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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
4 November 2010 (04.11.2010)

PCT

(10) International Publication Number
WO 2010/125566 A2

(51) International Patent Classification:
A61K 38/17 (2006.01)

(21) International Application Number:
PCT/IL2010/000342

(22) International Filing Date:
27 April 2010 (27.04.2010)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/172,800 27 April 2009 (27.04.2009) US

(71) Applicant *(for all designated States except US)*: **TECHNION RESEARCH AND DEVELOPMENT FOUNDATION LTD.** [IL/IL]; Technion City, 32000 Haifa (IL).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **ADMON, Arie** [IL/IL]; 3 Irit Street, 36532 Tivon (IL). **BARNEA, Eilon** [IL/IL]; 10/5 Hasavion Street, 36822 Nesher (IL). **BEER, Ilan** [IL/IL]; 3 Petel Street, 34556 Haifa (IL). **ZIV, Tamar** [IL/IL]; 10 Aharonson Street, 34785 Haifa (IL).

(74) Agents: **WEBB, Cynthia** et al.; Webb & CO., P.O. Box 2189, 76121 Rehovot (IL).

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

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

Published:

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

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

WO 2010/125566 A2

(54) **Title**: MARKERS FOR CANCER DETECTION

(57) **Abstract**: The present invention relates to methods for detecting, prognosing and staging cancers, in particular cancers of the gastrointestinal tract. The methods of the invention comprise detecting specific protein markers in a tissue of interest, wherein the detected levels thereof may be indicative of pre-cancerous or cancerous tissue, or the stage or prognosis of a cancer. Further provided are methods of treating cancer, and cancer detection kits.

MARKERS FOR CANCER DETECTION

FIELD OF THE INVENTION

The present invention relates to methods for detecting, prognosing and staging cancers, in particular cancers of the gastrointestinal tract. The methods of the invention comprise detecting specific protein markers in a tissue of interest, wherein the detected levels thereof may be indicative of pre-cancerous or cancerous tissue, or the stage or prognosis of a cancer. Further provided are methods of using the markers for treating cancer, and cancer detection kits.

BACKGROUND OF THE INVENTION

Colorectal cancer, also referred to as colon cancer or large bowel cancer, is a malignant neoplastic disease associated with tumors in the colon, rectum and appendix. With 655,000 deaths worldwide per year, it is the third most common form of cancer and the second leading cause of cancer-related death in the Western world.

Colorectal cancers originate in the colorectal epithelium and typically are not extensively vascularized (and therefore not invasive) during the early stages of development. The transition to a highly vascularized, invasive and ultimately metastatic cancer, which spreads throughout the body, commonly takes ten years or longer. If the cancer is detected prior to invasion, surgical removal of the cancerous tissue is an effective cure. However, colorectal cancer is often detected only upon manifestation of clinical symptoms, such as pain and black tarry stool. Generally, such symptoms are present only when the disease is well established, often after metastasis has occurred, and the prognosis for the patient is poor, even after surgical resection of the cancerous tissue. For example, patients diagnosed with early colon cancer generally have a much greater five-year survival rate as compared to the survival rate for patients diagnosed with distant metastasized colon cancer. Accordingly, early detection of colorectal cancer is of critical importance for reducing its morbidity.

Diagnostic methods for colon cancer most frequently depend on direct visual inspection of the gastrointestinal (GI) tract. Endoscopy involves inspection with a miniaturized light source at a probe end of a coherent bundle fiber optic cable. Reflected light beam images are returned through the fiber optic cable for detection by an external digital camera and display on an external monitor or for recording on an external video recorder or both. While this

technique allows for identification, removal, and biopsy of potentially cancerous growths such as polyps, its use is associated with certain disadvantages, such as being expensive, uncomfortable, inherently risky due to its invasive nature, and the inability to access some portions of the large intestine and most of the small intestine.

Swallowable endoscopy capsules containing miniaturized optical, digital camera and radio transmission systems have been subsequently developed along with complementary external monitoring systems for inspecting the GI tract. For example, the capsule marketed under the trade name PillCam® SB was initially approved by the U.S. Food and Drug Administration in 2001 for detection and diagnosis of disorders of the small intestine. U.S. Patent No. 5,604,531 discloses an *in vivo* video camera system comprising a swallowable capsule. The transit of endoscopy capsules through the small intestine is dependent on peristalsis, meaning that some areas with lesions may be missed if the capsule is not retained in that area for a sufficient amount of time. Further, the endoscopy capsules in current use are not capable of identifying molecular markers, which may be early indicators of colorectal cancer, even prior to the development of pre-cancerous polyps.

U.S. Patent No. 7,468,044 discloses a system and method for *in vivo* and *in situ* detection of body lumen conditions, such as in the GI tract. The system comprises an interaction chamber comprising an indicator; a light source for illuminating the interaction chamber; and an optical detector for detecting in vivo optical changes occurring in the interaction chamber upon reaction of the indicator with an endo-luminal sample.

U.S. Patent No. 7,515,953 discloses a method for detecting fluorescence emitted by cells in a wall of a body lumen, such as an intestinal wall, the method comprising use of a swallowable capsule, the capsule comprising a light source and a fluorescent-labeled probe, which is released from a reservoir in the capsule. According to the disclosure, an electric field generated from an electrode in the capsule enhances uptake of the probe, and a detector in the capsule detects the fluorescent signal emitted. By determining the intensity and/or position in the lumen wall of the fluorescent signal, a drug for killing abnormal cells is released from a second reservoir in the capsule. According to the disclosure, the abnormal cells may be cancer cells, colon polyps or precancerous cells.

U.S. Patent Application Publication No. 2008/0146896 discloses a device, such as an autonomous capsule, for *in vivo* analysis which includes a reaction chamber to store a detecting reagent able to react with a sample collected *in vivo*; and optionally a labeled-substance chamber to store a labeled substance able to bind to a compound resulting from a

reaction of the detecting reagent and the sample. According to the disclosure, the detecting reagent may be an antibody.

U.S. Patent Application Publication No. 2009/0216082 discloses a device system for *in vivo* detection of target molecules in an endo-luminal sample, and a method for *in vivo* magnetic immunoassay, which may be used for the detection of cancer in the gastrointestinal tract.

Yet other methods of colon cancer detection are based on detection of particular proteins or genes which are considered to be specifically or differentially expressed in colon cancer.

U.S. Patent No. 7,507,541 discloses a method of detecting the presence of *inter alia* colon cancer that is based on determining the level of 36P6D5 protein expressed by cells in a test tissue sample from an individual, and comparing the level to that expressed in a corresponding normal tissue sample.

U.S. Patent No. 7,501,242 discloses a method of detecting colon cancer that is based on detecting levels of expression of tyrosine threonine kinase (TTK) in a test sample, such as a colon sample, that are increased by at least two fold relative to the level of expression in a normal non-cancer sample of the same type.

U.S. Patent No. 7,452,727 discloses a automatable method for identifying cancer cells and their precursor cells that is based on detecting at least two molecular markers, wherein the detection of each marker alone is not a reliable indicator of the presence of cancer cells and their precursor cells. According to the disclosure, the molecular markers may be selected from her2/neu, Ki67, p53, her2/neu, bcl-2, MN, mdm-2, EGF receptor, bcl-2 and p16.

U.S. Patent No.7,402,403 discloses a method for the detection of cancer or early neoplastic change that is based on detecting autoantibodies directed to tumor marker antigens in a sample of bodily fluids, wherein the tumor marker antigens are selected from MUC1, p53, c-erbB2, Ras, c-myc, BRCA1, BRCA2, PSA, APC and CA125.

U.S. Patent No. 7,129,043 discloses a method of identifying a human subject having an increased risk of developing colon cancer that is based on detecting upregulation of the CLN3 gene.

U.S. Patent No. 7,115,368 relates to a method of detecting epithelial cancer cells *inter alia* colon cancer that is based on detection of expression in a biological sample of pellino proteins.

U.S. Patent No. 7,098,008 relates to a method for detection of cancer *inter alia* colon cancer that is based on detecting expression of melanoma antigen gene (MAGE).

U.S. Patent No. 7,078,180 relates to a method of diagnosing a cancer *inter alia* colon cancer that is based on detection of a ZEB (zfh-1/delta EF1) polypeptide.

U.S. Patent No. 6,949,339 relates to methods for detecting, diagnosing, monitoring, staging, and prognosticating colon cancers, based on detection of Colon Specific Genes.

U.S. Patent No. 6,919,176 discloses a method of detecting cancer *inter alia* colon cancer that is based on detection of expression of specific G-protein coupled receptors.

There remains an unmet need for methods of early detection, prognosis and treatment of colon cancer.

SUMMARY OF THE INVENTION

The present invention provides methods of detecting cancer that are based on the qualitative or quantitative identification of particular proteins, also referred to herein as molecular markers. Further provided are methods of cancer prevention, prognosis and treatment.

Disclosed herein for the first time is a specific group of protein markers which are indicative of both pre-cancerous and cancerous lesions. Further disclosed herein for the first time is an additional specific group of protein markers which are primarily indicative of pre-cancerous lesions.

The invention is based in part, on the unexpected discovery that the level of expression within the gastrointestinal tract of certain proteins is significantly increased in both pre-cancerous and cancerous tissue relative to the level of expression of the same proteins in healthy tissue of the same type Surprisingly, the expression of these markers is elevated, even compared to healthy tissue bordering tumor growth. It has also been surprisingly found that the expression of yet other markers is significantly increased in early stage cancer, and significantly decreased in later stages of cancer

Without wishing to be bound by any particular theory or mechanism of action, the invention enables identification of individuals at risk of developing cancer, in particular colorectal cancer, even prior to observable histological changes in affected tissue. Since the methods of the invention are based on changes in protein expression patterns in cells, rather than later occurring pathological changes in tissue, a level of sensitivity is obtained that is

greater by many orders of magnitude than current techniques of cancer detection. Thus, in the case of colorectal cancer, the invention provides a means of predicting the disease well in advance of the possibility of detecting potentially cancerous polyps by conventional endoscopic examination, the latter being the current yet inadequate standard for early detection.

The principles of the current invention are exemplified herein by quantitative mass spectroscopy analysis of healthy, pre-cancerous and cancerous tissue obtained from the gastrointestinal tract of patients during surgical excision of early stage (e.g. polyps) or more advanced stage tumors, which has resulted in the identification of a specific group of proteins, the expression of each of which is significantly increased in diseased colon tissue, as compared to healthy colon tissue. This group of proteins, which includes KIAA0152 (SEQ ID NO:1), NAMPT (SEQ ID NO:2), PYCR1 (SEQ ID NO:3), GPX2 (SEQ ID NO:4), PRKDC (SEQ ID NO:5), ALDH18A1 (SEQ ID NO:6), OCIAD2 (SEQ ID NO:7), GCS1 (SEQ ID NO:8), GMDS (SEQ ID NO:9), ARF4 (SEQ ID NO:10), ARF5 (SEQ ID NO:11), LRPPRC (SEQ ID NO:12), CTNNB1 (SEQ ID NO:13), ARF3 (SEQ ID NO:14), GCN1L1 (SEQ ID NO:15), BDH1 (SEQ ID NO:16), RPL9 (SEQ ID NO:17), UGCGL1 (SEQ ID NO:18), FAM3D (SEQ ID NO:19) and CCT4 (SEQ ID NO:20), has not been previously disclosed, either in part or as a whole, to be useful for detection of colorectal cancer at any stage of the disease.

It is to be specifically understood however, that the current method of the invention need not be limited to examination of colon tissue obtained by surgical means, nor should it be limited to detection and quantification of the subject molecular markers using mass spectrometry techniques. Rather, the invention may be advantageously practiced using for example immunological techniques and reagents for detection of the subject molecular markers, either *in vivo* or *ex vivo*. For example, labeled monoclonal antibodies can be used for *in vivo* detection and quantitation of such proteins in different tissue compartments and regions. The detection may be accomplished for example, using pharmaceutical compositions or endoscopy probes which incorporate specifically designed chemical, immunological or nucleic acid reagents. Advantageously, labeled reagents such as antibodies, which specifically interact with the subject molecular markers may be prepared as injectable or ingestible pharmaceutical compositions and following administration the interaction with their molecular targets may be externally monitored. Alternately or in addition, the invention may be practiced by analysis of biological samples obtained from a subject, such as biopsy

tissue.

In a first aspect, the invention provides a method of detecting cancer in a subject, the method comprising: (i) detecting in a biological sample from the subject at least one protein selected from the group consisting of KIAA0152 (SEQ ID NO:1), NAMPT (SEQ ID NO:2), PYCR1 (SEQ ID NO:3), GPX2 (SEQ ID NO:4), PRKDC (SEQ ID NO:5), ALDH18A1 (SEQ ID NO:6), OCIAD2 (SEQ ID NO:7), GCS1 (SEQ ID NO:8), GMDS (SEQ ID NO:9), ARF4 (SEQ ID NO:10), ARF5 (SEQ ID NO:11), LRPPRC (SEQ ID NO:12), CTNNB1 (SEQ ID NO:13), ARF3 (SEQ ID NO:14), GCN1L1 (SEQ ID NO:15), BDH1 (SEQ ID NO:16), RPL9 (SEQ ID NO:17), UGCGL1 (SEQ ID NO:18), FAM3D (SEQ ID NO:19) and CCT4 (SEQ ID NO:20), so as to determine the level of the at least one protein; and (ii) comparing the level determined in (i) to a reference level of the same at least one protein, wherein detection of a level of said at least one protein in the biological sample which is significantly different from the reference level, is indicative of cancer in the subject.

Each of the aforementioned proteins represents a separate embodiment of the invention and may be used independently from or in combination with any of the others.

In another aspect, the invention provides a method for determining the stage of a cancerous or pre-cancerous growth in a subject, the method comprising: (i) detecting in a test sample from the subject at least one protein selected from the group consisting of KIAA0152 (SEQ ID NO:1), NAMPT (SEQ ID NO:2), PYCR1 (SEQ ID NO:3), GPX2 (SEQ ID NO:4), PRKDC (SEQ ID NO:5), ALDH18A1 (SEQ ID NO:6), OCIAD2 (SEQ ID NO:7), GCS1 (SEQ ID NO:8), GMDS (SEQ ID NO:9), ARF4 (SEQ ID NO:10), ARF5 (SEQ ID NO:11), LRPPRC (SEQ ID NO:12), CTNNB1 (SEQ ID NO:13), ARF3 (SEQ ID NO:14), GCN1L1 (SEQ ID NO:15), BDH1 (SEQ ID NO:16), RPL9 (SEQ ID NO:17), UGCGL1 (SEQ ID NO:18), FAM3D (SEQ ID NO:19) and CCT4 (SEQ ID NO:20), so as to determine the level of the at least one protein; and (ii) comparing the level determined in (i) to a reference level of the same at least one protein; wherein the level detected in the test sample is indicative of the stage of the growth.

Each of the aforementioned proteins represents a separate embodiment of the invention and may be used independently from or in combination with any of the others.

In another aspect, the invention provides a method for determining the prognosis of a cancerous disease in a subject, the method comprising: (i) detecting in a test sample from the subject at least one protein selected from the group consisting of KIAA0152 (SEQ ID NO:1),

NAMPT (SEQ ID NO:2), PYCR1 (SEQ ID NO:3), GPX2 (SEQ ID NO:4), PRKDC (SEQ ID NO:5), ALDH18A1 (SEQ ID NO:6), OCIAD2 (SEQ ID NO:7), GCS1 (SEQ ID NO:8), GMDS (SEQ ID NO:9), ARF4 (SEQ ID NO:10), ARF5 (SEQ ID NO:11), LRPPRC (SEQ ID NO:12), CTNNB1 (SEQ ID NO:13), ARF3 (SEQ ID NO:14), GCN1L1 (SEQ ID NO:15), BDH1 (SEQ ID NO:16), RPL9 (SEQ ID NO:17), UGCGL1 (SEQ ID NO:18), FAM3D (SEQ ID NO:19) and CCT4 (SEQ ID NO:20), so as to determine the level of the at least one protein; and (ii) comparing the level determined in (i) to a reference level of the same at least one protein; wherein the level detected in the test sample is indicative of the prognosis of the cancerous disease.

Each of the aforementioned proteins represents a separate embodiment of the invention and may be used independently from or in combination with any of the others.

In particular embodiments, the biological sample is selected from the group consisting of blood, serum, nipple aspirate fluid, lymph node aspirate, a biopsy sample, a tumor sample, a tissue sample, mucosal fluid, cervical wash, lacrimal duct fluid, urine, saliva, pleural effusion and sputum. Each of the aforementioned materials represents a separate embodiment of the invention and may be used independently from or in combination with any of the others.

In currently preferred embodiments, the biological sample comprises a tissue sample. In currently preferred embodiments, the tissue sample comprises gastrointestinal tissue, particularly colorectal tissue or pre-cancerous tissue such as polyps. In currently preferred embodiments, the biological sample is selected from a tumor within the gastrointestinal tract, particularly a colorectal tumor. In particular embodiments, the gastrointestinal tissue is from an area or organ selected from the group consisting of the esophagus, the stomach, the small intestine, the large intestine (colon), the rectum, the appendix and a combination thereof.

In particular embodiments, the cancer being detected is selected from the group consisting of adrenal cancer, bladder cancer, bone cancer, brain cancer, breast cancer, cervical cancer, colorectal cancer, fallopian tube cancer, gastric cancer, head and neck cancer, hepatic cancer, lung cancer including small cell lung cancer and non-small cell lung cancer, melanoma, neuroblastoma, oral cancer, ovarian cancer, pancreatic cancer, prostate cancer, thyroid and parathyroid cancer, renal cancer, sarcoma, thymoma, hematological malignancies and germ cell tumors. In a currently preferred embodiment, the cancer is colorectal cancer. Each of the aforementioned cancers represents a separate embodiment of the invention and may be used independently from any of the others.

In a particular embodiment, the at least one protein detected is indicative of a disorder selected from the group consisting of pre-cancerous polyps, early stage colorectal cancer and advanced stage colorectal cancer. In particular embodiments, a method of detecting pre-cancerous polyps or early stage colorectal cancer comprises detecting at least one protein selected from the group consisting of KIAA0152 (SEQ ID NO:1), NAMPT (SEQ ID NO:2), PYCR1 (SEQ ID NO:3), GPX2 (SEQ ID NO:4), PRKDC (SEQ ID NO:5), ALDH18A1 (SEQ ID NO:6), OCIAD2 (SEQ ID NO:7), GCS1 (SEQ ID NO:8), GMDS (SEQ ID NO:9), ARF4 (SEQ ID NO:10), ARF5 (SEQ ID NO:11), LRPPRC (SEQ ID NO:12), CTNNB1 (SEQ ID NO:13), ARF3 (SEQ ID NO:14), GCN1L1 (SEQ ID NO:15), BDH1 (SEQ ID NO:16), RPL9 (SEQ ID NO:17), UGCGL1 (SEQ ID NO:18), FAM3D (SEQ ID NO:19) and CCT4 (SEQ ID NO:20); and further comprises detecting at least one protein selected from the group consisting of CPT2 (SEQ ID NO:21), ARL1 (SEQ ID NO:22), PFKL (SEQ ID NO:23), GOT2 (SEQ ID NO:24), AP1G1 (SEQ ID NO:25), STRBP (SEQ ID NO:26), CLCA1 (SEQ ID NO:27), CYFIP1 (SEQ ID NO:28), COQ9 (SEQ ID NO:29), NDUFA9 (SEQ ID NO:30), ALDH7A1 (SEQ ID NO:31), HMGCS1 (SEQ ID NO:32), NNT (SEQ ID NO:33), PRDX5 (SEQ ID NO:34), PCCB (SEQ ID NO:35), COPZ1 (SEQ ID NO:36), BAX (SEQ ID NO:37), ACAD9 (SEQ ID NO:38), UBXD8 (SEQ ID NO:39), HMGCS2 (SEQ ID NO:40), SLC25A3 (SEQ ID NO:41), SLC25A11 (SEQ ID NO:42), PDCD6 (SEQ ID NO:43), UCRC (SEQ ID NO:44), DEFA6 (SEQ ID NO:45), DYNC1H1 (SEQ ID NO:46), HK1 (SEQ ID NO:47), CYFIP2 (SEQ ID NO:48), DCI (SEQ ID NO:49) and CISD1 (SEQ ID NO:50). Each of the aforementioned proteins represents a separate embodiment of the invention and may be used independently from or in combination with any of the others.

