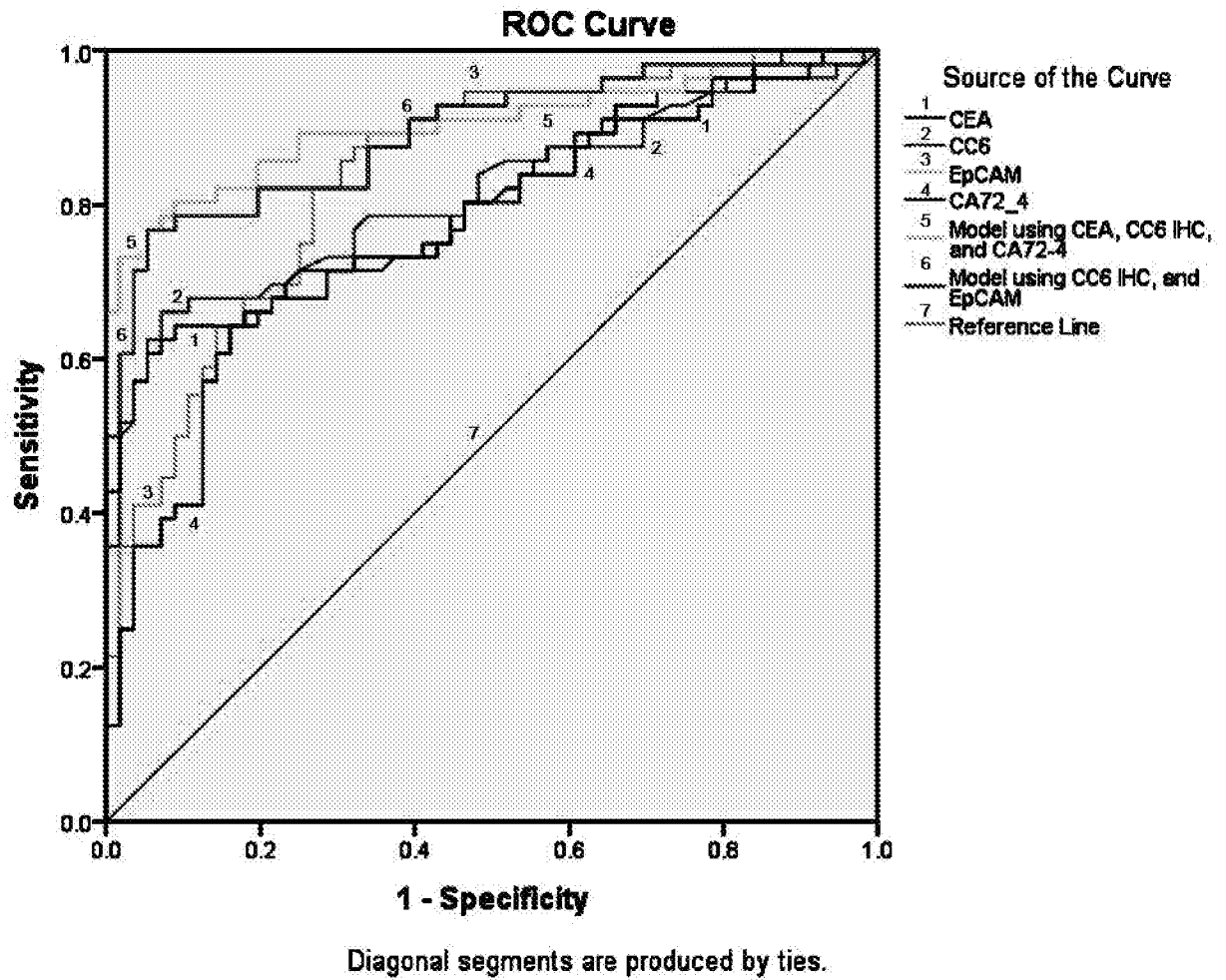


FIG. 4B

Area Under the Curve					
Test Result Variable(s)	Area	Std. Error ^a	Asymptotic Sig. ^b	Asymptotic 95% Confidence Interval	
				Lower Bound	Upper Bound
CEA	.799	.043	.000	.715	.882
CC6	.813	.042	.000	.731	.894
EpCAM	.842	.037	.000	.770	.914
CA72_4	.773	.044	.000	.687	.859
Model using CEA, CC6 IHC, and CA72-4	.902	.031	.000	.840	.963
Model using CC6 IHC, and EpCAM	.892	.031	.000	.831	.953

The test result variable(s): CEA, CC6, CA72_4 has at least one tie between the positive actual state group and the negative actual state group. Statistics may be biased.