In particular embodiments, a method of detecting pre-cancerous polyps or early stage colorectal cancer comprises detecting at least one of KIAA0152 (SEQ ID NO:1) and NAMPT (SEQ ID NO:2); and further comprises detecting at least one protein selected from the group consisting of CPT2 (SEQ ID NO:21), ARL1 (SEQ ID NO:22), PFKL (SEQ ID NO:23), GOT2 (SEQ ID NO:24), AP1G1 (SEQ ID NO:25), STRBP (SEQ ID NO:26), CLCA1 (SEQ ID NO:27), CYFIP1 (SEQ ID NO:28), COQ9 (SEQ ID NO:29), NDUFA9 (SEQ ID NO:30), ALDH7A1 (SEQ ID NO:31), HMGCS1 (SEQ ID NO:32), NNT (SEQ ID NO:33), PRDX5 (SEQ ID NO:34), PCCB (SEQ ID NO:35), COPZ1 (SEQ ID NO:36), BAX (SEQ ID NO:37), ACAD9 (SEQ ID NO:38), UBXD8 (SEQ ID NO:39), HMGCS2 (SEQ ID NO:40), SLC25A3 (SEQ ID NO:41), SLC25A11 (SEQ ID NO:42), PDCD6 (SEQ ID NO:43), UCRC (SEQ ID NO:44), DEFA6 (SEQ ID NO:45), DYNC1H1 (SEQ ID NO:46), HK1 (SEQ ID NO:47), CYFIP2 (SEQ ID NO:48), DCI (SEQ ID NO:49) and CISD1 (SEQ

ID NO:50). Each of the aforementioned proteins represents a separate embodiment of the invention and may be used independently from or in combination with any of the others.

In particular embodiments, a method of the invention comprises use of at least one reagent suitable for detecting the level of at least one protein selected from the group consisting of KIAA0152 (SEQ ID NO:1), NAMPT (SEQ ID NO:2), PYCR1 (SEQ ID NO:3), GPX2 (SEQ ID NO:4), PRKDC (SEQ ID NO:5), ALDH18A1 (SEQ ID NO:6), OCIAD2 (SEQ ID NO:7), GCS1 (SEQ ID NO:8), GMDS (SEQ ID NO:9), ARF4 (SEQ ID NO:10), ARF5 (SEQ ID NO:11), LRPPRC (SEQ ID NO:12), CTNNB1 (SEQ ID NO:13), ARF3 (SEQ ID NO:14), GCN1L1 (SEQ ID NO:15), BDH1 (SEQ ID NO:16), RPL9 (SEQ ID NO:17), UGCGL1 (SEQ ID NO:18), FAM3D (SEQ ID NO:19) and CCT4 (SEQ ID NO:20).

In particular embodiments, said reagent is suitable for detecting the level of at least one of KIAA0152 (SEQ ID NO:1) and NAMPT (SEQ ID NO:2). Each of the aforementioned reagents represents a separate embodiment of the invention and may be used independently from or in combination with any of the others.

In particular embodiments, a method of the invention further comprises use of at least one reagent suitable for detecting the level of at least one protein selected from the group consisting of CPT2 (SEQ ID NO:21), ARL1 (SEQ ID NO:22), PFKL (SEQ ID NO:23), GOT2 (SEQ ID NO:24), AP1G1 (SEQ ID NO:25), STRBP (SEQ ID NO:26), CLCA1 (SEQ ID NO:27), CYFIP1 (SEQ ID NO:28), COQ9 (SEQ ID NO:29), NDUFA9 (SEQ ID NO:30), ALDH7A1 (SEQ ID NO:31), HMGCS1 (SEQ ID NO:32), NNT (SEQ ID NO:33), PRDX5 (SEQ ID NO:34), PCCB (SEQ ID NO:35), COPZ1 (SEQ ID NO:36), BAX (SEQ ID NO:37), ACAD9 (SEQ ID NO:38), UBXD8 (SEQ ID NO:39), HMGCS2 (SEQ ID NO:40), SLC25A3 (SEQ ID NO:41), SLC25A11 (SEQ ID NO:42), PDCD6 (SEQ ID NO:43), UCRC (SEQ ID NO:44), DEFA6 (SEQ ID NO:45), DYNC1H1 (SEQ ID NO:46), HK1 (SEQ ID NO:47), CYFIP2 (SEQ ID NO:48), DCI (SEQ ID NO:49) and CISD1 (SEQ ID NO:50). Each of the aforementioned reagents represents a separate embodiment of the invention and may be used independently from or in combination with any of the others.

In particular embodiments, the biological sample or test sample is obtained from the subject by a procedure selected from the group consisting of biopsy, flexible endoscopy, double balloon endoscopy and surgical colorectal re-sectioning.

In particular embodiments, the biological sample or test sample is assessed *in vivo* in the subject. In particular embodiments, the method comprises contacting a body tissue, cavity

or fluid with at least one of a pharmaceutical composition and an endoscopy apparatus.

In particular embodiments, the method comprises contacting a body tissue, cavity or fluid with at least one reagent suitable for detecting the level of at least one protein selected from the group consisting of KIAA0152 (SEQ ID NO:1), NAMPT (SEQ ID NO:2), PYCR1 (SEQ ID NO:3), GPX2 (SEQ ID NO:4), PRKDC (SEQ ID NO:5), ALDH18A1 (SEQ ID NO:6), OCIAD2 (SEQ ID NO:7), GCS1 (SEQ ID NO:8), GMDS (SEQ ID NO:9), ARF4 (SEQ ID NO:10), ARF5 (SEQ ID NO:11), LRPPRC (SEQ ID NO:12), CTNNB1 (SEQ ID NO:13), ARF3 (SEQ ID NO:14), GCN1L1 (SEQ ID NO:15), BDH1 (SEQ ID NO:16), RPL9 (SEQ ID NO:17), UGCGL1 (SEQ ID NO:18), FAM3D (SEQ ID NO:19) and CCT4 (SEQ ID NO:20)4. In particular embodiments, the body tissue, cavity or fluid is contacted with a reagent is suitable for detecting the level of at least one of KIAA0152 (SEQ ID NO:1) and NAMPT (SEQ ID NO:2). Each of the aforementioned reagents represents a separate embodiment of the invention and may be used independently from or in combination with any of the others.

In particular embodiments, the method further comprises contacting a body tissue, cavity or fluid with at least one reagent suitable for detecting the level of at least one protein selected from the group consisting of CPT2 (SEQ ID NO:21), ARL1 (SEQ ID NO:22), PFKL (SEQ ID NO:23), GOT2 (SEQ ID NO:24), AP1G1 (SEQ ID NO:25), STRBP (SEQ ID NO:26), CLCA1 (SEQ ID NO:27), CYFIP1 (SEQ ID NO:28), COQ9 (SEQ ID NO:29), NDUFA9 (SEQ ID NO:30), ALDH7A1 (SEQ ID NO:31), HMGCS1 (SEQ ID NO:32), NNT (SEQ ID NO:33), PRDX5 (SEQ ID NO:34), PCCB (SEQ ID NO:35), COPZ1 (SEQ ID NO:36), BAX (SEQ ID NO:37), ACAD9 (SEQ ID NO:38), UBXD8 (SEQ ID NO:39), HMGCS2 (SEQ ID NO:40), SLC25A3 (SEQ ID NO:41), SLC25A11 (SEQ ID NO:42), PDCD6 (SEQ ID NO:43), UCRC (SEQ ID NO:44), DEFA6 (SEQ ID NO:45), DYNC1H1 (SEQ ID NO:46), HK1 (SEQ ID NO:47), CYFIP2 (SEQ ID NO:48), DCI (SEQ ID NO:49) and CISD1 (SEQ ID NO:50). Each of the aforementioned reagents represents a separate embodiment of the invention and may be used independently from or in combination with any of the others.

In particular embodiments, the method comprises administering a diagnostic pharmaceutical composition. In particular embodiments, the administering of the pharmaceutical composition is by a route selected from the group consisting of oral, parenteral, subcutaneous, intramuscular, intrathoracic and intraarticular.

In particular embodiments, the pharmaceutical composition or the endoscopy apparatus

comprise at least one reagent suitable for detecting the level of at least one of the aforementioned proteins. In particular embodiments, the pharmaceutical composition or the endoscopy apparatus comprise at least one reagent suitable for detecting the level of at least one of KIAA0152 (SEQ ID NO:1) and NAMPT (SEQ ID NO:2).

In particular embodiments, the reagent specifically interacts with at least one of the aforementioned proteins or with nucleic acid encoding at least one of the aforementioned proteins, or a fragment of said protein. In particular embodiments, the pharmaceutical composition or the endoscopy apparatus comprises a multiplicity of reagents, wherein each reagent of the multiplicity specifically interacts with one distinct protein or with nucleic acid encoding one distinct protein or a fragment thereof. In particular embodiments, the reagent is selected from an antibody, an antibody mimetic and a nucleic acid. In particular embodiments, the antibody is selected from a monoclonal antibody, a humanized antibody, a single chain antibody, an antibody fragment and combinations thereof. In particular embodiments, the pharmaceutical composition or the endoscopy apparatus comprise a multiplicity of antibody mimetics, wherein each antibody mimetic of the multiplicity specifically interacts with one of the aforementioned proteins.

In particular embodiments, the reagent comprises a detectable label, such as a fluorescent label, a radiolabel or an enzymatic label.

In particular embodiments, the pharmaceutical composition or the endoscopy apparatus comprise at least one least one nucleic acid, wherein the at least one nucleic acid is complementary to a nucleic acid encoding at least one of the aforementioned proteins, or a fragment of said protein. In particular embodiments, the at least one nucleic acid comprises a multiplicity of nucleic acids, wherein each nucleic acid of the multiplicity is complementary to a nucleic acid encoding one distinct protein of the aforementioned proteins, or a fragment of said protein. In particular embodiments, the nucleic acid comprises a detectable label, such as a fluorescent label, a radiolabel or an enzymatic label.

In particular embodiments, the detecting in step (i) comprises use of an assay system. In particular embodiments, the assay system comprises an immunoassay, a nucleic acid hybridization assay, a binding assay, an array, a phage display library or combinations thereof. In particular embodiments, the array is a protein array, or a phage display library array. In particular embodiments, the assay system comprises at least one reagent suitable for detecting the level of at least one of the aforementioned proteins, as described herein. In particular embodiments, the assay system comprises a multiplicity of such reagents, for

example antibodies, wherein each reagent in the multiplicity specifically interacts with one of the aforementioned proteins.

In particular embodiments, the detecting comprises use of an external monitoring system. In particular embodiments, the external monitoring system is configured to display the level of the at least one protein detected.

In particular embodiments, a method of the invention further comprises detecting in the sample at least one protein selected from the group consisting of FAM62B (SEQ ID NO:51), SLC1A5 (SEQ ID NO:52), RSL1D1 (SEQ ID NO:53), LYZ (SEQ ID NO:54), THBS1 (SEQ ID NO:55), LMO7 (SEQ ID NO:56), TNC (SEQ ID NO:57), RBM39 (SEQ ID NO:58), ILVBL (SEQ ID NO:59), ERO1L (SEQ ID NO:60), LOC442497 (SEQ ID NO:61), TCOF1 (SEQ ID NO:62), SERPINB9 (SEQ ID NO:63), HSDL2 (SEQ ID NO:64), ADAMDEC1 (SEQ ID NO:65), AMACR (SEQ ID NO:66), AMACR;C1QTNF3 (SEQ ID NO:67), ARID1A (SEQ ID NO:68), CEBPZ (SEQ ID NO:69), COL5A1 (SEQ ID NO:70), EFEMP2 (SEQ ID NO:71), FAM84B (SEQ ID NO:72), FKBP10 (SEQ ID NO:73), FKBP9 (SEQ ID NO:74), GPRC5A (SEQ ID NO:75), KPNA2 (SEQ ID NO:76), MMP1 (SEQ ID NO:77), PNMA5 (SEQ ID NO:78), POLR1C (SEQ ID NO:79), SPARC (SEQ ID NO:80), UBAP2 (SEQ ID NO:81), UCK2 (SEQ ID NO:82) and WDR74 (SEQ ID NO:83). Each of the aforementioned proteins represents a separate embodiment of the invention and may be used independently from or in combination with any of the others.

In particular embodiments, a method of the invention further comprises use of at least one reagent suitable for detecting the level of at least one protein selected from the group consisting of FAM62B (SEQ ID NO:51), SLC1A5 (SEQ ID NO:52), RSL1D1 (SEQ ID NO:53), LYZ (SEQ ID NO:54), THBS1 (SEQ ID NO:55), LMO7 (SEQ ID NO:56), TNC (SEQ ID NO:57), RBM39 (SEQ ID NO:58), ILVBL (SEQ ID NO:59), ERO1L (SEQ ID NO:60), LOC442497 (SEQ ID NO:61), TCOF1 (SEQ ID NO:62), SERPINB9 (SEQ ID NO:63), HSDL2 (SEQ ID NO:64), ADAMDEC1 (SEQ ID NO:65), AMACR (SEQ ID NO:66), AMACR;C1QTNF3 (SEQ ID NO:67), ARID1A (SEQ ID NO:68), CEBPZ (SEQ ID NO:69), COL5A1 (SEQ ID NO:70), EFEMP2 (SEQ ID NO:71), FAM84B (SEQ ID NO:72), FKBP10 (SEQ ID NO:73), FKBP9 (SEQ ID NO:74), GPRC5A (SEQ ID NO:75), KPNA2 (SEQ ID NO:76), MMP1 (SEQ ID NO:77), PNMA5 (SEQ ID NO:78), POLR1C (SEQ ID NO:79), SPARC (SEQ ID NO:80), UBAP2 (SEQ ID NO:81), UCK2 (SEQ ID NO:82) and WDR74 (SEQ ID NO:83). Each of the aforementioned reagents represents a separate embodiment of the invention and may be used independently from or in combination with

any of the others.

In particular embodiments, the cancer is a cancer other than colorectal cancer and the at least one protein includes at least one of FAM62B (SEQ ID NO:51), SLC1A5 (SEQ ID NO:52), RSL1D1 (SEQ ID NO:53), LYZ (SEQ ID NO:54), TNC(SEQ ID NO:57), RBM39 (SEQ ID NO:58), ERO1L (SEQ ID NO:60), LOC442497 (SEQ ID NO:61), TCOF1 (SEQ ID NO:62), NAMPT (SEQ ID NO:2), SERPINB9 (SEQ ID NO:63), HSDL2 (SEQ ID NO:64).

In particular embodiments, the level of the at least one protein in the biological or test sample is increased by at least 2-fold, or at least 3-fold, or at least 5-fold, or at least 10-fold, or at least 20-fold, or at least 50-fold, relative to the reference level. In particular embodiments, the reference level is representative of the level of the same protein in non-diseased tissue. In particular embodiments, the reference level is representative of the level of the same protein in a particular stage or form of cancer. In particular embodiments, the reference level is representative of the level of the same protein in a cancer having a known prognosis.

In another aspect, the invention provides a method of treating cancer, the method comprising administering to a subject in need thereof at least one pharmaceutical agent, wherein the pharmaceutical agent specifically interacts with at least one protein selected from the group consisting of KIAA0152 (SEQ ID NO:1), NAMPT (SEQ ID NO:2), PYCR1 (SEQ ID NO:3), GPX2 (SEQ ID NO:4), PRKDC (SEQ ID NO:5), ALDH18A1 (SEQ ID NO:6), OCIAD2 (SEQ ID NO:7), GCS1 (SEQ ID NO:8), GMDS (SEQ ID NO:9), ARF4 (SEQ ID NO:10), ARF5 (SEQ ID NO:11), LRPPRC (SEQ ID NO:12), CTNNB1 (SEQ ID NO:13), ARF3 (SEQ ID NO:14), GCN1L1 (SEQ ID NO:15), BDH1 (SEQ ID NO:16), RPL9 (SEQ ID NO:17), UGCGL1 (SEQ ID NO:18), FAM3D (SEQ ID NO:19), CCT4 (SEQ ID NO:20), CPT2 (SEQ ID NO:21), ARL1 (SEQ ID NO:22), PFKL (SEQ ID NO:23), GOT2 (SEQ ID NO:24), AP1G1 (SEQ ID NO:25), STRBP (SEQ ID NO:26), CLCA1 (SEQ ID NO:27), CYFIP1 (SEQ ID NO:28), COQ9 (SEQ ID NO:29), NDUFA9 (SEQ ID NO:30), ALDH7A1 (SEQ ID NO:31), HMGCS1 (SEQ ID NO:32), NNT (SEQ ID NO:33), PRDX5 (SEQ ID NO:34), PCCB (SEQ ID NO:35), COPZ1 (SEQ ID NO:36), BAX (SEQ ID NO:37), ACAD9 (SEQ ID NO:38), UBXD8 (SEQ ID NO:39), HMGCS2 (SEQ ID NO:40), SLC25A3 (SEQ ID NO:41), SLC25A11 (SEQ ID NO:42), PDCD6 (SEQ ID NO:43), UCRC (SEQ ID NO:44), DEFA6 (SEQ ID NO:45), DYNC1H1 (SEQ ID NO:46), HK1 (SEQ ID NO:47), CYFIP2 (SEQ ID NO:48), DCI (SEQ ID NO:49), CISD1 (SEQ ID NO:50), LYZ (SEQ ID NO:54), LOC442497;SLC3A2 (SEQ ID NO:61), DMBT1 (SEQ ID NO:84), NUCB1 (SEQ

ID NO:.85), GGH (SEQ ID NO:86), AGR3 (SEQ ID NO:87), TM9SF2 (SEQ ID NO:88), SYK (SEQ ID NO:89), GCA (SEQ ID NO:90), HDLBP (SEQ ID NO:91), C1QBP (SEQ ID NO:92) and CLIC1 (SEQ ID NO:93). Each of the aforementioned agents represents a separate embodiment of the invention and may be used independently from or in combination with any of the others.

In a particular embodiment, the at least one protein is selected from the group consisting of KIAA0152 (SEQ ID NO:1), LYZ (SEQ ID NO:54), LOC442497;SLC3A2 (SEQ ID NO:61), DMBT1 (SEQ ID NO:84), NUCB1 (SEQ ID NO:.85), GGH (SEQ ID NO:86), AGR3 (SEQ ID NO:87), TM9SF2 (SEQ ID NO:88), SYK (SEQ ID NO:89), GCA (SEQ ID NO:90), HDLBP (SEQ ID NO:91), C1QBP (SEQ ID NO:92) and CLIC1 (SEQ ID NO:93). In particular embodiments, the pharmaceutical agent specifically interacts with KIAA0152 (SEQ ID NO:1). Each of the aforementioned reagents represents a separate embodiment of the invention and may be used independently from or in combination with any of the others.

In particular embodiments, the pharmaceutical agent is selected from an antibody and an antibody mimetic. In particular embodiments, the pharmaceutical agent further comprises a cytotoxic moiety, such as a plant toxin, a bacterial toxin, a radioactive moiety or a chemotherapeutic agent. In particular embodiments, the pharmaceutical agent is selected from a chemical conjugate and fusion protein.

In another aspect, the invention provides an antigen composition comprising at least one protein selected from the group consisting of KIAA0152 (SEQ ID NO:1), NAMPT (SEQ ID NO:2), PYCR1 (SEQ ID NO:3), GPX2 (SEQ ID NO:4), PRKDC (SEQ ID NO:5), ALDH18A1 (SEQ ID NO:6), OCIAD2 (SEQ ID NO:7), GCS1 (SEQ ID NO:8), GMDS (SEQ ID NO:9), ARF4 (SEQ ID NO:10), ARF5 (SEQ ID NO:11), LRPPRC (SEQ ID NO:12), CTNNB1 (SEQ ID NO:13), ARF3 (SEQ ID NO:14), GCN1L1 (SEQ ID NO:15), BDH1 (SEQ ID NO:16), RPL9 (SEQ ID NO:17), UGCGL1 (SEQ ID NO:18), FAM3D (SEQ ID NO:19), CCT4 (SEQ ID NO:20), CPT2 (SEQ ID NO:21), ARL1 (SEQ ID NO:22), PFKL (SEQ ID NO:23), GOT2 (SEQ ID NO:24), AP1G1 (SEQ ID NO:25), STRBP (SEQ ID NO:26), CLCA1 (SEQ ID NO:27), CYFIP1 (SEQ ID NO:28), COQ9 (SEQ ID NO:29), NDUFA9 (SEQ ID NO:30), ALDH7A1 (SEQ ID NO:31), HMGCS1 (SEQ ID NO:32), NNT (SEQ ID NO:33), PRDX5 (SEQ ID NO:34), PCCB (SEQ ID NO:35), COPZ1 (SEQ ID NO:36), BAX (SEQ ID NO:37), ACAD9 (SEQ ID NO:38), UBXD8 (SEQ ID NO:39), HMGCS2 (SEQ ID NO:40), SLC25A3 (SEQ ID NO:41), SLC25A11 (SEQ ID NO:42),

PDCD6 (SEQ ID NO:43), UCRC (SEQ ID NO:44), DEFA6 (SEQ ID NO:45), DYNC1H1 (SEQ ID NO:46), HK1 (SEQ ID NO:47), CYFIP2 (SEQ ID NO:48), DCI (SEQ ID NO:49), CISD1 (SEQ ID NO:50), LYZ (SEQ ID NO:54), LOC442497;SLC3A2 (SEQ ID NO:61), DMBT1 (SEQ ID NO:84), NUCB1 (SEQ ID NO:.85), GGH (SEQ ID NO:86), AGR3 (SEQ ID NO:87), TM9SF2 (SEQ ID NO:88), SYK (SEQ ID NO:89), GCA (SEQ ID NO:90), HDLBP (SEQ ID NO:91), C1QBP (SEQ ID NO:92) and CLIC1 (SEQ ID NO:93), or an immunogenic fragment of said at least one protein. Each of the aforementioned proteins represents a separate embodiment of the invention and may be used independently from or in combination with any of the others.

In a particular embodiment, the composition comprises at least one of CPT2 (SEQ ID NO:21), ARL1 (SEQ ID NO:22), PFKL (SEQ ID NO:23), GOT2 (SEQ ID NO:24), AP1G1 (SEQ ID NO:25), STRBP (SEQ ID NO:26), CLCA1 (SEQ ID NO:27), CYFIP1 (SEQ ID NO:28), COQ9 (SEQ ID NO:29), NDUFA9 (SEQ ID NO:30), ALDH7A1 (SEQ ID NO:31), HMGCS1 (SEQ ID NO:32), NNT (SEQ ID NO:33), PRDX5 (SEQ ID NO:34), PCCB (SEQ ID NO:35), COPZ1 (SEQ ID NO:36), BAX (SEQ ID NO:37), ACAD9 (SEQ ID NO:38), UBXD8 (SEQ ID NO:39), HMGCS2 (SEQ ID NO:40), SLC25A3 (SEQ ID NO:41), SLC25A11 (SEQ ID NO:42), PDCD6 (SEQ ID NO:43), UCRC (SEQ ID NO:44), DEFA6 (SEQ ID NO:45), DYNC1H1 (SEQ ID NO:46), HK1 (SEQ ID NO:47), CYFIP2 (SEQ ID NO:48), DCI (SEQ ID NO:49) and CISD1 (SEQ ID NO:50), or an immunogenic fragment thereof. Each of the aforementioned proteins represents a separate embodiment of the invention and may be used independently from or in combination with any of the others.

In a particular embodiment, the composition comprises KIAA0152 (SEQ ID NO:1) or an immunogenic fragment thereof. In a particular embodiment, a method of preventing or treating cancer in a subject in need thereof comprises administering to the subject an effective amount of said antigen composition. Further provided is an antigen composition as previously specified; for use in preventing or treating cancer.

In another aspect, the invention provides a kit for detecting, staging or prognosing cancer, the kit comprising at least one reagent suitable for detecting the level of at least one protein selected from the group consisting of KIAA0152 (SEQ ID NO:1), NAMPT (SEQ ID NO:2), PYCR1 (SEQ ID NO:3), GPX2 (SEQ ID NO:4), PRKDC (SEQ ID NO:5), ALDH18A1 (SEQ ID NO:6), OCIAD2 (SEQ ID NO:7), GCS1 (SEQ ID NO:8), GMDS (SEQ ID NO:9), ARF4 (SEQ ID NO:10), ARF5 (SEQ ID NO:11), LRPPRC (SEQ ID NO:12), CTNNB1 (SEQ ID NO:13), ARF3 (SEQ ID NO:14), GCN1L1 (SEQ ID NO:15),

BDH1 (SEQ ID NO:16), RPL9 (SEQ ID NO:17), UGCGL1 (SEQ ID NO:18), FAM3D (SEQ ID NO:19), CCT4 (SEQ ID NO:20), CPT2 (SEQ ID NO:21), ARL1 (SEQ ID NO:22), PFKL (SEQ ID NO:23), GOT2 (SEQ ID NO:24), AP1G1 (SEQ ID NO:25), STRBP (SEQ ID NO:26), CLCA1 (SEQ ID NO:27), CYFIP1 (SEQ ID NO:28), COQ9 (SEQ ID NO:29), NDUFA9 (SEQ ID NO:30), ALDH7A1 (SEQ ID NO:31), HMGCS1 (SEQ ID NO:32), NNT (SEQ ID NO:33), PRDX5 (SEQ ID NO:34), PCCB (SEQ ID NO:35), COPZ1 (SEQ ID NO:36), BAX (SEQ ID NO:37), ACAD9 (SEQ ID NO:38), UBXD8 (SEQ ID NO:39), HMGCS2 (SEQ ID NO:40), SLC25A3 (SEQ ID NO:41), SLC25A11 (SEQ ID NO:42), PDCD6 (SEQ ID NO:43), UCRC (SEQ ID NO:44), DEFA6 (SEQ ID NO:45), DYNC1H1 (SEQ ID NO:46), HK1 (SEQ ID NO:47), CYFIP2 (SEQ ID NO:48), DCI (SEQ ID NO:49) and CISD1 (SEQ ID NO:50). Each of the aforementioned proteins represents a separate embodiment of the invention and may be used independently from or in combination with any of the others.

In particular embodiments, the kit is for detecting pre-cancerous polyps or early stage colorectal cancer and comprises at least one reagent suitable for detecting the level of at least one protein selected from the group consisting of CPT2 (SEQ ID NO:21), ARL1 (SEQ ID NO:22), PFKL (SEQ ID NO:23), GOT2 (SEQ ID NO:24), AP1G1 (SEQ ID NO:25), STRBP (SEQ ID NO:26), CLCA1 (SEQ ID NO:27), CYFIP1 (SEQ ID NO:28), COQ9 (SEQ ID NO:29), NDUFA9 (SEQ ID NO:30), ALDH7A1 (SEQ ID NO:31), HMGCS1 (SEQ ID NO:32), NNT (SEQ ID NO:33), PRDX5 (SEQ ID NO:34), PCCB (SEQ ID NO:35), COPZ1 (SEQ ID NO:36), BAX (SEQ ID NO:37), ACAD9 (SEQ ID NO:38), UBXD8 (SEQ ID NO:39), HMGCS2 (SEQ ID NO:40), SLC25A3 (SEQ ID NO:41), SLC25A11 (SEQ ID NO:42), PDCD6 (SEQ ID NO:43), UCRC (SEQ ID NO:44), DEFA6 (SEQ ID NO:45), DYNC1H1 (SEQ ID NO:46), HK1 (SEQ ID NO:47), CYFIP2 (SEQ ID NO:48), DCI (SEQ ID NO:49) and CISD1 (SEQ ID NO:50). In particular embodiments, the kit comprises a reagent suitable for detecting the level of KIAA0152 (SEQ ID NO:1).

According to some embodiments the present invention excludes proteins known to be associated with colorectal cancers. According to other embodiments the methods of the present invention exclude proteins that were known to be associated with other types of cancers.

Other objects, features and advantages of the present invention will become clear from the following description and examples.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 illustrates methods used for the identification of the protein diagnostic markers of cancer according to the invention.

Figure 2 illustrates quantitative data analysis of protein markers identified in diseased and non-diseased tissues from human subjects with colorectal cancer, including patients with polyps. Sample numbers indicated along the top of the figure correspond to different patients, and in some cases different samples from a single patient. Proteins are indicated along the left side. The height of each rectangle corresponds to the relative signal intensity, and the shading corresponds to the quantitative ratio (protein in diseased tissue:protein in non-diseased tissue) from the same patient. Black shaded rectangles, ratio >10; white rectangles, ratio in the range 3-10; gray shaded rectangles, ratio in the range 0.3-3; diamond filled rectangles, ratio <0.3.

DETAILED DESCRIPTION OF THE INVENTION

The present invention provides methods of identification of cancerous cells by detection of levels of particular proteins, also referred to herein as "cancer-associated proteins or "molecular markers", which have been found to be differentially expressed in cancer tissue, in particular, colorectal cancer tissue. Details of the subject cancer-associated proteins disclosed herein are provided in Tables 1 and 3-7. Some of the identified proteins for which significantly elevated levels were detected in colorectal cancer tissue as compared to normal colorectal tissue have not been previously disclosed to be associated with cancerous disease. Examples of such proteins are indicated in Table 1 as denoted by the symbol "+" in the column labeled "Newly identified as cancer-associated" and include the proteins FAM62B, SLC1A5. RSL1D1, LYZ, TNC, RBM39, ILVBL, ERO1L, LOC442497, TCOF1, NAMPT, SERPINB9 and HSDL2. Yet other proteins, in particular THBS1 and LMO7, have been disclosed in the prior art to be down-regulated in cancerous conditions, whereas the inventors of the present invention now disclose significantly increased levels of these proteins in cancer tissue, indicative of up-regulation. These proteins are indicated in Table 1 as denoted by the symbol "UR" in the column labeled "Newly identified as cancer-associated". Yet other proteins were detected at significant levels in colorectal cancer tissue but were completely absent from healthy colorectal tissue in the same patients. Such proteins are indicated in Table 1 as denoted by the symbol "+" in the column labeled "Detected in cancerous but not healthy" and include the proteins AMACR, AMACR;C1QTNF3, COL5A1, FKBP10, FKBP9, GPRC5A, POLR1C, SPARC and UBAP2.

Table 3 list proteins which are highly expressed in colorectal cancer, both in pre-cancerous polyps and in more advanced stages of the disease. Accordingly, any of the proteins listed in Table 3 represent candidates for use as diagnostic markers of colorectal cancer, which can be indicative of any of pre-cancerous lesions, early stage or late stage forms of the disease. Two currently preferred protein markers are KIAA0152 and NAMPT.

A quantitative analysis of some of the proteins listed in Table 3 is shown in Figure 2. The analysis shows that in a large number of the patients examined, the quantitative ratio of protein in diseased tissue (i.e. polyps and/or cancer): protein in non-diseased tissue from the same patient, is 3 or greater.

Table 4 lists proteins which appear to be highly expressed in pre-cancerous polyps, yet their expression tends to be diminished in more advanced stages of the disease. Accordingly, proteins listed in Table 4 represent promising candidates for use as very early diagnostic markers to identify individuals at risk of developing colorectal cancer.

Identification and quantification of a panel of tumor-associated proteins such as those listed in Tables 1 and 3-6, provides a specific means of detecting colorectal cancer in a subject, as well as means of prognosing and staging previously diagnosed colorectal cancers. Such methods can conveniently be carried out using detectably-labeled reagents which specifically bind the selected target proteins. Typically such reagents comprise monoclonal antibodies or small chemical mimetics thereof. In some cases however, it may be advantageous to employ nucleic acids complementary to mRNA species encoding the target proteins so as to quantitate expression activity which may be correlated with the corresponding protein levels. Furthermore, therapeutic pharmaceutical agents directed against tumor-associated proteins, for example, those disclosed in Table 7 herein, may be prepared, for example monoclonal antibodies conjugated to cytotoxic moieties, for use in targeted therapy of colorectal cancer.

Table 1.

Protein Name	IPI Acc. No.	Newly identified as cancer-associated	Detected in cancerous but not healthy
ABCF2 ATP-binding cassette, sub-family F, member 2 isoform b	68506		
ACIN1 Isoform 1 of Apoptotic chromatin condensation inducer in the nucleus	7334		
ACOT7 Isoform 1 of Cytosolic acyl coenzyme A thioester hydrolase	10415		

Protein Name	IPI Acc. No.	Newly identified as cancer-associated	Detected in cancerous but not healthy
ADAMDEC1 ADAM DEC1 precursor	4480		+
ADO 2-aminoethanethiol dioxygenase	45939		
ADSS Adenylosuccinate synthetase isozyme 2	26833		
AK3 GTP:AMP phosphotransferase mitochondrial	465256		
ALG5 Dolichyl-phosphate beta-glucosyltransferase	2506		
AMACR Alpha-methylacyl-CoA racemase	847727		+
AMACR;C1QTNF3 alpha-methylacyl-CoA racemase isoform 1	5918		+
ANXA3 Annexin A3	24095		
ARHGEF1 Isoform 1 of Rho guanine nucleotide exchange factor 1	647786		
ARID1A Isoform 1 of AT-rich interactive domain-containing protein 1A	643722		+
ARRDC1 Arrestin domain containing 1	514937		
ATAD3A Isoform 2 of ATPase family AAA domain-containing protein 3A	295992		
ATP6AP2 Renin receptor precursor	168884		
ATP6V0A1 97 kDa protein	892784		
ATP6V1E1 vacuolar H+ ATPase E1 isoform b	719806		
BMS1 Ribosome biogenesis protein BMS1 homolog	6099		
BPI Bactericidal/permeability-increasing protein	552280		
BUD31 Protein BUD31 homolog	13180		
BXDC1 Brix domain containing 1	644504		
C14orf21 Pumilio domain-containing protein C14orf21	216999		
C1orf116 Isoform 1 of Specifically androgen-regulated gene protein	28392		
C1R;ACYP1;C17orf13 Complement C1r subcomponent precursor	296165		
C20orf43 UPF0549 protein C20orf43	297121		
C2orf47 Uncharacterized protein C2orf47, mitochondrial precursor	291751		
C3orf64 Isoform 1 of Uncharacterized glycosyltransferase AER61 precursor	396231		

Protein Name	IPI Acc. No.	Newly identified as cancer-associated	Detected in cancerous but not healthy
C7orf24 Uncharacterized protein C7orf24	31564		
C8orf55 Uncharacterized protein C8orf55 precursor	171421		
CAD Putative uncharacterized protein CAD	893035		
CASP8 Uncharacterized protein CASP8	220725		
CBFB Core-binding factor subunit beta	16746		
CEACAM5 Carcinoembryonic antigen-related cell adhesion molecule 5 precursor	27486		
CEBPZ CCAAT/enhancer-binding protein zeta	306723		+
CHD4 Isoform 2 of Chromodomain-helicase-DNA-binding protein 4	455210		
CLPB Isoform 2 of Caseinolytic peptidase B protein homolog	216192		
COL12A1 Isoform 1 of Collagen alpha-1(XII) chain precursor	329573		
COL5A1 Collagen alpha-1(V) chain precursor	844090		+
COMT Isoform Soluble of Catechol O-methyltransferase	375513		
CSTF1 Cleavage stimulation factor 50 kDa subunit	11528		
CTSG Cathepsin G precursor	28064		
DCK Deoxycytidine kinase	20454		
DDX18 ATP-dependent RNA helicase DDX18	301323		
DEK 48 kDa protein	871695		
DHCR7 7-dehydrocholesterol reductase	294501		
DHODH Dihydroorotate dehydrogenase, mitochondrial precursor	24462		
DHX30 DEAH (Asp-Glu-Ala-His) box polypeptide 30 isoform 2	477295		
DIAPH1 Diaphanous homolog 1	884341		
DMBT1 Isoform 1 of Deleted in malignant brain tumors 1 protein precursor	99110		
DNAJA3 Isoform 2 of DnaJ homolog subfamily A member 3, mitochondrial precursor	179187		
DNAJC19 10 kDa protein	795263		
DPEP1 Dipeptidase 1 precursor	59476		

Protein Name	IPI Acc. No.	Newly identified as cancer-associated	Detected in cancerous but not healthy
EFEMP2 Mutant p53 binding protein 1 variant (Fragment)	556657		+
EI24 Isoform 2 of Etoposide-induced protein 2.4 homolog	23185		
EIF2B3 Isoform 1 of Translation initiation factor eIF-2B subunit gamma	6504		
EIF2S2 Eukaryotic translation initiation factor 2 subunit 2	21728		
ERO1L ERO1-like protein alpha precursor	386755	+	
EXOC2 Exocyst complex component 2	783559		
F11R Junctional adhesion molecule A precursor	1754		
FAM62B Isoform 2 of Extended synaptotagmin-2	409635	+	
FAM84B Protein FAM84B	64666		+
FAP Isoform 1 of Seprase	295461		
FASN Fatty acid synthase	26781		
FASTKD5 FAST kinase domain-containing protein 5	414973		
FDFT1 Squalene synthetase	20944		
FERMT1 Isoform 1 of Fermitin family homolog 1	304754		
FKBP10 FK506-binding protein 10 precursor	303300		+
FKBP9 FK506-binding protein 9 precursor	182126		+
FOXK1 Isoform 1 of Forkhead box protein K1	556645		
FRG1 Protein FRG1	4655		
FYB FYN binding protein (FYB-120/130) isoform 1	73110		
GCA Grancalcin	4524		
GEMIN5 Gem-associated protein 5	291783		
GGH Gamma-glutamyl hydrolase precursor	23728		
GLRX3 Glutaredoxin-3	8552		
GLT25D1 Glycosyltransferase 25 family member 1 precursor	168262		
GMPS GMP synthase	29079		
GNL3 Isoform 1 of Guanine nucleotide-binding protein-like 3	306380		
GPR89B;GPR89A Isoform 1 of Protein GPR89	472858		

Protein Name	IPI Acc. No.	Newly identified as cancer-associated	Detected in cancerous but not healthy
GPRC5A Retinoic acid-induced protein 3	22624		+
GPX2 Glutathione peroxidase 2	298176		
GTF2F1 General transcription factor IIF subunit 1	17450		
GTF2I Isoform 2 of General transcription factor II-I	293242		
HCFC1 Uncharacterized protein HCFC1	641743		
HM13 Isoform 1 of Minor histocompatibility antigen H13	152441		
HSDL2 Isoform 1 of Hydroxysteroid dehydrogenase-like protein 2	414384	+	
HSPH1 Isoform Beta of Heat shock protein 105 kDa	218993		
ICAM1 Intercellular adhesion molecule 1 precursor	8494		
IGFBP7 Insulin-like growth factor-binding protein 7 precursor	16915		
IKIP IKIP2	401791		
ILVBL Isoform 1 of Acetolactate synthase-like protein	554541		
IPO7 Importin-7	7402		
ISLR Immunoglobulin superfamily containing leucine-rich repeat protein precursor	23648		
KIAA0020 Pumilio domain-containing protein KIAA0020	791325		
KIAA0241 Isoform 1 of Protein KIAA0241	397348		
KIAA1219 Isoform 3 of Protein KIAA1219	410120		
KPNA2 Karyopherin alpha 2	789457		+
LACTB2 Beta-lactamase-like protein 2	6952		
LCN2 Lipocalin 2	643623		
LEPRE1 Isoform 3 of Prolyl 3-hydroxylase 1 precursor	45839		
LMO7 Isoform 3 of LIM domain only protein 7	291802	UR	
LOC442497;SLC3A2 solute carrier family 3 (activators of dibasic and neutral amino acid transport), member 2 isoform e	554722	+	
LOC731605 similar to BCL2-associated transcription factor 1 isoform 2	886854		
LTBP2 Latent-transforming growth	292150		

Protein Name	IPI Acc. No.	Newly identified as cancer-associated	Detected in cancerous but not healthy
factor beta-binding protein 2 precursor			
LTF Lactotransferrin precursor	848342		
LYZ Lysozyme C precursor	19038	+	
MAN1A2 Uncharacterized protein MAN1A2	743100		
MCM2 DNA replication licensing factor MCM2	184330		
MCM3 DNA replication licensing factor MCM3	13214		
MCM5 DNA replication licensing factor MCM5	18350		
MCM6 DNA replication licensing factor MCM6	31517		
MCM7 Isoform 1 of DNA replication licensing factor MCM7	299904		
MMP1 Interstitial collagenase precursor	8561		+
MMP2 72 kDa type IV collagenase precursor	27780		
MMP9 Matrix metalloproteinase-9 precursor	27509		
MOCS3 Molybdenum cofactor synthesis protein 3	4489		
MPO Isoform H7 of Myeloperoxidase precursor	236556		
MRPS6 Mitochondrial 28S ribosomal protein S6	305668		
MTA2 Metastasis-associated protein MTA2	171798		
MUC16 Mucin-16	103552		
NAMPT Isoform 1 of Nicotinamide phosphoribosyltransferase	18873	+	
NAT10 N-acetyltransferase 10	300127		
NCBP1 Nuclear cap-binding protein subunit 1	19380		
NCF4 Isoform 1 of Neutrophil cytosol factor 4	14338		
NEBL Nebulette variant 4	872370		
NIP7 Isoform 1 of 60S ribosome subunit biogenesis protein NIP7 homolog	7175		
NOC2L Nucleolar complex protein 2 homolog	411886		
NOC4L Nucleolar complex protein 4 homolog	31661		
NOL10 Isoform 1 of Nucleolar protein 10	29513		