a. Under the nonparametric assumption

b. Null hypothesis: true area = 0.5

FIG. 4C

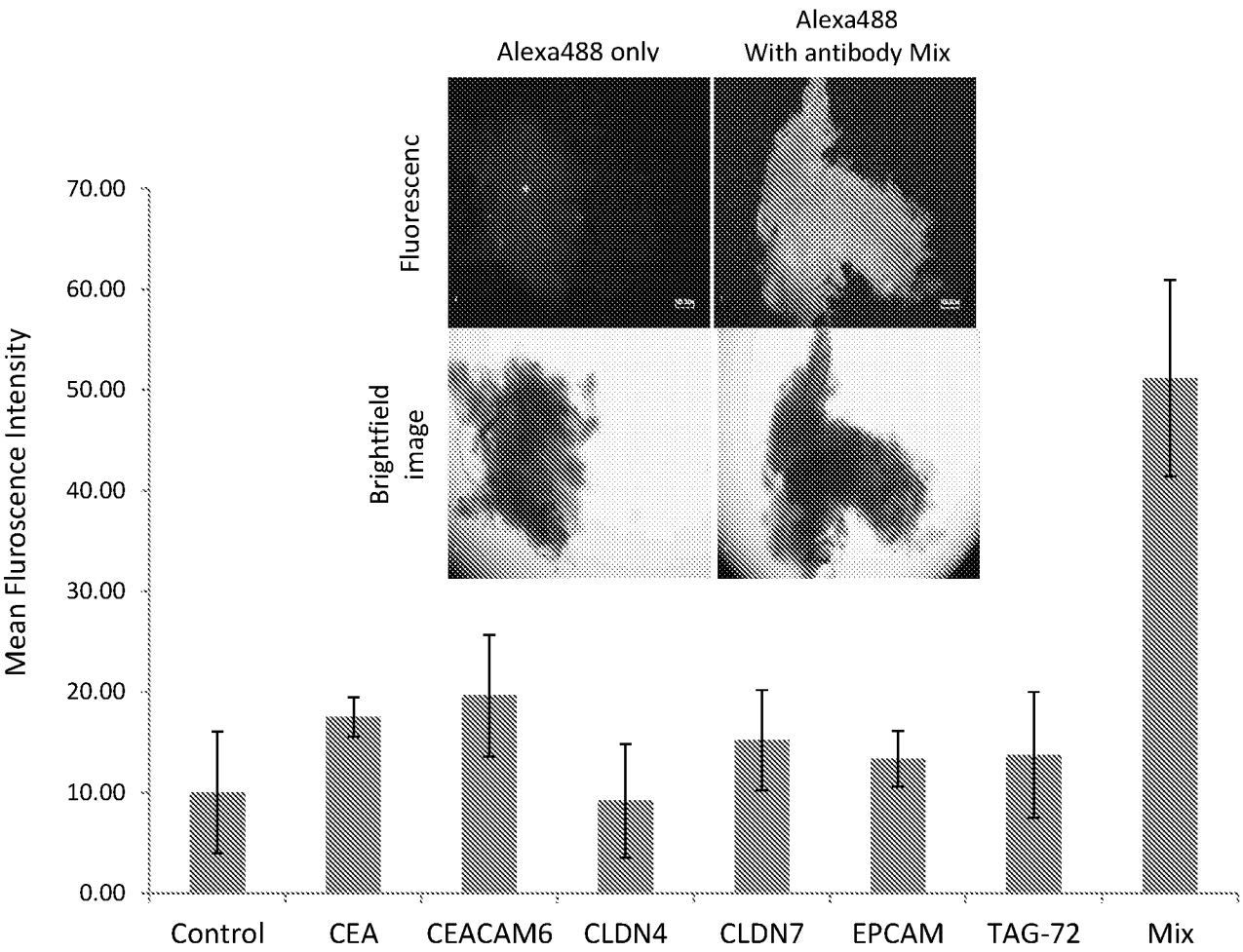


FIG. 5

INTERNATIONAL SEARCH REPORT

International application No.
PCT/SG2016/050337

A. CLASSIFICATION OF SUBJECT MATTER

G01N 33/68 (2006.01) G01N 33/574 (2006.01)

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

TXPEA, TXPEB, TXPEC, TXPEE, TXPEF, TXPEI, TXPEP, TXPEPEA, TXPES, TXPUS01A, TXPUSE1A, TXPUSEA, TXPUSEB, TXPW0EA, WPIAP, EPODOC, MEDLINE, CAPLUS, BIOSIS, EMBASE, Google Scholar: CEA, CEACAM6, EPCAM, CA72-4, claudin-7, diagnosis, gastric cancer, kit and like terms

Inventor/applicant search in Google Scholar, Espacenet, PAMS NOSE and INTESS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
	Documents are listed in the continuation of Box C	