Protein Name	IPI Acc. No.	Newly identified as cancer-associated	Detected in cancerous but not healthy
NQO1 NAD	12069		
NUDCD1 Isoform 2 of NudC domain-containing protein 1	306398		
NUP188 Isoform 1 of Nucleoporin NUP188 homolog	477040		
NUP210 Isoform 1 of Nuclear pore membrane glycoprotein 210 precursor	291755		
NXF1 Nuclear RNA export factor 1	33153		
OCIAD2 Isoform 1 of OCIA domain-containing protein 2	555902		
OLFM4 Olfactomedin-4 precursor	22255		
OSBPL2 Isoform 1 of Oxysterol-binding protein-related protein 2	14137		
OTUD6B OTU domain containing 6B	182180		
PAICS Multifunctional protein ADE2	217223		
PARP14 poly (ADP-ribose) polymerase family, member 14	291215		
PCNA Proliferating cell nuclear antigen	21700		
PEX14 Peroxisomal membrane protein PEX14	25346		
PHF6 Isoform 1 of PHD finger protein 6	395568		
PKP3 Plakophilin-3	26952		
PLOD3 Procollagen-lysine,2-oxoglutarate 5-dioxygenase 3 precursor	30255		
PNMA5 Paraneoplastic antigen-like protein 5	514588		+
POLR1C Isoform 1 of DNA-directed RNA polymerases I and III subunit RPAC1	5179		+
PROM1 Prominin-1 precursor	12540		
PSAT1 Isoform 1 of Phosphoserine aminotransferase	1734		
PTDSS1 Phosphatidylserine synthase 1	10746		
PTK7 Tyrosine-protein kinase-like 7 precursor	298292		
PUF60 Isoform 5 of Poly	856076		
PYCR1 pyrroline-5-carboxylate reductase 1 isoform 2	376503		
PYCR2 Pyrroline-5-carboxylate reductase 2	470610		
RAD21 Double-strand-break repair protein rad21 homolog	6715		

Protein Name	IPI Acc. No.	Newly identified as cancer-associated	Detected in cancerous but not healthy
RANBP3 Isoform 2 of Ran-binding protein 3	456728		
RBM12 RNA-binding protein 12	550308		
RBM39 Isoform 2 of RNA-binding protein 39	215801	+	
RCC1 regulator of chromosome condensation 1 isoform b	787306		
RCC2 Protein RCC2	465044		
RCN1 Reticulocalbin-1 precursor	15842		
RDBP Negative elongation factor E	858		
REEP6 Receptor expression-enhancing protein 6	647161		
REG1A Lithostathine-1-alpha precursor	9027		
REG1B Lithostathine-1-beta precursor	9197		
RFC2 Isoform 2 of Replication factor C subunit 2	218280		
RFC5 Replication factor C subunit 5	31514		
RRM1 Ribonucleoside-diphosphate reductase large subunit	13871		
RRP1 RRP1-like protein	550766		
RRS1 Ribosome biogenesis regulatory protein homolog	14253		
RSL1D1 RSL1D1 protein	642046	+	
S100A11 Protein S100-A11	13895		
S100A12 Protein S100-A12	218131		
S100A8 Protein S100-A8	7047		
S100A9 Protein S100-A9	27462		
SAE1 SUMO-activating enzyme subunit 1	33130		
SDCCAG10 Isoform 1 of Peptidyl-prolyl cis-trans isomerase SDCCAG10	25174		
SERPINB5 Serpin B5 precursor	783625		
SERPINB9 Serpin B9	32139	+	
SERPINH1 Serpin H1 precursor	32140		
SET Isoform 2 of Protein SET	301311		
SLC1A5 Neutral amino acid transporter B	19472	+	
SLC25A15 Mitochondrial ornithine transporter 1	3389		
SLC2A1 Solute carrier family 2,	220194		

Protein Name	IPI Acc. No.	Newly identified as cancer-associated	Detected in cancerous but not healthy
facilitated glucose transporter member 1			
SORD 11 kDa protein	791243		
SPARC SPARC precursor	14572		+
SQSTM1 Isoform 1 of Sequestosome-1	179473		
SRM Spermidine synthase	292020		
SRRM2 Isoform 1 of Serine/arginine repetitive matrix protein 2	782992		
SSBP1 Single-stranded DNA-binding protein, mitochondrial precursor	29744		
SYK Isoform Long of Tyrosine-protein kinase SYK	18597		
TBC1D2B Isoform 1 of TBC1 domain family member 2B	550733		
TCOF1 Isoform 2 of Treacle protein	298696	+	
TFRC Transferrin receptor protein 1	22462		
TH1L Isoform NELF-D of Negative elongation factor C/D	759539		
THBS1 Thrombospondin-1 precursor	296099	UR	
THBS2 Thrombospondin-2 precursor	18769		
THOC6 Isoform 1 of THO complex subunit 6 homolog	328985		
TIMP1 TIMP metallopeptidase inhibitor 1	642739		
TJP2 Isoform A1 of Tight junction protein ZO-2	3843		
TM9SF4 Isoform 2 of Transmembrane 9 superfamily member 4 precursor	885106		
TNC Isoform 1 of Tenascin precursor	31008	+	
TOMM34 Mitochondrial import receptor subunit TOM34	9946		
TOP1 DNA topoisomerase 1	413611		
TOP2A Isoform 2 of DNA topoisomerase 2-alpha	414101		
TPR nuclear pore complex-associated protein TPR	742682		
TRMT6 Isoform 1 of tRNA	99311		
UBAP2 Ubiquitin-associated protein 2	171127		+
UBE2O Ubiquitin-conjugating enzyme E2 O	783378		
UCK2 Isoform 1 of Uridine-cytidine kinase 2	65671		+

Protein Name	IPI Acc. No.	Newly identified as cancer-associated	Detected in cancerous but not healthy
UQCC 34 kDa protein	872061		
URB1 Nucleolar pre-ribosomal-associated protein 1	297241		
UTP20 Small subunit processome component 20 homolog	4970		
VAMP7 Isoform 3 of Vesicle-associated membrane protein 7	401804		
WDR43 WD repeat-containing protein 43	55954		
WDR74 Isoform 1 of WD repeat-containing protein 74	18192		+
XPO5 Isoform 1 of Exportin-5	640703		
XPOT Exportin-T	306290		
YLPM1 YLP motif containing 1	165434		

Definitions

The terms "subject" and "patient" as used herein refer to any single subject for which cancer detection, prognosis, staging or therapy is desired, including humans and non-human mammals, such as primate, bovine, ovine, canine, feline and rodent mammals. Also included are subjects involved in clinical research trials not showing any clinical sign of disease, or subjects involved in epidemiological studies, or subjects used as controls.

The terms "cancer detection" and "cancer diagnosis" and related grammatical terms, such as "detecting cancer" and "diagnosing cancer", respectively, are used herein interchangeably to refer to any of: determination of a subject's susceptibility to a malignant cancer disease; determination as to whether a subject is presently affected by a malignant cancer disease; determination of a subject's stage of cancer, determination of and monitoring the effect on the cancer in response to anti-cancer therapy.

The term "characteristics of a cancerous or pre-cancerous growth" as used herein refers to one or more molecular, physiological, histological, clinical or other properties which may be used to define the nature and behavior of the growth.

The terms "cancer", "neoplasm", "tumor", "growth" and the like are used interchangeably herein to refer to cells which exhibit relatively autonomous growth, so that they exhibit an aberrant growth phenotype characterized by a significant loss of control of cell proliferation. In general, cells of interest for detection or treatment in the present

application include pre-malignant (e.g., benign hyperplasic), malignant, metastatic, and non-metastatic cells.

The term "prognosis" as used herein refers to the expected or predicted outcome of a disease, such as a cancer, in a patient following diagnosis. A prognosis may predict the relative chance of disease progression, arrest or cure. A prognosis may be established on the basis of prognostic indicators specific for a particular disease. Prognostic indicators in cancer may include for example, the grade and stage of cancer at initial diagnosis, the genetic make-up of the patient, the presence and level of cancer-associated antigens in the tumor, and patient responsiveness to a particular therapy.

The terms "biological sample" and "test sample" as used herein encompass a variety of types of biological materials that can be used in the methods of the invention. The sample may be procured from the body of an individual or investigated without removal from the body of an individual. The term encompasses solid tissue samples, such as from biopsy specimens, tumors or tumor metastases, or tissue cultures or cells derived there from and the progeny thereof. The terms also encompass blood and other liquid samples of biological origin, such as nipple aspirate fluid, lymph node aspirate, mucosal fluid, cervical wash, lacrimal duct fluid, urine, saliva, pleural effusion and sputum. The terms encompass samples that have been manipulated in any way after their procurement, such as by lysis, treatment with reagents, solubilization, or enrichment for certain components. Also included are clinical samples, cells in cell culture, cell supernatants and cell lysates. It is to be explicitly understood that in accordance with the invention, a biological or test sample may be obtained i.e. removed, from the body of a subject, or accessed *in vivo*, for example by contacting with a specific reagent or apparatus.

As used herein, a biological or test sample that is "obtained from a subject", means that the sample is removed from the body of the subject, and any subsequent analysis thereof may be performed outside the body for example under *in vitro* or *ex vivo* conditions. When however, a biological or test sample is "assessed *in vivo*" or "accessed *in vivo*" it means that the sample is maintained within the body of the subject and direct or indirect contact is established using any suitable means, such as via a reagent, composition, device or apparatus. Subsequent analysis of the accessed sample is performed under *in vivo* conditions, which can include conditions of local or general anaesthesia. Means of accessing and assessing a sample *in vivo* include for example, contacting the tissue or organ or region thereof of interest with a pharmaceutical composition, a swallowable endoscopy capsule or an endoscopy probe.

The term "a normal biological sample of the same type" as used herein refers to a non-diseased sample consisting of the same biological material or type of tissue and/or cells e.g. blood, colorectal tissue, as that of the test sample. The normal biological sample may be that from a single individual, including the subject in which cancer detection, prognosing or characterizing is performed, or from a group of individuals of known healthy status. Accordingly, the level of a protein in a normal biological sample may be obtained from a single determination or may advantageously represent a statistical average of multiple determinations, such as from a group of healthy individuals or from multiple healthy tissue sites in a single individual.

The term "a control sample" as used herein refers to the standard provided by either a normal i.e. non-diseased sample or group of samples, or a sample or group of samples corresponding to an established form, type, stage or grade of a disease, in particular a cancer disease. Accordingly, the level of a protein in a control sample may be obtained from a single determination or may advantageously represent a statistical average of multiple determinations, for example from a group of healthy individuals, or from a group of diseased individuals established to have the same form, type, stage or grade of a cancer disease.

The term "non-diseased tissue" as used herein refers to tissue which is determined to be free of a cancer disease, on the basis of any technique known in the art, for example, histological and immuno-histochemical investigation. It is to be understood that diseased and non-diseased tissue of the same tissue type may be present in close proximity in a individual patient.

The term "immunogenic fragment" as used herein in reference to a protein refers to a portion of the protein which is capable of inducing an immune response, such as antibody production, following administration to an individual.

As used herein, the terms "a protein associated with cancer", "tumor associated protein", "molecular marker" and the like, interchangeably refer to a protein that is present at relatively higher or lower levels in a cancer cell relative to a normal cell of the same type (e.g., as in protein associated with colon cancer).

As used herein the terms "nucleic acid encoding a protein," encompass polynucleotides and polypeptides respectively having sequence similarity or sequence identity to the respective gene and gene products having the accession numbers of the particular protein referred to, of at least about 65%, preferably at least about 80%, more preferably at least

about 85%, and can be about 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more. Sequence similarity and sequence identity are calculated based on a reference sequence, which may be a subset of a larger sequence, such as a conserved motif, coding region, flanking region, etc. A reference sequence will usually be at least about 18 nt long, more usually at least about 30 nt long, and may extend to the complete sequence that is being compared. In general, percent sequence identity is calculated by counting the number of residue matches (e.g., nucleotide residue or amino acid residue) between the query and test sequence and dividing total number of matches by the number of residues of the individual sequences found in the region of strongest alignment, as is known in the art. Algorithms for computer-based sequence analysis are known in the art, such as BLAST (see, e.g., Altschul et al., J. Mol. Biol., 215:403-10 (1990)), particularly the Smith-Waterman homology search algorithm as implemented in MPSRCH program (Oxford Molecular).

The terms "protein" and "polypeptide" are used interchangeably herein to refer to polymeric forms of amino acids of any length, which can include coded and non-coded amino acids, chemically or biochemically modified or derivatized amino acids, and polypeptides having modified peptide backbones. The term includes fusion proteins, including, but not limited to, fusion proteins with a heterologous amino acid sequence, fusions with heterologous and homologous leader sequences, with or without N-terminal methionine residues; immunologically tagged proteins; and the like.

As used herein, a "fusion protein" or "chimeric peptide" refers to a protein or polypeptide which comprises at least a portion of a first naturally occurring protein or polypeptide fused to least a portion of a second protein or polypeptide. For example, a fusion protein for targeting FAM62B may include a portion or all of an anti-FAM62B antibody fused with another peptide or polypeptide such as a protein label moiety or a cytotoxic protein.

The term "antibody" as used herein is used in the broadest sense and specifically encompasses monoclonal antibodies, humanized antibodies, polyclonal antibodies, multispecific antibodies (e.g., bispecific antibodies), single chain antibodies and antibody fragments (e.g., $F(ab')_2$, Fab', Fab, Fv) so long as they bind specifically to a target antigen or epitope of interest.

The term "monoclonal antibody" as used herein refers to an antibody obtained from a population of substantially homogeneous antibodies, i.e., the individual antibodies comprising the population are identical except for possible naturally occurring mutations that

may be present in minor amounts. Monoclonal antibodies are highly specific, being directed against a single antigenic site. Furthermore, in contrast to conventional (polyclonal) antibody preparations that typically include different antibodies directed against different determinants (epitopes), each monoclonal antibody is directed against a single determinant on the antigen. The modifier "monoclonal" indicates the character of the antibody as being obtained from a substantially homogeneous population of antibodies, and is not to be construed as requiring production of the antibody by any particular method. For example, the monoclonal antibodies to be used in accordance with the present invention may be made by the hybridoma method first described by Kohler et al., Nature 256:495 (1975), or may be made by recombinant DNA methods (see, e.g., U.S. Pat. No.4,816,567). The "monoclonal antibodies" may also be isolated from phage antibody libraries, as is known in the art, for example using techniques such as those described in Clackson et al., Nature 352:624-628 (1991) and Marks et al., J. Mol. Biol. 222:581-597 (1991).

Furthermore, monoclonal antibodies specifically include "chimeric" antibodies (immunoglobulins) in which a portion of the heavy and/or light chain is identical with or homologous to corresponding sequences in antibodies derived from a particular species or belonging to a particular antibody class or subclass, while the remainder of the chain(s) is identical with or homologous to corresponding sequences in antibodies derived from another species or belonging to another antibody class or subclass, as well as fragments of such antibodies, so long as they exhibit the desired biological activity (see for example U.S. Pat. No. 4,816,567; and Morrison et al., Proc. Natl. Acad. Sci. USA 81:6851-6855 (1984)).

"Humanized" forms of non-human (e.g., murine) antibodies are chimeric antibodies that contain minimal sequence derived from non-human immunoglobulin. For the most part, humanized antibodies are human immunoglobulins (recipient antibody) in which residues from a hypervariable region of the recipient are replaced by residues from a hypervariable region of a non-human species (donor antibody) such as mouse, rat, rabbit or non-human primate having the desired specificity, affinity, and capacity. In some instances, Fv framework region (FR) residues of the human immunoglobulin are replaced by corresponding non-human residues. Furthermore, humanized antibodies may comprise residues that are not found in the recipient antibody or in the donor antibody. These modifications are made to further refine antibody performance. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the hypervariable loops correspond to those of a non-

human immunoglobulin and all or substantially all of the FRs are those of a human immunoglobulin sequence. The humanized antibody optionally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin. For further details, see Jones et al., Nature 321:522-525 (1986); Riechmann et al., Nature 332:323-329 (1988); and Presta, Curr. Op. Struct. Biol. 2:593-596 (1992).

The term "reagent suitable for detecting the level of a protein" as used herein refers to a reagent which either specifically binds the protein itself or is complementary to a nucleic acid molecule that is involved in expression of the protein. Particularly suitable examples of reagents which specifically bind to proteins are antibodies and chemical mimetics thereof having similar level of binding affinity and/or avidity. Nucleic acid reagents which are complementary to mRNA encoding a protein of interest are suitable for determining the level of the protein, typically by correlating the amount of mRNA detected with the corresponding level of protein product produced in a suitable translation system.

The terms "specifically interacts" and "specifically binds" are used herein interchangeably to refer to high avidity and/or high affinity binding of a reagent, such as an antibody to a specific polypeptide or epitope thereof, such as for example, an epitope of any of the proteins FAM62B, SLC1A5, RSL1D1, LYZ, RBM39 and TCOF1. Antibody binding to its epitope is stronger than binding of the same antibody to any other epitope, particularly those which may be present in molecules in association with, or in the same sample, as the specific polypeptide of interest. Antibodies which bind specifically to a polypeptide of interest may be capable of binding other polypeptides at a weak, yet detectable, level (e.g., 10% or less of the binding shown to the polypeptide of interest). Such weak binding, or background binding, is readily discernible from the specific antibody binding to the compound or polypeptide of interest, e.g., by use of appropriate controls.

The term "primary antibody" as used herein refers to an antibody which binds specifically to the target protein antigen in a biological or test sample. A primary antibody is generally the first antibody used in an immunoassay procedure. In some embodiments, the primary antibody is the only antibody used in an immunoassay procedure.

The term "secondary antibody" as used herein refers to an antibody which binds specifically to a primary antibody, thereby forming a bridge between the primary antibody and a subsequent reagent, if any. The secondary antibody is typically directed against the Fc portion of the immunoglobulin type of the primary antibody (e.g., anti-mouse Fc antibody).

The terms "polynucleotide" and "nucleic acid" are used interchangeably herein to refer to polymeric forms of nucleotides of any length, either ribonucleotides or deoxynucleotides, including but are not limited to, single-, double-, or multi-stranded DNA or RNA, genomic DNA, cDNA, DNA-RNA hybrids, or a polymer comprising purine and pyrimidine bases or other natural, chemically or biochemically modified, non-natural, or derivatized nucleotide bases. Further includee are mRNA or cDNA that comprise intronic sequences (see, e.g., Niwa et al. (1999) Cell 99(7):691-702). The backbone of the polynucleotide can comprise sugars and phosphate groups (as typically be found in RNA or DNA), or modified or substituted sugar or phosphate groups. Alternatively, the backbone of the polynucleotide can comprise a polymer of synthetic subunits such as phosphoramidites and thus can be an oligodeoxynucleoside phosphoramidate or a mixed phosphoramidate-phosphodiester oligomer. Peyrottes et al. (1996) Nucl. Acids Res. 24:1841-1848; Chaturvedi et al. (1996) Nucl. Acids Res. 24:2318-2323. A polynucleotide may comprise modified nucleotides, such as methylated nucleotides and nucleotide analogs, uracyl, other sugars, and linking groups such as fluororibose and thioate, and nucleotide branches. The sequence of nucleotides may be interrupted by non-nucleotide components. A polynucleotide may be further modified after polymerization, such as by conjugation with a labeling component, capping, substitution of one or more of naturally occurring nucleotides with an analog, and introduction of means for attaching the polynucleotide to proteins, metal ions, labeling components, other polynucleotides, or a solid support.

The term "complementary" as used herein refers to the broad concept of subunit sequence complementarity between two nucleic acids, e.g., two DNA molecules or a DNA/RNA hybrid. When a nucleotide position in both of the molecules is occupied by nucleotides normally capable of base pairing with each other, then the nucleic acids are considered to be complementary to each other at this position. Thus, two nucleic acids are complementary to each other when a substantial number (at least 50%) of corresponding positions in each of the molecules are occupied by nucleotides which normally base pair with each other (e.g., A:T and G:C nucleotide pairs).

The term "hybridization" as used herein encompasses any process by which a strand of nucleic acid joins with a complementary strand through base pairing (see for example, Coombs, Dictionary of Biotechnology, Stockton Press, New York N.Y. (1994)). Accordingly, a hybridization assay is a quantitative means of determining the extent of hybridization between a nucleic acid in a test sample, such as a tissue sample, and a nucleic

acid probe corresponding to at least a fragment of a gene (DNA or RNA) encoding a protein of interest, such as a cancer associated protein. "Stringency" typically occurs in a range from about Tm -5° C (5° C below the Tm of the nucleic acid probe) to about 20° C to 25° C below Tm. As will be understood by those of skill in the art, a stringency hybridization can be used to identify or detect identical polynucleotide sequences or to identify or detect similar or related polynucleotide sequences.

Amplification as carried out in the polymerase chain reaction technologies is described in Dieffenbach et al., PCR Primer, a Laboratory Manual, Cold Spring Harbor Press, Plainview N.Y. (1995), and may be performed prior to or as part of a hybridization assay.

As used herein, the term "differentially expressed" generally refers to a polynucleotide and/or the corresponding protein that is expressed at levels in a test cell that differ significantly from levels in a reference or control cell, e.g., a cancer-associated protein as disclosed herein is found at levels at least about 50% to about 100% increased, generally at least about 1.5-fold, at least about 2-fold, at least about 3-fold, at least about 5-fold, at least about 10-fold, at least about 20-fold, or at least about 30-fold or more increased in a cancerous cell when compared with a cell of the same type that is not cancerous. The comparison can be made between two tissues, for example, if one is using *in situ* hybridization or another assay method that allows some degree of discrimination among cell types in the tissue. The comparison may also be made between cells removed from their tissue source, or cells maintained in their native state *in vivo*. "Differential expression" refers to both quantitative, as well as qualitative, differences in the genes' temporal and/or cellular expression patterns among, for example, normal and neoplastic tumor cells, and/or among tumor cells which have undergone different tumor progression events.

The terms "correspond to" or "represents" as used in, for example, the phrase "polynucleotide corresponds to a differentially expressed gene" are used to refer to the relationship between a given polynucleotide and the gene from which the polynucleotide sequence is derived (e.g., a polynucleotide that is derived from a coding region of the gene, a splice variant of the gene, an exon, and the like) or to which the polynucleotide hybridizes to under stringent conditions.

The term "label" as used herein refers to a compound or composition which is conjugated or fused directly or indirectly to a reagent such as an antibody, a nucleic acid probe or a chemical agent and facilitates detection of the reagent to which it is conjugated or fused. The label may itself be detectable (e.g., radioisotope labels or fluorescent labels) or, in

the case of an enzymatic label, may catalyze chemical alteration of a substrate compound or composition which is detectable.

The term "mimetic" as used herein refers to any entity, including natural and synthesized inorganic or organic molecules, including recombinant molecules, that mimic the properties of the molecule of which it is a mimetic. Accordingly, a mimetic of a particular antibody has the same, similar or enhanced epitope binding properties of that antibody.

The term "gene" as used herein refers to any nucleic acid sequence or portion thereof with a functional role in encoding or transcribing a protein or regulating other gene expression. The gene may consist of all the nucleic acids responsible for encoding a functional protein or only a portion of the nucleic acids responsible for encoding or expressing a protein. The nucleic acid sequence may contain a genetic abnormality within exons, introns, initiation or termination regions, promoter sequences, other regulatory sequences or unique adjacent regions to the gene.

The singular forms "a", "and", and "the" include plural referents unless the context clearly dictates otherwise. Thus, for example, reference to "a cell" includes a plurality of such cells and reference to "the reagent" includes reference to one or more reagents and equivalents thereof known to those skilled in the art, and so forth.

Where a range of values is provided, it is understood that each intervening value, to the tenth of the unit of the lower limit unless the context clearly dictates otherwise, between the upper and lower limits of that range is also specifically disclosed. Each smaller range between any stated value or intervening value in a stated range and any other stated or intervening value in that stated range is encompassed within the invention. The upper and lower limits of these smaller ranges may independently be included or excluded in the range, and each range where either, neither or both limits are included in the smaller ranges is also encompassed within the invention, subject to any specifically excluded limit in the stated range. Where the stated range includes one or both of the limits, ranges excluding either or both of those included limits are also included in the invention.

Detection, prognosis and management of colorectal cancer

The present invention is based on the discovery that specific proteins, including for example, KIAA0152, NAMPT, PYCR1, GPX2, PRKDC, ALDH18A1, OCIAD2, GCS1, GMDS, ARF4, ARF5, LRPPRC, CTNNB1, ARF3, GCN1L1, BDH1, RPL9, UGCGL1, FAM3D and CCT4, CPT2, ARL1, PFKL, GOT2, AP1G1, STRBP, CLCA1, CYFIP1, COQ9,

NDUFA9, ALDH7A1, HMGCS1, NNT, PRDX5, PCCB, COPZ1, BAX, ACAD9, UBXD8, HMGCS2, SLC25A3, SLC25A11, PDCD6, UCRC, DEFA6, DYNC1H1, HK1, CYFIP2, DCI, CISD1, FAM62B, S100A8, S100A9, MPO, MCM2, LTF, OLFM4, FERMT1, CEACAM5, SLC1A5, THBS1, NAT10, RSL1D, LMO7, LYZ and MCM3 are present at significantly higher levels in cancerous or pre-cancerous colorectal tissue as compared to normal tissue of the same cell type, even in the same individual..

This discovery serves as the basis for identification of cancerous tissue, as well as for staging and prognosing tumors, and development of therapeutic agents which target the particular cancer-associated markers. Detection of the markers disclosed herein and/or the genes encoding them, enables early diagnosis based on molecular changes leading to carcinogenesis and/or decision making in disease management. For example, a relatively increased level of a particular protein, such as FAM62B, compared to normal cells or tissues of the same type can be indicative of a poorer prognosis, and therefore warrant more aggressive therapy (e.g., chemotherapy following surgery). The correlation of tumor specific markers with response to treatment and outcome in patients can define prognostic indicators that allow the design of tailored therapy based on the molecular profile of the tumor. These therapies include antibody targeting, antagonists (e.g., small molecules), and gene therapy. Determining colon cancer-specific protein levels and comparison of a patient's profile with known levels in normal tissue and variants of the disease allows a determination of optimal treatment strategies. The marker expression pattern can also be used to better classify (i.e. stage and grade), and thus diagnose and treat different forms of cancer. Furthermore, a protein identified as being differentially or specifically expressed in, one type of cancer may also have implications for development or risk of development of other types of cancer.

Colorectal cancer is one of the most common neoplasms in humans and perhaps the most frequent form of hereditary neoplasia. Prevention and early detection are key factors in controlling and curing colorectal cancer. Colorectal cancer begins as polyps, which are small, benign growths of cells that form on the inner lining of the colon. Over a period of several years, some of these polyps accumulate additional mutations and become cancerous. Multiple familial colorectal cancer disorders have been identified, as follows: 1) Familial adenomatous polyposis (FAP); 2) Gardner's syndrome; 3) Hereditary nonpolyposis colon cancer (HNPCC); and 4) Familial colorectal cancer in Ashkenazi Jews.

Staging is a process used in the medical arts to describe how advanced the cancerous state is in a patient. While staging systems vary with the types of cancer, they generally

involve the "TNM" system: "T" indicates the type of tumor, "N" indicates whether the cancer has metastasized to nearby lymph nodes; and "M" indicates whether the cancer has metastasized to other parts of the body. Generally, if a cancer is only detectable in the area of the primary lesion without having spread to any lymph nodes it is called Stage I. If it has spread only to the closest lymph nodes, it is called Stage II. In Stage III, the cancer has generally spread to the lymph nodes in near proximity to the site of the primary lesion. Cancers that have spread to a distant part of the body, such as the liver, bone, brain or other site, are Stage IV, the most advanced stage.

The grade of a cancer describes how closely a tumor resembles normal tissue of its same type. The microscopic appearance of a tumor is used to identify tumor grade based on parameters such as cell morphology, cellular organization, and other markers of differentiation. As a general rule, the grade of a tumor corresponds to its rate of growth or aggressiveness, with undifferentiated or high-grade tumors generally being more aggressive than well differentiated or low-grade tumors. The following guidelines are generally used for grading tumors: GX, Grade cannot be assessed; G1, Well differentiated; G2, Moderately well differentiated; G3, Poorly differentiated; G4, Undifferentiated.

Methods of detection, prognosis, and characterization as disclosed herein are based on levels of at least one, and preferably a panel or multiplicity of cancer-associated proteins, in comparison to levels of the same protein(s) in suitable non-cancerous or cancerous control samples. For example, a detection of cancer may be enabled by the detection of a level of one or more proteins of interest in a sample, such as FAM62b, that is increased at least by 50% or 100% or greater or, alternatively by 2-fold, 3-fold 5-fold, 10-fold, 30-fold, or greater, relative to a normal non-cancerous sample of the same tissue type. The normal non-cancerous sample may be from the same individual as the test sample or from one or more different individuals. Preferably, the level in the normal non-cancerous sample is a statistical average of multiple determinations, such as from a group of healthy individuals or from multiple healthy tissue sites in a single individual.

Similarly, prognosis and characterization of the cancer may be established on the levels of proteins of interest relative to reference standards established for particular grades, stages and forms of cancer. For example, detection of a level of a particular protein which is 1.5-fold compared to that of the control may be taken to indicate a relatively positive prognosis and/or relatively non-aggressive type of cancer, whereas detection of a level of the same protein of 10-fold or more compared to that of the control may be taken to indicate a poorer

prognosis and/or a substantially aggressive type of cancer.

Assay systems and methods

The methods of the invention may be carried out using various assay systems and methods for detection of the level of the protein of interest in a test biological sample. Suitable systems include those employing an immunoassay, a nucleic acid hybridization assay, a binding assay, an array, a phage display library, or a combination thereof.

Immunoassays

Immunoassays for detecting levels of specific binding between an antibody and its target antigen are known in the art and include for example, radioimmunoassay, (RIA), fluorescent immunoassay, (FIA) enzyme-linked immunosorbant assay (ELISA), immunohistochemistry (IHC) and fluorescent activated cell sorting (FACS) (see, e.g., Harlow and Lane, Using Antibodies, A Laboratory Manual, Cold Spring Harbor Laboratory, New York (1999)).. In general, an immunoassay may be either direct or indirect. In a direct assay, the binding of antibody to the target antigen is determined directly using a labeled reagent, such as a fluorescent labeled or an enzyme-labeled primary antibody, which can be detected without further antibody interaction. In a typical indirect assay, unconjugated primary antibody binds to the antigen and then a labeled secondary antibody binds to the primary antibody. Signal amplification occurs because several secondary antibodies may react with different epitopes on the primary antibody. Alternately, both the primary and secondary antibodies may be unlabeled and labeled tertiary antibody is employed. Where the antibody (primary, secondary or tertiary) is conjugated to an enzymatic label, a chromagenic or fluorogenic substrate is added to provide detection of the antigen.

The primary antibody used for detection of the protein(s) of interest is stably associated with e.g., directly or indirectly bound to, the surface of a solid support, e.g. column, microtiter plate, beads, membrane, typically made of glass, plastic, polysaccharides, nylon or nitrocellulose. A multiplicity of antibody specificities for detecting a panel of tumor-associated proteins may be simultaneously bound to the same support, such as in an array.

The test sample is allowed to contact the support during a period of incubation, generally following blocking of non-specific binding sites with non-interfering proteins such as bovine serum albumin. After incubation with each reagent e.g. blocking agent, primary antibody, secondary antibody, the support is washed to remove non-bound components. Determination of suitable reagents, conditions for washing, incubation etc. is within the

ability of one of average skill in the art.

Immunoassays can also be employed histologically, as in immunofluorescence or immunoelectron microscopy, for *in situ* detection of tumor associated antigen. In situ detection can be accomplished by removing a histological sample from a subject, and contacting the sample with a labeled antibody. The antibody is typically contacted with the sample by overlaying the labeled antibody onto the sample. Through the use of such a procedure, the presence of the tumor associated antigen can be determined and/or the distribution of the antigen in the histological sample can be examined. Those of ordinary skill in the art will readily appreciate that any of a wide variety of histological methods (such as staining procedures) can be modified in order to achieve such in situ detection.

Detectable labels suitable for conjugation to antibodies and other binding reagents include radioisotopes, fluorescent labels, enzyme-substrate labels, chromogenic labels, chemiluminescent labels and colloidal gold particles.

Radioisotopes include for example, ^{35}S, ^{14}C, ^{125}I, ^{3}H, ^{32}P and ^{131}I. Fluorescent labels include for example, fluorescent molecules fluorescein isothiocyanate (FITC), rhodamine, phycoerythrin (PE), phycocyanin, allophycocyanin, ortho-phthaldehyde, fluorescamine, peridinin-chlorophyll a (PerCP), Cy3 (indocarbocyanine), Cy5 (indodicarbocyanine), lanthanide phosphors, and the like.

Enzymatic labels include luciferases (e.g. firefly luciferase and bacterial luciferase), luciferin, 2,3-dihydrophthalazinediones, malate dehydrogenase, urease, peroxidase such as horseradish peroxidase (HRPO), alkaline phosphatase, β-galactosidase, glucoamylase, lysozyme, saccharide oxidases (e.g., glucose oxidase, galactose oxidase, and glucose-6-phosphate dehydrogenase), heterocyclic oxidases (such as uricase and xanthine oxidase), lactoperoxidase, microperoxidase, and the like. Examples of enzyme-substrate combinations include, for example: horseradish peroxidase (HRPO) with hydrogen peroxidase as a substrate, wherein the hydrogen peroxidase oxidizes a dye precursor (e.g., orthophenylene diamine (OPD) or 3,3',5,5'-tetramethyl benzidine hydrochloride (TMB)); alkaline phosphatase (AP) with para-nitrophenyl phosphate as chromogenic substrate; and β-D-galactosidase β-D-Gal) with a chromogenic substrate (e.g., p-nitrophenyl-β-D-galactosidase) or fluorogenic substrate (e.g., 4-methylumbelliferyl-β-D-galactosidase).

A label may be indirectly conjugated with an antibody or other reagent, as is known in the art. For example, an antibody can be conjugated with biotin and any of the types of labels

mentioned above can be conjugated with avidin, or vice versa. Biotin binds selectively to avidin and thus, the label can be conjugated with the antibody in an indirect manner. In some cases, detectable labels are attached by spacer arms of various lengths to reduce potential steric hindrance.

Detection of bound, labeled antibody can be carried out by standard colorimetric, radioactive, photometric and/or fluorescent detection means. For fluorescent labels, signal can be detected by, for example, a scanning confocal microscope in photon counting mode. Appropriate scanning devices are described by, for example, U.S. Pat. Nos. 5,578,832 and 5,631,734. For antibodies labeled with biotin, the reaction can be treated with the appropriate streptavidin-conjugate (e.g., streptavidin-horseradish peroxidase, streptavidin-alkaline phosphatase, streptavidin-luciferase, and the like) and then treated with the appropriate reagents for calorimetric or photometric detection. For radiolabeled antibody, signal can be detected using a scintillation counter, phosphoimager or similar device.

Arrays

Cancer associated proteins in a sample may be detected using an array-based binding assay system. Such an array-based system may incorporate an immunoassay as described above, or incorporate small molecule chemical entities which specifically interact with particular cancer associated proteins. In either case, the solid substrate used for the array comprises a plurality of binding reagents attached to the substrate, wherein each binding reagent has specificity for a different cancer associated protein. The protein panel or "set" for which the array is predetermined to specifically bind and detect is characteristic of a particular disease, or stage of form of a disease. The binding reagents, are immobilized onto the substrate surface, preferably in a spatially addressable manner. The binding reagents may be antibodies, antibody fragments or small molecule chemical entities.

The nature and geometry of the solid substrate will depend upon a variety of factors, including, among others, the type of array (e.g., one-dimensional, two-dimensional or three-dimensional). Generally, the surface can be composed of any material which will permit immobilization of the binding reagents and which will not substantially degrade under the conditions used in the applications of the array.

The solid substrate used for the array may be in the form of beads, particles or sheets, and may be permeable or impermeable, depending on the type of array, wherein the surface is coated with a suitable material enabling binding of the binding reagents at high affinity.

For example, for linear or three-dimensional arrays the surface may be in the form of beads or particles, fibers (such as glass wool or other glass or plastic fibers) or glass or plastic capillary tubes. For two-dimensional arrays, the solid surface may be in the form of plastic, micromachined chips, membranes, slides, plates or sheets in which at least one surface is substantially flat, wherein these surfaces may comprise glass, plastic, silicon, low cross-linked and high cross-linked polystyrene, silica gel, polyamide, and the like.

Fluorescence tagged beads are also an addressable (liquid) array in which each bead is tagged with a different set of fluorescent colors and bound with an antibody; specific to an array is detected with devices such as fluorescence scanners for arrays or FACS for beads.

The arrays used for the present invention may be of any desired size. The upper and lower limits on the size of the array are determined solely by the practical considerations of resolution, size of molecules expressed at each address and the like.

Either a population of discrete proteins is employed to form the array, such that each address presents a different molecule, or a single or a few addresses are employed with a similar protein. In many applications, redundancies in the spots are desirable for the purposes of acting as internal controls.

Technologies for the deposition of droplets containing protein binding reagents onto a suitable solid surface are known in the art. An ink-jet printing technology for deposition of small droplets while avoiding overlap or splatter is disclosed in U.S. Patent No. 5,449,754.

In order to conduct array-based binding assays, the test sample is allowed to contact the array comprising a coated surface containing the anchored binding reagents. Following contact, the array is optionally washed, typically under conditions such that any complexes formed will remain immobilized on the solid surface and unbound material will be removed.

The detection of complexes anchored on the solid surface can be accomplished in a number of ways. In some embodiments, the non-immobilized sample is pre-labeled, and the detection is directed to label immobilized on the surface indicating that complexes were formed. In other embodiments, an indirect label can be used to detect complexes anchored on the surface; e.g., using a labeled antibody specific for the previously non-immobilized sample (the antibody, in turn, may be directly labeled or indirectly labeled with a labeled anti-Ig antibody). In another embodiment, the immobilized molecules of the microarray are labeled, the array can be scanned or otherwise analyzed for detectable assay signal, and the signal from each labeled spot, or alternatively from all spots, quantified.

An important consideration is the presence of an amount of a label at each position within the array that is proportional to the amount of molecule immobilized at that particular spot. Thus, it is important that the efficiencies of the coupling reactions which are used to immobilize the labeled molecules are substantially similar.

Virtually any label that produces a detectable, quantifiable signal and that is capable of being attached to an immobilized binding reagent on a substrate can be used in conjunction with the array of the invention. Suitable labels include: radioisotopes, fluorophores, chromophores, chemiluminescent moieties, as described above.

Preferably, the position of the label will not interfere with interaction between a desired sample and the immobilized molecules and with the detection in case of an interaction between the desired sample and an immobilized molecule of the array. Suitable methods of making labeled molecules are well known in the art.

In the case where each spot in the array contains an amount of a label or "tracer" proportional to the amount of molecules immobilized at the particular spot, the signals obtained from the arrays of the invention can be normalized. As a consequence, signal intensities from spots within a single array, or across multiple arrays, can be directly compared. A normalized signal of a particular spot may be defined by $(I_t - I_o)/I_o$, where I_t is the intensity of the signal of the spot after contacting with a sample of interest and I_o is the intensity of the background signal of the spot before contacting with a sample of interest.

Various methods and devices for detection and analysis of the array are known in the art. Practically, any imaging system that is capable of detecting with a resolution appropriate to the size of the array features can be utilized. For example, a method for screening an array of proteins for interactions with a fluid sample is disclosed in U.S. Patent No. 6,475,809. Imaging apparatus may be selected, for example, from ScanArray 4000 (General Scanning), Biochip Imager (Hewlett Packard), GMS 418 Array Scanner (Genetic Microsystems), GeneTAC 1000 (Genomic Solutions), Chip Reader (Virtek). Phosphorimager systems are available for detecting radiolabels, e.g. Cyclone (Packard Instrument Co.) and BAS-5000 (Fujifilm).