Further documents are listed in the continuation of Box C



See patent family annex

* "A"	Special categories of cited documents: document defining the general state of the art which is not considered to be of particular relevance	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"E"	earlier application or patent but published on or after the international filing date	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"O"	document referring to an oral disclosure, use, exhibition or other means	"&"	document member of the same patent family
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search
28 September 2016

Date of mailing of the international search report
28 September 2016

Name and mailing address of the ISA/AU

AUSTRALIAN PATENT OFFICE
PO BOX 200, WODEN ACT 2606, AUSTRALIA
Email address: pct@ipaustalia.gov.au

Authorised officer

Jonathan Konrad
AUSTRALIAN PATENT OFFICE
(ISO 9001 Quality Certified Service)
Telephone No. +61 2 6283 2601

INTERNATIONAL SEARCH REPORT		International application No.
C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		PCT/SG2016/050337
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	Ordóñez, N. G., 'Broad-spectrum immunohistochemical epithelial markers: a review', Human Pathology, 2013, Vol. 44, pages 1195-1215 see page 1203 right column lines 1-3 and second paragraph final sentence, page 1205 left column third paragraph and right column third paragraph, page 1208 left column second paragraph and Tables 6, 7, 9 and 10	1-34
Y	Tan, J. et al., 'Advances in detection technologies of early cancer', Cell Communications, 2015 (18 March), Vol. 2, No. 2, pages 36-45 see Table 2 and page 41 second paragraph	1-34
Y	Shimada, S. et al., 'Comparison of the interferon-gamma-mediated regulation of tumor-associated antigens expressed by human gastric carcinoma cells', In Vivo, 1993, Vol. 7, No. 1, pages 1-8 see the abstract, page 1 right column first paragraph final sentence and page 3 left column	1-34
Y	Yasui, W. et al., 'Search for new biomarkers of gastric cancer through serial analysis of gene expression and its clinical implications', Cancer Science, 2004, Vol. 95, No. 5, pages 385-392 see the abstract and Table 1	1-34
Y	Zhang, Y. et al., 'CEACAM6 promotes tumor migration, invasion, and metastasis in gastric cancer', Acta Biochim Biophys Sin, 2014, Vol. 46, pages 283-290 see the abstract and Figure 1	1-34
Y	Chan, C. H. F. and Stanners, C. P., 'Recent advances in the tumour biology of the GPI-anchored carcinoembryonic antigen family members <i>CEACAM5</i> and <i>CEACAM6</i> ', Current Oncology, 2007, Vol. 14, No. 2, pages 70-73 see page 70 right column second paragraph	1-34
Y	Hewitt, K. J. et al., 'The claudin gene family: expression in normal and neoplastic tissues', BMC Cancer, 2006, Vol. 6, No. 186 see the abstract, page 2 left column second paragraph and page 3 right column first paragraph	2-3, 17, 30-31
A	US 2012/0028269 A1 (Lim et al.) 02 February 2012	
A	WO 2011/010969 A1 (NATIONAL UNIVERSITY OF SINGAPORE) 27 January 2011	
P,A	Roy, R. K. et al., 'CEACAM6 is upregulated by <i>Helicobacter pylori</i> CagA and is a biomarker for early gastric cancer', Oncotarget, advance publication 11 July 2016	
P,A	Jeyasekharan, A. D. et al., 'CEACAM6 is upregulated by the H pylori Cag-A oncoprotein, and is a candidate biomarker for the diagnosis of early gastric cancer', Cancer Research, 2015 (1 August), Vol. 75, No. 15 Supplement, Abstract 542, Proceedings: AACR 106th Annual Meeting 2015; April 18-22, 2015; Philadelphia, PA	
Form PCT/ISA/210 (fifth sheet) (July 2009)		

INTERNATIONAL SEARCH REPORT Information on patent family members		International application No. PCT/SG2016/050337	
This Annex lists known patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.			
Patent Document/s Cited in Search Report		Patent Family Member/s	
Publication Number	Publication Date	Publication Number	Publication Date
US 2012/0028269 A1 02 February 2012 US 2012028269 A1 02 Feb 2012 WO 2011/010969 A1 27 January 2011 WO 2011010969 A1 27 Jan 2011 CN 102687011 A 19 Sep 2012 CN 102687011 B 31 Dec 2014 EP 2457092 A1 30 May 2012 EP 2457092 B1 22 Oct 2014 SG 168431 A1 28 Feb 2011 SG 177702 A1 28 Feb 2012 US 2012190050 A1 26 Jul 2012 US 8669067 B2 11 Mar 2014			
End of Annex			
Due to data integration issues this family listing may not include 10 digit Australian applications filed since May 2001. Form PCT/ISA/210 (Family Annex)(July 2009)			