Hybridization assays

Hybridization assays generally comprise contacting a sample containing nucleic acids (target nucleic acids) with a nucleic acid probe capable of hybridizing to tumor associated antigen nucleic acids, under conditions such that hybridization can occur, and detecting or

measuring any resulting hybridization.

Suitable hybridization assays include, for example, Northern blots, dot blots, RT-PCR, and quantitative PCR. Such procedures can be performed *in situ* directly for example, in tissue sections (e.g., fixed and/or frozen) of subject tissue obtained from biopsies or resections, such that no nucleic acid purification is necessary. Tumor associated antigen nucleic acids can be used as probes and/or primers for such procedures (see e.g., Nuovo, PCR In Situ Hybridization: Protocols and Applications, Raven Press, NY (1992)).

Detection of tumor associated antigen nucleic acids typically involves contacting and incubating nucleic acids from a test sample with one or more labeled nucleic acids, (i.e. "probes") under conditions favorable for the specific annealing of the nucleic acids to their complementary sequences. Typically, the lengths of the nucleic acid reagents are at least 15 to 30 nucleotides. After incubation, all non-annealed nucleic acids are removed. The presence of bound i.e. hybridized, nucleic acids from the sample, if any such molecules exist, is then detected. Using such a detection scheme, the nucleic acid from the tissue or cell type of interest can be immobilized, to a solid support such as a membrane, or a plastic surface such as that on a microtiter plate or polystyrene beads.

Nucleic acid arrays can be used to monitor the expression of tumor associated genes, such as, for example, those corresponding to KIAA0152, NAMPT, PYCR1, GPX2, PRKDC, ALDH18A1, OCIAD2, GCS1, GMDS, ARF4, ARF5, LRPPRC, CTNNB1, ARF3, GCN1L1, BDH1, RPL9, UGCGL1, FAM3D and CCT4, CPT2, ARL1, PFKL, GOT2, AP1G1, STRBP, CLCA1, CYFIP1, COQ9, NDUFA9, ALDH7A1, HMGCS1, NNT, PRDX5, PCCB, COPZ1, BAX, ACAD9, UBXD8, HMGCS2, SLC25A3, SLC25A11, PDCD6, UCRC, DEFA6, DYNC1H1, HK1, CYFIP2, DCI, CISD1, FAM62B, S100A8, S100A9, MPO, MCM2, LTF, OLFM4, FERMT1, CEACAM5, SLC1A5, THBS1, NAT10, RSL1D, LMO7, LYZ and/or MCM3. For detection of a multiplicity of genes encoding distinct cancer associated proteins, an array of polynucleotide probes may be contacted with a sample of target nucleic acids to produce a hybridization pattern. The binding of the target nucleic acids to one or more probes of the array is then detected to obtain a qualitative and/or quantitative profile of expression of the tumor associated antigen gene.

A variety of different arrays can be used and are known in the art. The polymeric or probe molecules of the arrays can be polynucleotides or hybridizing derivatives or analogs thereof, including: nucleic acids in which the phosphodiester linkage has been replaced with a substitute linkage, such as phophorothioate, methylimino, methyl-phosphonate,

phosphoramidate, guanidine, and the like; nucleic acids in which the ribose subunit has been substituted, for example, hexose phosphodiester; peptide nucleic acids; and the like. The length of the probes will generally range from about 10 to about 1000 nucleotides, typically, from about 15 to about 150 nucleotides in length, but also possibly from about 150 to about 1000 nucleotides in length. The probes can be single or double stranded, usually single stranded, and can be PCR fragments amplified from cDNA. The probe molecules on the surface of the substrates will typically correspond to at least one of the tumor associated antigen genes and be positioned on the array at known locations so that positive hybridization events can be correlated to expression of a particular gene in the physiological source from which the target nucleic acid sample is derived. Because of the manner in which the target nucleic acid sample is generated, the arrays of probes will generally have sequences that are complementary to the non-template strands of the gene to which they correspond.

The substrate for the array can be fabricated from a variety of materials, including plastics, ceramics, metals, gels, membranes, glasses, and the like. The arrays can be produced according to any convenient methodology, such as pre-forming the probes and then stably associating them with the surface of the support or growing the probes directly on the support. A number of different array configurations and methods for their production are known to those of skill in the art and disclosed in, for example, U.S. Pat. Nos. 5,445,934; 5,532,128; 5,556,752; 5,242,974; 5,384,261; 5,405,783; 5,412,087; 5,424,186; 5,429,807; 5,436,327; 5,472,672; 5,527,681; 5,529,756; 5,545,531; 5,554,501; 5,561,071; 5,571,639; 5,593,839; 5,599,695; 5,624,711; 5,658,734; and 5,700,637.

The target nucleic acid is typically contacted with the array under conditions sufficient for hybridization of target nucleic acid to probe to occur. Suitable hybridization conditions are well known to those of skill in the art (see, e.g., Sambrook et al., Molecular Cloning, A Laboratory Manual, 3rd ed., Cold Spring Harbor, Cold Spring Harbor, N.Y. (2001)).

The amount of tumor associated antigen nucleic acids in the sample can be quantitated (see, e.g., U.S. Pat. No. 6,004,755). For example, the target nucleic acids in the sample can be end-labeled in a manner such that each of the target nucleic acids in the sample produces a signal of the same specific activity. By generating the same specific activity is meant that each individual target polynucleotide in the sample being assayed is labeled in a manner such that the molecule is capable of providing the same signal (e.g., the same intensity of signal) as every other labeled target in the sample. Each of the target nucleic acids generates a signal of the same specific activity because the number of labeled nucleotide bases in each of the

target molecules is either identical or substantially the same.

The label is capable of providing a detectable signal, either directly or through interaction with one or more additional members of a signal producing system. Suitable detectable labels include radioactive, fluorescent and enzymatic labels as described above.

In some applications, it is desired to analyze populations of target nucleic acids from two or more samples. Such samples can be differentially labeled. Alternatively, target nucleic acids from different samples are separately contacted to identical probe arrays under conditions of hybridization, typically stringent hybridization conditions, such that labeled nucleic acids hybridize to their complementary probes on the substrate surface, and the target nucleic acids bound to the array separately detected. A set of standard nucleic acid molecules can optionally be used. For example, the standard nucleic acids can be provided by reverse transcribing standard RNA.

Following hybridization, a washing step is usually employed to remove non-specifically bound nucleic acid from the support surface, generating a pattern of hybridized nucleic acid on the substrate surface. Various wash solutions and protocols for their use are known to those of skill in the art.

Where the label on the target nucleic acid is not directly detectable, the array can be contacted with the other member(s) of the signal producing system that is being employed. For example, where the label on the target is biotin, the array can be contacted with streptavidin-fluorophore conjugate under conditions sufficient for binding between the specific binding member pairs to occur. Following contact, any unbound members of the signal producing system will then be removed (e.g., by washing).

The resultant hybridization pattern(s) of target nucleic acids bound to the array can be visualized or detected in a variety of ways, with the particular manner of detection being chosen based on the particular label of the nucleic acid. For example, detection means can include scintillation counting, autoradiography, fluorescence measurement, colorimetric measurement, light emission measurement, and the like.

Prior to detection or visualization, the array of hybridized target/probe complexes can be optionally treated with an endonuclease, for example, mung bean nuclease, S1 nuclease, and the like. The endonuclease degrades single stranded, but not double stranded DNA.

Following detection or visualization, the hybridization pattern can be used to determine qualitative and/or quantitative information about the expression of tumor associated antigen

genes. The hybridization patterns of different samples can be compared with each other, or with a control sample, to identify differences between the patterns. The hybridization arrays can also be used to identify differential gene expression, in the analysis of diseased and normal tissue.

Antibody production

Antibodies directed against cancer associated proteins include polyclonal antibodies, monoclonal antibodies, chimeric antibodies, single chain antibodies, antibody fragments (e.g., Fab, Fab', $F(ab')_2$, Fv, or hypervariable regions), bi-specific antibodies and humanized antibodies, methods of production of which are known in the art.

For the production of polyclonal antibodies, a host animal (e.g., rabbits, mice, rats, sheep, goats, and the like) can be immunized by injection with a tumor associated antigen, fragment, derivative or analog. Various adjuvants can be used to increase the immunological response, depending on the host species. Such adjuvants include, for example, Freund's adjuvant (complete and incomplete), mineral gels such as aluminum hydroxide, surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanins, dinitrophenol, and other adjuvants, such as BCG (bacille Calmette-Guerin) and Corynebacterium parvum.

Techniques for preparation of monoclonal antibodies include the original Kohler and Milstein hybridoma technique (see, e.g., Nature 256:495 97 (1975)), the trioma technique (see, e.g., Hagiwara and Yuasa, Hum. Antibodies Hybridomas 4:15 19 (1993); Hering et al., Biomed. Biochim. Acta 47:211 16 (1988)), the human B-cell hybridoma technique (see, e.g., Kozbor et al., Immunology Today 4:72 (1983)), and the EBV-hybridoma technique (see, e.g., Cole et al., In: Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, Inc., pp. 77 96 (1985)). Human antibodies can be obtained using human hybridomas (see, e.g., Cote et al., Proc. Natl. Acad. Sci. USA 80:2026 30 (1983)) or by transforming human B cells with EBV virus *in vitro* (see, e.g., Cole et al., supra).

Chimeric antibodies are typically prepared by splicing the genes (of one species) for an antibody molecule specific for tumor associated antigen together with genes from another species of antibody molecule of appropriate biological activity. It can be desirable to transfer the antigen binding regions (e.g., Fab', $F(ab')_2$, Fab, Fv, or hypervariable regions) of antibodies from one species into the framework of an antibody from another species by recombinant DNA techniques to produce a chimeric molecule. Methods for producing such

molecules are described in, for example, U.S. Pat. Nos. 4,816,567; 4,816,397; 5,693,762, and 5,712,120. A human monoclonal antibody or portion(s) thereof can be identified by screening a human B-cell cDNA library for nucleic acid molecules that encode antibodies that specifically bind to a tumor associated antigen according to the method generally set forth by Huse et al. (Science 246:1275 81 (1989)). The nucleic acid molecule can then be cloned and amplified to obtain sequences that encode the antibody (or antigen-binding domain) of the desired specificity. Phage display technology offers another technique for selecting antibodies that bind to tumor associated antigens, fragments, derivatives or analogs thereof (see, e.g., International Patent Publications WO 91/17271 and WO 92/01047; Huse et al., supra.)

Techniques for the production of single chain antibodies are described for example in U.S. Pat. Nos. 4,946,778 and 5,969,108. A Fab expression library (see, e.g., Huse et al., supra) allows rapid and easy identification of monoclonal Fab fragments with the desired specificity for tumor associated antigens, fragments, derivatives, or analogs thereof.

$F(ab')_2$ antibody fragments can be produced by pepsin digestion of an antibody molecule. Fab' fragments can be generated by reducing the disulfide bridges of a $F(ab')_2$ fragment, Fab and Fv fragments can be generated by treating an antibody molecule with papain and a reducing agent. Recombinant Fv fragments can also be produced in eukaryotic cells using, for example, as described in U.S. Pat. No. 5,965,405.

Bi-specific antibodies can be monoclonal antibodies that have binding specificities for at least two different antigens. For example, one of the binding specificities can be for a tumor associated antigen and the other one is for any other antigen. Alternatively, one specificity is for a first tumor associated antigen, while the other specificity is for a second, different tumor associated antigen. Methods for making bi-specific antibodies may be based on the co-expression of two immunoglobulin heavy-chain/light-chain pairs, where the two heavy chains have different specificities (see, e.g., Milstein and Cuello, Nature 305:537 39 (1983)). Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas produce a potential mixture of different antibody molecules, some of which have the desired bi-specific structure. The purification of the correct molecule is usually accomplished by affinity chromatography steps.

Antibody variable domains with the desired binding specificities (antibody-antigen combining sites) can be fused to immunoglobulin constant domain sequences. The fusion typically is with an immunoglobulin heavy-chain constant domain, comprising at least part of

the hinge, CH2, and CH3 regions. The first heavy-chain constant region (CH1) containing the site necessary for light-chain binding is usually present in at least one of the fusions. DNAs encoding the immunoglobulin heavy-chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors, and are co-transfected into a suitable host organism. For further details of generating bi-specific antibodies see, for example, Suresh et al (Methods in Enzymology 121:210 (1986)).

To generate a phage antibody library, a cDNA library is first obtained from mRNA which is isolated from cells, e.g., the hybridoma, which expresses the desired protein to be expressed on the phage surface, e.g., the desired antibody. cDNA copies of the mRNA are produced using reverse transcriptase. cDNA which specifies immunoglobulin fragments are obtained by PCR and the resulting DNA is cloned into a suitable bacteriophage vector to generate a bacteriophage DNA library comprising DNA specifying immunoglobulin genes. The procedures for making a bacteriophage library comprising heterologous DNA are well known in the art and are described, for example, in Sambrook et al., supra.

Pharmaceutical compositions and agents

A pharmaceutical or antigen composition according to the invention comprises at least one protein selected from the group consisting of KIAA0152, NAMPT, PYCR1, GPX2, PRKDC, ALDH18A1, OCIAD2, GCS1, GMDS, ARF4, ARF5, LRPPRC, CTNNB1, ARF3, GCN1L1, BDH1, RPL9, UGCGL1, FAM3D, CCT4, CPT2, ARL1, PFKL, GOT2, AP1G1, STRBP, CLCA1, CYFIP1, COQ9, NDUFA9, ALDH7A1, HMGCS1, NNT, PRDX5, PCCB, COPZ1, BAX, ACAD9, UBXD8, HMGCS2, SLC25A3, SLC25A11, PDCD6, UCRC, DEFA6, DYNC1H1, HK1, CYFIP2, DCI, CISD1, LYZ, LOC442497;SLC3A2, DMBT1, NUCB1, GGH, AGR3, TM9SF2, SYK, GCA, HDLBP, C1QBP and CLIC1, or an immunogenic fragment thereof.

The composition may be used in a therapeutic method for preventing or treating cancer, wherein an effective amount of the composition is administered to a subject in need thereof. A subject in need thereof may be for example, an individual in which colorectal polyps have been identified, and/or has been determined to have other risk factors for development of colorectal cancer. Accordingly, a method of preventing colorectal cancer may comprise administering an effective amount of a pharmaceutical composition comprising at least one protein highly expressed in polyps, for example any of CPT2, ARL1, PFKL, GOT2, AP1G1, STRBP, CLCA1, CYFIP1, COQ9, NDUFA9, ALDH7A1, HMGCS1, NNT, PRDX5, PCCB, COPZ1, BAX, ACAD9, UBXD8, HMGCS2, SLC25A3, SLC25A11, PDCD6, UCRC,

DEFA6, DYNC1H1, HK1, CYFIP2, DCI and CISD1.

Further provided are pharmaceutical compositions comprising one or more reagents suitable for detecting the level of at least one protein selected from the group consisting of KIAA0152, NAMPT, PYCR1, GPX2, PRKDC, ALDH18A1, OCIAD2, GCS1, GMDS, ARF4, ARF5, LRPPRC, CTNNB1, ARF3, GCN1L1, BDH1, RPL9, UGCGL1, FAM3D, CCT4, CPT2, ARL1, PFKL, GOT2, AP1G1, STRBP, CLCA1, CYFIP1, COQ9, NDUFA9, ALDH7A1, HMGCS1, NNT, PRDX5, PCCB, COPZ1, BAX, ACAD9, UBXD8, HMGCS2, SLC25A3, SLC25A11, PDCD6, UCRC, DEFA6, DYNC1H1, HK1, CYFIP2, DCI, CISD1, LYZ, LOC442497;SLC3A2, DMBT1, NUCB1, GGH, AGR3, TM9SF2, SYK, GCA, HDLBP, C1QBP and CLIC1.

As hereinbefore described, such reagents may comprise labeled antibodies, antibody mimetics or nucleic acids as the active ingredients. Advantageously such reagents are provided as a "cocktail" tailored to detect *in vivo* a particular panel of cancer associated proteins, or genes expressing same, Following administration, the localization of such reagents to specific body and organ regions may be detected by means of an external monitoring apparatus or system appropriate for detection and quantitation of the specific label incorporated into the reagents.

In yet other embodiments, pharmaceutical agents, in particular antibodies and an antibody mimetics, are provided, which may be used for targeted therapy of cancers expressing particular cancer associated proteins. Such agents may advantageously incorporate a cytotoxic moiety, such as a plant toxin, a bacterial toxin, a radioactive atom or a chemotherapeutic agent. Further, the pharmaceutical agent may be in the form of a chemical conjugate or a fusion protein and provided in a suitably formulated pharmaceutical composition.

Plant toxins include for example, ricin, abrin, pokeweed antiviral protein, saporin, gelonin, and derivatives thereof. Bacterial toxins include for example, Pseudomonas exotoxin, diphtheria toxin and derivatives thereof, iodine-125 and radon. Radioactive atoms include for example, radium, cesium-137, iridium-192, americium-241 and gold-198.

Chemotherapeutic agents include, without limitation, alkylating agents, for example cyclophosphamide; nitrosoureas, for example carmustine and lomustine; antimetabolites, for example 5-fluorouracil, capecitabine, 6-mercaptopurine, methotrexate, gemcitabine, cytarabine, fludarabine and pemetrexed; anthracyclines, for example daunorubicin,

doxorubicin respinomycin D and idarubicin; topoisomerase inhibitors, for example topotecan, irinotecan, etoposide and teniposide; mitotic inhibitors, for example paclitaxel, docetaxel, etoposide, vinblastine and vincristine; platinum based drugs, for example cisplatin, carboplatin and oxaliplatin; steroids for example hydrocortisone, dexamethasone, methylprednisolone and prednisolone; and anti-angiogenic agents, for example bevacizumab, thalidomide, dopamine and tetrathiomolybdate.

The pharmaceutical composition can be used to administer diagnostic reagents designed to detect levels of proteins disclosed herein, in particular KIAA0152 (SEQ ID NO:1), NAMPT (SEQ ID NO:2), PYCR1 (SEQ ID NO:3), GPX2 (SEQ ID NO:4), PRKDC (SEQ ID NO:5), ALDH18A1 (SEQ ID NO:6), OCIAD2 (SEQ ID NO:7), GCS1 (SEQ ID NO:8), GMDS (SEQ ID NO:9), ARF4 (SEQ ID NO:10), ARF5 (SEQ ID NO:11), LRPPRC (SEQ ID NO:12), CTNNB1 (SEQ ID NO:13), ARF3 (SEQ ID NO:14), GCN1L1 (SEQ ID NO:15), BDH1 (SEQ ID NO:16), RPL9 (SEQ ID NO:17), UGCGL1 (SEQ ID NO:18), FAM3D (SEQ ID NO:19) and CCT4 (SEQ ID NO:20), including antibodies, antibody mimetics or nucleic acids, or therapeutic reagents designed to exert a cytotoxic affect on their cancer targets. Therapeutic compositions for targeting tumors may include a reagent which specifically interacts with a protein selected from NAMPT, PYCR1, GPX2, PRKDC, ALDH18A1, OCIAD2, GCS1, GMDS, ARF4, ARF5, LRPPRC, CTNNB1, ARF3, GCN1L1, BDH1, RPL9, KIAA0152, UGCGL1, FAM3D, CCT4, LYZ, LOC442497;SLC3A2, DMBT1, NUCB1, GGH, AGR3, TM9SF2, SYK, GCA, HDLBP, C1QBP, KIAA0152 and CLIC1.

The pharmaceutical composition may be prepared by suspending the desired reagent in any pharmaceutically acceptable carrier, for example, HEPES buffered saline at a pH of about 7.8. Other pharmaceutically acceptable carriers which are useful include, but are not limited to, glycerol, water, saline, ethanol and other pharmaceutically acceptable salt solutions such as phosphates and salts of organic acids. Examples of these and other pharmaceutically acceptable carriers are described in Remington's Pharmaceutical Sciences (1991, Mack Publication Co., New Jersey).

The pharmaceutical compositions may be prepared in the form of a sterile injectable aqueous or oily suspension or solution. This suspension or solution may be formulated according to the known art, and may comprise, in addition to the active ingredient, additional ingredients such as the dispersing agents, wetting agents, or suspending agents. Sterile injectable formulations may be prepared using a non-toxic parenterally-acceptable diluent or

solvent, such as water or 1,3-butane diol, for example. Other acceptable diluents and solvents include, but are not limited to, Ringer's solution, isotonic sodium chloride solution, and fixed oils such as synthetic mono- or di-glycerides.

Pharmaceutical compositions that are useful in the methods of the invention may be administered, prepared, packaged, and/or sold in formulations suitable for oral, rectal, vaginal, parenteral, topical, pulmonary, intranasal, buccal, ophthalmic, or another route of administration. Other contemplated formulations include projected nanoparticles, liposomal preparations, resealed erythrocytes containing the active ingredient, and immunologically-based formulations.

As used herein, the term "physiologically acceptable" ester or salt means an ester or salt form of the active ingredient which is compatible with any other ingredients of the pharmaceutical composition, which is not deleterious to the subject to which the composition is to be administered.

The formulations of the pharmaceutical compositions described herein may be prepared by any method known or hereafter developed in the art of pharmacology. In general, such preparatory methods include the step of bringing the active ingredient into association with a carrier or one or more other accessory ingredients, and then, if necessary or desirable, shaping or packaging the product into a desired single- or multi-dose unit.

Controlled- or sustained-release formulations of a pharmaceutical composition of the invention may be made using conventional technology.

A formulation of a pharmaceutical composition of the invention suitable for oral administration may be prepared in the form of a discrete solid dose unit including, but not limited to, a tablet, a hard or soft capsule, a cachet, a troche, or a lozenge, each containing a predetermined amount of the active ingredient. Other formulations suitable for oral administration include, but are not limited to, a powdered or granular formulation, an aqueous or oily suspension, an aqueous or oily solution, or an emulsion. As used herein, an "oily" liquid is one which comprises a carbon-containing liquid molecule and which exhibits a less polar character than water.

As used herein, "additional ingredients" include, but are not limited to, one or more of the following: excipients; surface active agents; dispersing agents; inert diluents; granulating and disintegrating agents; binding agents; lubricating agents; sweetening agents; flavoring agents; coloring agents; preservatives; physiologically degradable compositions such as

gelatin; aqueous vehicles and solvents; oily vehicles and solvents; suspending agents; dispersing or wetting agents; emulsifying agents, demulcents; buffers; salts; thickening agents; fillers; emulsifying agents; antioxidants; antibiotics; antifungal agents; stabilizing agents; and pharmaceutically acceptable polymeric or hydrophobic materials. Other "additional ingredients" which may be included in the pharmaceutical compositions of the invention are known in the art.

Kits

The invention also includes kits for detecting, diagnosing, prognosing or staging a cancer or a tumor in a mammal. The cancer or tumor can be of any of the types described herein, and is preferably colorectal cancer. The kit comprises a container or a sample tube, or the like, for storing a sample of a cell, a population of cells, a tissue or a body fluid obtained from the mammal.

The kit also comprises one or more detection reagents selected from: an antibody or antibody mimetic which specifically bind with a cancer associated protein disclosed herein; a nucleic acid such as an oligonucleotide which specifically binds a nucleic acid (such as mRNA) encoding said cancer associated protein, and a PCR primer pair specific for a nucleic acid encoding said cancer associated protein. These detection reagents are as described herein. The kit comprises the one or more detection reagents in an amount effective to permit detection of the protein(s) of interest or a corresponding nucleic acid in the sample. Detection of the proteins or the nucleic acids is accomplished using any of the methods described herein or known to a skilled artisan for detecting a specific protein or specific nucleic acid molecule within a biological sample.

Reagents in the kit may be directed to a protein selected from KIAA0152, NAMPT, PYCR1, GPX2, PRKDC, ALDH18A1, OCIAD2, GCS1, GMDS, ARF4, ARF5, LRPPRC, CTNNB1, ARF3, GCN1L1, BDH1, RPL9, UGCGL1, FAM3D, CCT4, CPT2, ARL1, PFKL, GOT2, AP1G1, STRBP, CLCA1, CYFIP1, COQ9, NDUFA9, ALDH7A1, HMGCS1, NNT, PRDX5, PCCB, COPZ1, BAX, ACAD9, UBXD8, HMGCS2, SLC25A3, SLC25A11, PDCD6, UCRC, DEFA6, DYNC1H1, HK1, CYFIP2, DCI and CISD1. The kit may be intended for detecting pre-cancerous or early stage colorectal cancer and include reagents detecting any protein selected from the group consisting of CPT2, ARL1, PFKL, GOT2, AP1G1, STRBP, CLCA1, CYFIP1, COQ9, NDUFA9, ALDH7A1, HMGCS1, NNT, PRDX5, PCCB, COPZ1, BAX, ACAD9, UBXD8, HMGCS2, SLC25A3, SLC25A11, PDCD6, UCRC, DEFA6, DYNC1H1, HK1, CYFIP2, DCI and CISD1. In particular

embodiments, the kit comprises a reagent suitable for detecting KIAA0152.

The kit also comprises at least one control biological sample, such as from non-diseased tissue, for comparison to a biological sample obtained from a subject under investigation. Control biological samples which correspond to samples from a cancer of a known stage or having a known prognosis, may also be included.

The kit also comprises an instructional material which directs the use of the reagents and the samples for the determining the amount and the location of the proteins or the nucleic acids in one or more cells of the sample. The instructional material also directs the correlation of the amount and the location of the protein or the nucleic acid in the cells of the sample with the diagnosis, prognosis and/or stage of a cancer or a tumor in the mammal.

As used herein, an "instructional material" includes a publication, a recording, a diagram, or any other medium of expression which directs or dictates the use of the components of a kit for performing the function of a method of the invention described herein. The instructional material of the kit of the invention may, for example, be affixed to a container which contains the composition or be shipped together with a container which contains the composition.

The following examples are presented in order to more fully illustrate certain embodiments of the invention. They should in no way, however, be construed as limiting the broad scope of the invention. One skilled in the art can readily devise many variations and modifications of the principles disclosed herein without departing from the scope of the invention.

EXAMPLES

The following methods were used in the Examples.

Protein extraction from frozen tissues

Surplus snap-frozen tumor and margin non-tumoral tissue were collected from patients undergoing surgery for colorectal cancer at either of two university teaching hospitals in Israel. Around 30 mg of frozen tissue were extracted from every tissue sample by mixing the tissue sample with in 0.5 ml of 8 M urea, 400 mM ammonium bicarbonate, and homogenized by high speed tissue homogenizer (Omni-TH) for 1 min, activated at full speed.

Identification of proteins

1) Gel-slicing method. About 50 μg of the proteins extracted from the tumor tissues or the healthy tissues were resolved by 10% of SDS-PAGE. Each gel lane was subsequently stained with Coomassie blue. The stained gel lanes were cut into 12 slices. The gel pieces were de-stained by extensive washing with acetonitrile and ammonium bicarbonate. The proteins in each gel slice were proteolyzed inside the stained gel: stained slices were reduced with 10 mM DTT, incubated at 60 °C for 30 min, alkylated with 10 mM iodoacetamide, at room temperature for 30 min and digested with trypsin overnight at 37 °C, using modified trypsin (Promega) at a 1:100 enzyme-to-substrate ratio. The tryptic peptides were analyzed by μLC-MS/MS using the OrbitrapXL mass spectrometer (Thermo-Fisher) fitted with a capillary HPLC (Eksigent). The peptides were resolved on homemade capillary columns (75 micron ID) packed with reversed phase 3.5 micron beads Reprosil C18-Aqua, using a method described by (Ishihama, Rappsilber et al. 2002). The HPLC separations of the peptides were at flow rates of about 250 nanoliters per minute during 2 hrs and with 7-40% gradients of acetonitrile in the presence of 0.1% formic acid. The capillary columns were connected on line to the Orbitrap mass spectrometer through an electrospray interface. The mass spectrometer was operated in a data-dependent mode where the masses of the eluting peptides were measured at high accuracy in the Orbitrap part of the machine and the seven most intense masses, detected at each full MS spectra whose charge states were determined to be double and triple, were selected for fragmentation by CID in the linear trap at the subsequent seven CID fragmentations.

2) Multidimensional chromatography coupled with mass spectrometry. 50 μg of total protein extracts mixed with 100 μg in 8 M urea, 100 mM ammonium bicarbonate were treated for blocking all the sulfhydryls by first reducing disulfides by addition of 10 mM DTT, incubation at 60 °C for 30 min. The free disulfides were next blocked by carboxymethylation using 10 mM iodoacetamide and incubation at room temperature for 30 min. The denatured and carboxymethylated protein mixtures were diluted three-fold with water to reduce urea concentration to about 2M followed by digestion in solution, overnight at 37 °C using modified trypsin (Promega) at a 1:100 enzyme-to-substrate ratio. The resulting peptides from the trypsinized proteins were desalted with reversed-phase with a C18 tip disposable micro-columns (Harvard), eluted with 90% acetonitrile, dried and dissolved in 0.1% formic acid. The resulting peptides were resolved by multi-dimensional chromatography with on-line first dimension of strong cation exchange (SCX) chromatography 0.3x5 mm columns (LC Packings) using ten salt steps of 20, 40, 60, 80, 100,

120, 160, 200, 300 and 500 mM ammonium acetate in 5% acetonitrile with 0.1% acetic acid. The peptides from each increased salt elution from the SCX columns were transferred on-line to a C18 trap column (0.3x5 mm, LC-Packings), which was connected on-line to a Reprosil C18 homemade capillary columns (75 micron ID), resolved by 7-40% acetonitrile gradients, during 2 hrs, in the presence of 0.1% formic acid as described before for the gel-slicing method.

Resolving peptides by two-dimensional capillary chromatography has been described (Link, Eng et al. 1999), reviewed in (Link 2002). The mass spectrometry analysis was performed on-line as described above using data-dependent LC-MS/MS analysis with full MS in the Orbitrap and subsequent seven dependent ion-trap CID spectra of the most abundant doubly and triply charged peptides, detected in the full MS.

3) Isotope labeling peptides to enable quantitative analysis. Labeling tryptic peptides with light or heavy stable isotope reagents may rely on commercial reagents, reviewed in (Ong and Mann 2005; Regnier and Julka 2006). To facilitate the accurate comparison of the relative amounts of each of the proteins in the different samples, the mixture of peptides produced by the trypsinization of the entire in-solution protein proteolysis were covalently modified with light and heavy stable-isotope reagents. Reductive dimethylation labeling was done with heavy and light formaldehyde as described by (Hsu, Huang et al. 2003).

50ug of proteolytic peptides were covalently modified with stable-isotope labeled (heavy and light formaldehyde) reagents (reductive dimethylation). The labeled peptides were resolved by multi-dimensional chromatography with on-line SCX column as described above.

Stable isotope labeling was also performed by iTRAQ (Ross, Huang et al. 2004) using a labeling kit that was purchased from Applied Biosystems and labeling was done according to the manufacturer's protocol. The labeled peptides were resolved by multi-dimensional chromatography with on-line SCX column as described above.

Bioinformatics

The mass spectrometry data of both the tryptic peptides obtained from proteolysis in the gel slices mentioned above and the tryptic peptides resolved by multidimensional chromatography were clustered and analyzed using the Pep-Miner software tool (Beer, Barnea et al. 2004). The search against the human part of the IPI database was done by using multiple search engines: Pep-Miner (Beer, Barnea et al. 2004), Mascot (Perkins, Pappin et al. 1999) and Sequest (Eng, McCormack et al. 1994). Both Mascot and Sequest were run together using the Protein Discoverer software tool (Thermo-Fisher).

Peptides were selected according to the following critera: 1) Mascot: ionScore > identityHigh and expValue < 0.05 and deltaScore = 0; 2) Sequest: ((xCorr > 2 and chg </= 2) or (xCorr > 2.5 and chg >/= 3)) and probability > 15 and deltaScore = 0 A peptide was used for the analysis if it was identified with the above criteria at least once for Sequest and once for Mascot, although not necessarily both in the same scan or run or patient.

Example 1. Clinical characteristics of patients undergoing surgical gastrointestinal re-sectioning.

The entire protein repertoires of healthy and diseased gastrointestinal tissues were analyzed from samples obtained from greater than 50 patients undergoing surgery for colorectal cancer.

The clinical characteristics of the patients studied are provided in Table 2.

Table 2.

Patient No.	**Age at Diagnosis**	**Gender**	**Appearance**	**Diagnosis**	**TNM**	**Stage**	**Grade**
201	78	M	primary	Adeno-carcinoma	T3N1M0	III	high
202	81	F	primary	Mucinous adeno-carcinoma	T3N1	III	low
203	85	M	primary	Adeno-carcinoma	T3N0M0	II	high
205	71	F	primary	Adeno-carcinoma	T3N1M0	III	high
207	81	F	primary	Adeno-carcinoma	T3N2	III	high
208	78	M	primary	Adeno-carcinoma	T3N2M1	IV	high
209	71	M	primary	Adeno-carcinoma	T3N0M1	II	high
211	57	F	primary	Adeno-carcinoma	T2N0M0	I	intermed
212	50	F	primary	Adeno-carcinoma	T3N1M0	III	high
214	54	F	primary	Adeno-carcinoma	T3N1M0	III	high
217	77	F	Local recurrance	Adeno-carcinoma	T3N0M0	II	high
218 (T+P+N)	79	M	primary	Adeno-carcinoma	T3N2M1	IV	low
219	60	M	primary	Adeno-carcinoma	T3N1M0	III	high
220	73	M	primary	Adeno-carcinoma	T3N2M1	IV	high
Patient No.	**Age at Diagnosis**	**Gender**	**Type of surgery**	**Diagnosis**	**Tumor location**	**Stage**	**Grade**
J1 (T+N)	61	F	Lt colectomy	Adeno-	left colon	IV	intermed

				carcinoma			
J2 (T+N)	67	M	Lt colectomy	Adeno-carcinoma	left colon	II	intermed
J3 (T+N)	70	M	Rt colectomy	Adeno-carcinoma	right colon	II	intermed
J4 (T+N)	83	F	Sigmoidecto my	Adeno-carcinoma	left colon	II	intermed
J5 (T+N)	41	M	Lt colectomy	Adeno-carcinoma	left colon	I	intermed
J6 (T+N)	67	M	Lt colectomy	Adeno-carcinoma	left colon	IV	intermed
J7 (T+N)	95	M	Rt colectomy	Adeno-carcinoma	right colon	III (T3N1)	intermed
J8 (T+N)	72	F	Rt colectomy	Adeno-carcinoma	right+left colon	III T3N4	high
J9 (T+N)	56	F	Rt colectomy	Adeno-carcinoma	right colon	II T4N0M 0	intermed
J10 (T+N)	64	M	Anterior Resection	Adeno-carcinoma of rectum	rectum	II T4N0M x	low
J11 (T+N)	61	F	Rt colectomy	Adeno-carcinoma	right colon	I	low
J12 (T+N)	61	M	colectomy	Adeno-carcinoma	rectum	II	intermed
J13 (T+N)	60	M	Rt colectomy	Adeno-carcinoma	right colon	I	intermed
J14 (T+N)	53	F	colectomy	Adeno-carcinoma	left colon	II	intermed
J15 (P+N)	44	M	Lt colectomy	Colonic polyps	left colon		
J16 (P+N +Pbig)	64	M	Rt colectomy	Tubulo-villus	right colon		
J17 (T+N)	81	M	Rt colectomy	Adeno-carcinoma	right colon	II	intermed
J18 (P+N)	58	F	Rt colectomy	Adeno-carcinoma in TVA	right colon	I	intermed
J19 (T+N+P)	72	M	Rt colectomy	Adeno-carcinoma	right colon	II	intermed
J20 (T+N)	60	F	Rt colectomy	Adeno-carcinoma	left colon	II	intermed
J21 (T+N)	29	F	Rt colectomy	Adeno-carcinoma	right colon	IV	low
J22 (T+N)	70	M	Rt colectomy	Adeno-carcinoma	right colon	II	high
J23 (T+N)	92	F	Rt colectomy	Adeno-carcinoma	right colon	III	intermed
J24 (P+N)	75	M	Rt colectomy	Villo-tubulos adenoma	right colon		
J25 (T+N)	47	M	Rt colectomy	Adeno-carcinoma	Rectac-toigmoid	III	intermed
J26 (T+N)	77	M	Rt colectomy	Adeno-carcinoma	right colon	III	intermed
J27	77	M	Lt colectomy	Adeno-	left colon	II	intermed

(T+N)				carcinoma			
J28 (Pbig+N+P)	77	M	ext Rt colectomy	Villo-tubulos adenoma	right colon		high
J29 (T+N)	78	M	Lt colectomy	Adeno-carcinoma	left colon	II	intermed
J30 (T+N)	67	F	Lt colectomy	Adeno-carcinoma	right colon	III	intermed
J31 (T+N)	83	M	Sigmoidecto my	Adeno-carcinoma	right colon	II	low
J32 (T+N)	65	F	Anterior Resaction	Adeno-carcinoma	right colon	II	intermed
J33 (N+P1-5)	83	M	Rt colectomy	Adeno-carcinoma	right colon	II	intermed
J34 (T+N)	82	M	Rt colectomy	Adeno-carcinoma	right colon	II	intermed-high
J35 (T+N)	81	M	Gastrectomy	Adeno-carcinoma	stomach	III	high
J36 (T+N)	62	M	Resaction	Adeno-carcinoma	recto signuoid	III	intermed
J37 (T+N)	55	M	Rt colectomy	Adeno-carcinoma		III	intermed-high
J38 (T+N+P1-P2)	80	M	Rt colectomy	Adeno-carcinoma	right colon	III	intermed-low
J39 (T+P)	62	M	Rt colectomy	Villo-tubulos adenoma	right colon	I	high
J40 (T+N)	52	M	Anterior Resection	Adeno-carcinoma	Sigmoid	III	intermed-low
J41 (T+N+P1+P2)	87	M	Rt colectomy	Adeno-carcinoma	right colon	III	intermed-low
J42 (T+N+P1+P2)	87	M	Rt colectomy	Adeno-carcinoma		I	intermed

T, tumor; N, normal; P, polyp; high grade=poorly differentiated; low grade=well differentiated; intermediate grade=moderately well differentiated.

Example 2. Identification of a group of protein markers that are found in polyps and in advanced stages of colorectal cancer.

Table 3 lists proteins that were observed to be highly expressed in polyps, as well as in early and advanced stages of various colorectal cancers. As shown in Table 3 and in Figure 2, in a large number of patients, the expression level of these proteins in polyps or cancerous tissue was at least 3 times greater than the expression level in healthy tissue from the same patient. Similarly, few if any, of the patients exhibited decreased expression levels of the same proteins in polyps or cancerous tissue as compared to the corresponding level in healthy tissue from the same patient. That is, a ratio of protein expression (polyps/tumor vs healthy

tissue from the same patient), that was less than 1:3, was rarely observed. Accordingly, a diagnostic array of reagents directed to detection of at least some of the proteins in this group may be used as a general screening test for colorectal cancers.

Table 3.

IPI Acc. No.	nPeptides	Protein	nBigT	nSmallT	nBigP	nSmallP
IPI00022649	18	SLC12A2 Isoform 1 of Solute carrier family 12 member 2	14	2	16	0
IPI00017526	3	S100P Protein S100-P	27	0	15	0
IPI00022255	16	OLFM4 Olfactomedin-4 precursor	21	2	15	0
IPI00024095	16	ANXA3 Annexin A3	19	1	15	0
IPI00555902	6	OCIAD2 Isoform 1 of OCIA domain-containing protein 2	15	0	15	0
IPI00376503	7	PYCR1 pyrroline-5-carboxylate reductase 1 isoform 2	20	0	14	0
IPI00298176	7	GPX2 Glutathione peroxidase 2	18	0	14	0
IPI00296337	55	PRKDC Isoform 1 of DNA-dependent protein kinase catalytic subunit	16	2	14	0
IPI00215919	9	ARF5 ADP-ribosylation factor 5	11	2	14	0
IPI00619903	21	UGCGL1 UDP-glucose:glycoprotein glucosyltransferase 1 precursor	9	2	14	0
IPI00018873	18	NAMPT Isoform 1 of Nicotinamide phosphoribosyltransferase	17	2	13	0
IPI00008982	16	ALDH18A1 Isoform	16	2	13	0

IPI Acc. No.	nPeptides	Protein	nBigT	nSmallT	nBigP	nSmallP
		Long of Delta-1-pyrroline-5-carboxylate synthetase				
IPI00328170	6	GCS1 Mannosyl-oligosaccharide glucosidase	15	0	13	0
IPI00216057	8	SORD Sorbitol dehydrogenase	14	2	13	0
IPI00030207	16	GMDS GDP-mannose 4,6 dehydratase	13	2	13	0
IPI00215918	10	ARF4 ADP-ribosylation factor 4	12	2	13	0
IPI00783271	45	LRPPRC Leucine-rich PPR motif-containing protein, mitochondrial precursor	11	3	13	0
IPI00017292	23	CTNNB1 Isoform 1 of Catenin beta-1	11	2	13	0
IPI00215917	10	ARF3 ADP-ribosylation factor 3	11	2	13	0
IPI00001159	18	GCN1L1 Translational activator GCN1	11	1	13	0
IPI00025341	10	BDH1 D-beta-hydroxybutyrate dehydrogenase, mitochondrial precursor	10	4	13	0
IPI00031691	9	RPL9 60S ribosomal protein L9	10	1	13	0
IPI00029046	7	KIAA0152 Uncharacterized protein KIAA0152 precursor	10	0	13	0
IPI00060143	5	FAM3D Protein FAM3D precursor	9	4	13	0
IPI00302927	18	CCT4 T-complex	8	1	13	0

IPI Acc. No.	nPeptides	Protein	nBigT	nSmallT	nBigP	nSmallP
		protein 1 subunit delta				
IPI00643623	6	LCN2 Lipocalin 2	23	0	12	0
IPI00217223	13	PAICS Multifunctional protein ADE2	17	2	12	0
IPI00012501	4	REG4 Isoform 1 of Regenerating islet-derived protein 4 precursor	16	0	12	0
IPI00793443	14	IPO5 RAN binding protein 5	15	1	12	0
IPI00025273	15	GART Isoform Long of Trifunctional purine biosynthetic protein adenosine-3	15	1	12	0
IPI00646687	19	POF1B Isoform 2 of Protein POF1B	14	2	12	0
IPI00023728	8	GGH Gamma-glutamyl hydrolase precursor	14	1	12	0
IPI00306301	13	PDHA1 Mitochondrial PDHA1	13	3	12	1
IPI00006379	12	NOP5/NOP58 Nucleolar protein 5	13	0	12	0
IPI00171692	7	ABHD11 Isoform 1 of Abhydrolase domain-containing protein 11	13	2	12	0
IPI00329719	25	MYO1D Isoform 1 of Myosin-Id	12	4	12	0
IPI00216225	16	ITGA6 Isoform Alpha-6X1X2A of Integrin alpha-6 precursor	12	2	12	0
IPI00103994	15	LARS Leucyl-tRNA synthetase, cytoplasmic	12	2	12	0
IPI00005198	8	ILF2 Interleukin enhancer-binding factor	12	1	12	0

IPI Acc. No.	nPeptides	Protein	nBigT	nSmallT	nBigP	nSmallP
		2				
IPI00783625	5	SERPINB5 Serpin B5 precursor	12	0	12	0
IPI00789324	22	JUP JUP protein	11	2	12	0
IPI00871852	21	EIF4A1 46 kDa protein	11	2	12	0
IPI00646493	28	COPA coatomer protein complex, subunit alpha isoform 1	11	2	12	0
IPI00783982	20	COPG Coatomer subunit gamma	11	2	12	0
IPI00295851	23	COPB1 Coatomer subunit beta	11	2	12	0
IPI00007928	17	PRPF8 Pre-mRNA-processing-splicing factor 8	11	1	12	0
IPI00008164	17	PREP Prolyl endopeptidase	11	1	12	0
IPI00848161	15	BAT1 Isoform 1 of Spliceosome RNA helicase BAT1	10	1	12	0
IPI00844578	26	DHX9 ATP-dependent RNA helicase A	10	2	12	0
IPI00383680	19	RPN2 Ribophorin II	10	2	12	0
IPI00011253	14	RPS3 40S ribosomal protein S3	10	1	12	0
IPI00300371	13	SF3B3 Isoform 1 of Splicing factor 3B subunit 3	10	0	12	0
IPI00893013	11	XPO1 123 kDa protein	10	1	12	0
IPI00790342	11	RPL6 60S ribosomal protein L6	10	2	12	0
IPI00011511	7	CECR5 Isoform 2 of Cat eye syndrome critical region protein 5	10	4	12	0

IPI Acc. No.	nPeptides	Protein	nBigT	nSmallT	nBigP	nSmallP
		precursor				
IPI00023542	4	TMED9 transmembrane emp24 protein transport domain containing 9	10	0	12	0
IPI00073772	8	FBP1 Fructose-1,6-bisphosphatase 1	9	4	12	0
IPI00893057	16	PDXDC1 87 kDa protein	8	3	12	0
IPI00029012	15	EIF3A Eukaryotic translation initiation factor 3 subunit A	8	1	12	0
IPI00003833	10	MTCH2 Mitochondrial carrier homolog 2	8	2	12	0
IPI00792875	3	SERPINB5 14 kDa protein	8	2	12	0
IPI00306960	9	NARS Asparaginyl-tRNA synthetase, cytoplasmic	7	1	12	0
IPI00012912	12	CPT2 Carnitine O-palmitoyltransferase 2, mitochondrial precursor	6	5	12	0
IPI00219518	4	ARL1 ADP-ribosylation factor-like protein 1	6	1	12	0
IPI00236556	23	MPO Isoform H7 of Myeloperoxidase precursor	24	2	11	0
IPI00218993	20	HSPH1 Isoform Beta of Heat shock protein 105 kDa	21	3	11	0
IPI00477179	8	DDX21 Isoform 2 of Nucleolar RNA helicase 2	17	1	11	0
IPI00215801	10	RBM39 Isoform 2 of RNA-binding protein 39	15	1	11	0

IPI Acc. No.	nPeptides	Protein	nBigT	nSmallT	nBigP	nSmallP
IPI00554788	32	KRT18 Keratin, type I cytoskeletal 18	13	2	11	0
IPI00893918	14	VARS Valyl-tRNA synthetase	13	2	11	0
IPI00877938	11	IARS isoleucyl-tRNA synthetase	13	1	11	0
IPI00641181	2	MARCKSL1 MARCKS-related protein	13	0	11	0
IPI00401990	17	ACSL5 acyl-CoA synthetase long-chain family member 5 isoform a	12	2	11	0
IPI00396435	16	DHX15 Putative pre-mRNA-splicing factor ATP-dependent RNA helicase DHX15	12	2	11	0
IPI00026089	14	SF3B1 Splicing factor 3B subunit 1	12	1	11	0
IPI00005158	14	LONP1 Lon protease homolog, mitochondrial precursor	12	2	11	0
IPI00479262	12	EIF4G1 Isoform B of Eukaryotic translation initiation factor 4 gamma 1	12	1	11	0
IPI00644431	12	DDX39 ATP-dependent RNA helicase DDX39	11	2	11	0
IPI00028931	26	DSG2 Desmoglein-2 precursor	11	4	11	0
IPI00013452	20	EPRS Bifunctional aminoacyl-tRNA synthetase	11	3	11	0
IPI00418313	17	ILF3 Isoform 4 of	11	3	11	0

IPI Acc. No.	nPeptides	Protein	nBigT	nSmallT	nBigP	nSmallP
		Interleukin enhancer-binding factor 3				
IPI00017376	7	SEC23B Protein transport protein Sec23B	11	2	11	0
IPI00297084	8	DDOST dolichyl-diphosphooligosaccharide-protein glycosyltransferase precursor	11	1	11	0
IPI00411937	5	NOL5A Nucleolar protein 5A	11	1	11	1
IPI00221091	5	RPS15A 40S ribosomal protein S15a	11	0	11	0
IPI00015872	3	TSPAN8 Tetraspanin-8	11	7	11	0
IPI00784044	11	MCCC2 Isoform 1 of Methylcrotonoyl-CoA carboxylase beta chain, mitochondrial precursor	10	3	11	0
IPI00013485	10	RPS2 40S ribosomal protein S2	10	2	11	0
IPI00006684	7	API5 58 kDa protein	10	2	11	0
IPI00028004	5	PSMB3 Proteasome subunit beta type-3	10	2	11	0
IPI00019385	5	SSR4 Translocon-associated protein subunit delta precursor	10	2	11	0
IPI00016608	5	TMED2 Transmembrane emp24 domain-containing protein 2 precursor	10	1	11	0
IPI00329791	5	DDX46 cDNA FLJ78679, highly similar to Homo sapiens DEAD (Asp-Glu-Ala-	10	0	11	0

IPI Acc. No.	nPeptides	Protein	nBigT	nSmallT	nBigP	nSmallP
		Asp) box polypeptide 46 (DDX46), mRNA				
IPI00215948	35	CTNNA1 Isoform 1 of Catenin alpha-1	9	3	11	0
IPI00217952	29	GFPT1 Isoform 1 of Glucosamine--fructose-6-phosphate aminotransferase [isomerizing] 1	9	2	11	0
IPI00220834	18	XRCC5 ATP-dependent DNA helicase 2 subunit 2	9	1	11	0
IPI00514622	3	RANBP6 Ran-binding protein 6	9	0	11	0
IPI00027252	16	PHB2 Prohibitin-2	9	2	11	0
IPI00020672	13	DPP3;BBS1 Isoform 1 of Dipeptidyl-peptidase 3	9	3	11	0
IPI00017895	13	GPD2 Isoform 1 of Glycerol-3-phosphate dehydrogenase, mitochondrial precursor	9	2	11	0
IPI00872756	10	ASL 58 kDa protein	9	4	11	0
IPI00329598	8	HSD17B11 Estradiol 17-beta-dehydrogenase 11 precursor	9	3	11	1
IPI00013296	6	RPS18;LOC100130553 40S ribosomal protein S18	9	1	11	0
IPI00477831	14	ERAP1 Isoform 1 of Endoplasmic reticulum aminopeptidase 1	8	2	11	0
IPI00303207	7	ABCE1 ATP-binding cassette sub-family E	8	0	11	0

IPI Acc. No.	nPeptides	Protein	nBigT	nSmallT	nBigP	nSmallP
		member 1				
IPI00847192	6	RPS9 protein (Fragment)	8	1	11	0
IPI00441344	5	GLB1 Isoform 1 of Beta-galactosidase precursor	8	1	11	0
IPI00418497	3	TIMM50 Isoform 2 of Import inner membrane translocase subunit TIM50, mitochondrial precursor	8	0	11	0
IPI00646182	40	ATP1A1 ATPase, Na+/K+ transporting, alpha 1 polypeptide	7	3	11	0
IPI00009634	20	SQRDL Sulfide:quinone oxidoreductase, mitochondrial precursor	7	3	11	0
IPI00304171	12	H2AFY Isoform 2 of Core histone macro-H2A.1	7	1	11	0
IPI00022887	6	ERGIC1 Isoform 1 of Endoplasmic reticulum-Golgi intermediate compartment protein 1	7	5	11	0
IPI00171626	2	LPCAT1 1-acylglycerophosphocholine O-acyltransferase 1	7	0	11	0
IPI00332371	20	PFKL Isoform 1 of 6-phosphofructokinase, liver type	6	2	11	0
IPI00018206	15	GOT2 Aspartate aminotransferase, mitochondrial precursor	6	3	11	0
IPI00643591	5	AP1G1 AP-1 complex	6	1	11	0

IPI Acc. No.	nPeptides	Protein	nBigT	nSmallT	nBigP	nSmallP
		subunit gamma-1				
IPI00413860	2	STRBP Isoform 2 of Spermatid perinuclear RNA-binding protein	6	0	11	0
IPI00014625	42	CLCA1 Calcium-activated chloride channel regulator 1 precursor	4	13	11	1
IPI00644231	14	CYFIP1 Isoform 1 of Cytoplasmic FMR1-interacting protein 1	4	2	11	0
IPI00470631	4	COQ9 Isoform 1 of Ubiquinone biosynthesis protein COQ9, mitochondrial precursor	4	3	11	0
IPI00003968	9	NDUFA9 NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 9, mitochondrial precursor	3	4	11	0
IPI00221234	14	ALDH7A1 Similar to Antiquitin	2	5	11	0
IPI00007047	6	S100A8 Protein S100-A8	29	0	10	0
IPI00027462	6	S100A9 Protein S100-A9	27	1	10	0
IPI00021700	8	PCNA Proliferating cell nuclear antigen	18	0	10	0
IPI00297579	7	CBX3;LOC653972 Chromobox protein homolog 3	18	1	10	0
IPI00642046	4	RSL1D1 RSL1D1 protein	18	1	10	0
IPI00002520	14	SHMT2 Serine	17	2	10	0

IPI Acc. No.	nPeptides	Protein	nBigT	nSmallT	nBigP	nSmallP
		hydroxymethyltransferase, mitochondrial precursor				
IPI00219871	2	LSM8 U6 snRNA-associated Sm-like protein LSm8	16	2	10	1
IPI00030275	17	TRAP1 Heat shock protein 75 kDa, mitochondrial precursor	15	2	10	0
IPI00893035	8	CAD Putative uncharacterized protein CAD	15	2	10	0
IPI00553131	9	GALE UDP-glucose 4-epimerase	14	2	10	0
IPI00217477	3	HMGB3 High mobility group protein B3	14	3	10	0
IPI00295992	9	ATAD3A Isoform 2 of ATPase family AAA domain-containing protein 3A	13	2	10	0
IPI00646721	9	USP7 Ubiquitin carboxyl-terminal hydrolase	13	9	10	0
IPI00465044	8	RCC2 Protein RCC2	13	1	10	0
IPI00420014	17	ASCC3L1 Isoform 1 of U5 small nuclear ribonucleoprotein 200 kDa helicase	12	0	10	0
IPI00027444	16	SERPINB1 Leukocyte elastase inhibitor	12	2	10	0
IPI00216308	14	VDAC1 Voltage-dependent anion-selective channel protein 1	12	2	10	0

IPI Acc. No.	nPeptides	Protein	nBigT	nSmallT	nBigP	nSmallP
IPI00334175	14	PTBP1 Isoform 2 of Polypyrimidine tract-binding protein 1	12	2	10	0
IPI00218466	8	SEC61A1 Isoform 1 of Protein transport protein Sec61 subunit alpha isoform 1	12	1	10	0
IPI00019912	24	HSD17B4 Peroxisomal multifunctional enzyme type 2	11	1	10	0
IPI00002557	4	COPG2 Coatomer subunit gamma-2	11	3	10	0
IPI00449049	16	PARP1 Poly [ADP-ribose] polymerase 1	11	1	10	0
IPI00014238	10	KARS Lysyl-tRNA synthetase	11	2	10	0
IPI00008530	10	RPLP0 60S acidic ribosomal protein P0	11	2	10	0
IPI00009950	8	LMAN2 Vesicular integral-membrane protein VIP36 precursor	11	2	10	0
IPI00026202	6	RPL18A 60S ribosomal protein L18a	11	2	10	0
IPI00018597	3	SYK Isoform Long of Tyrosine-protein kinase SYK	11	0	10	0
IPI00216951	21	DARS Aspartyl-tRNA synthetase, cytoplasmic	10	2	10	0
IPI00747497	14	EEF1G 50 kDa protein	10	2	10	0
IPI00032038	14	CPT1A Isoform 1 of Carnitine O-palmitoyltransferase I, liver isoform	10	7	10	0
IPI00297492	10	STT3A Dolichyl-	10	1	10	0

IPI Acc. No.	nPeptides	Protein	nBigT	nSmallT	nBigP	nSmallP
		diphosphooligosaccharide--protein glycosyltransferase subunit STT3A				
IPI00549672	8	PSMD13 HSPC027	10	0	10	0
IPI00221089	7	RPS13 40S ribosomal protein S13	10	1	10	0
IPI00152377	7	STT3B Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit STT3B	10	0	10	0
IPI00028055	6	TMED10 Transmembrane emp24 domain-containing protein 10 precursor	10	2	10	0
IPI00032139	3	SERPINB9 Serpin B9	10	1	10	0
IPI00290089	28	CDH17 Cadherin-17 precursor	9	5	10	0
IPI00797038	21	PCK2 mitochondrial phosphoenolpyruvate carboxykinase 2 isoform 1 precursor	9	2	10	0
IPI00004860	18	RARS Isoform Complexed of Arginyl-tRNA synthetase, cytoplasmic	9	2	10	0
IPI00026665	13	QARS Glutaminyl-tRNA synthetase	9	1	10	0
IPI00017283	10	IARS2 Isoleucyl-tRNA synthetase, mitochondrial precursor	9	2	10	0
IPI00030847	7	TM9SF3 Transmembrane 9	9	1	10	0

IPI Acc. No.	nPeptides	Protein	nBigT	nSmallT	nBigP	nSmallP
		superfamily member 3 precursor				
IPI00014361	5	TSTA3 GDP-L-fucose synthetase	9	0	10	0
IPI00008298	1	DEFA5 Defensin-5 precursor	9	0	10	0
IPI00033022	15	DNM2 Isoform 1 of Dynamin-2	8	1	10	0
IPI00220847	19	ITGB4 Isoform Beta-4D of Integrin beta-4 precursor	8	5	10	0
IPI00847318	15	PKP2 plakophilin 2 isoform 2a	8	5	10	0
IPI00215911	11	APEX1 DNA-(apurinic or apyrimidinic site) lyase	8	1	10	0
IPI00874185	10	HIBCH 46 kDa protein	8	5	10	0
IPI00182533	6	RPL28 60S ribosomal protein L28	8	6	10	0
IPI00296909	6	PARP4 Poly [ADP-ribose] polymerase 4	8	2	10	0
IPI00023876	5	CASP6 Isoform Alpha of Caspase-6 precursor	8	3	10	0
IPI00744194	5	Similar to Sodium/potassium-transporting ATPase alpha-1 chain precursor	7	4	10	0
IPI00302925	28	CCT8 59 kDa protein	7	2	10	0
IPI00100160	19	CAND1 Isoform 1 of Cullin-associated NEDD8-dissociated protein 1	7	1	10	0
IPI00337494	13	SLC25A24 Isoform 1 of Calcium-binding	7	3	10	0

IPI Acc. No.	nPeptides	Protein	nBigT	nSmallT	nBigP	nSmallP
		mitochondrial carrier protein SCaMC-1				
IPI00747849	6	ATP1B1 Isoform 1 of Sodium/potassium-transporting ATPase subunit beta-1	7	3	10	0
IPI00008433	5	RPS5 40S ribosomal protein S5	7	2	10	0
IPI00008475	3	HMGCS1 Hydroxymethylglutaryl-CoA synthase, cytoplasmic	6	4	10	0
IPI00337541	18	NNT NAD(P) transhydrogenase, mitochondrial precursor	6	2	10	0
IPI00876999	12	PRDX5 Uncharacterized protein PRDX5 (Fragment)	6	2	10	0
IPI00007247	12	PCCB Propionyl-CoA carboxylase beta chain, mitochondrial precursor	6	5	10	0
IPI00032851	5	COPZ1 Coatomer subunit zeta-1	6	4	10	0
IPI00845474	5	BAX BCL2-associated X protein isoform sigma	6	2	10	0
IPI00152981	3	ACAD9 Acyl-CoA dehydrogenase family member 9, mitochondrial precursor	6	0	10	0
IPI00172656	1	UBXD8 UBX domain-containing protein 8	6	1	10	0
IPI00008934	20	HMGCS2 Hydroxymethylglutaryl-CoA synthase,	5	12	10	0

IPI Acc. No.	nPeptides	Protein	nBigT	nSmallT	nBigP	nSmallP
		mitochondrial precursor				
IPI00790115	11	SLC25A3 cDNA FLJ90278 fis, clone NT2RP1000325, highly similar to Phosphate carrier protein, mitochondrialprecursor	5	1	10	0
IPI00219729	9	SLC25A11 Mitochondrial 2-oxoglutarate/malate carrier protein	5	4	10	0
IPI00025277	5	PDCD6 Programmed cell death protein 6	5	0	10	0
IPI00554701	2	UCRC Cytochrome b-c1 complex subunit 9	4	0	10	0
IPI00008301	2	DEFA6 Defensin-6 precursor	4	0	10	0
IPI00456969	88	DYNC1H1 Cytoplasmic dynein 1 heavy chain 1	3	1	10	0
IPI00220663	17	HK1 Isoform 2 of Hexokinase-1	3	1	10	0
IPI00719600	7	CYFIP2 Isoform 2 of Cytoplasmic FMR1-interacting protein 2	3	2	10	0
IPI00398758	5	DCI Isoform 2 of 3,2-trans-enoyl-CoA isomerase, mitochondrial precursor	3	5	10	0
IPI00020510	3	CISD1 CDGSH iron sulfur domain-containing protein 1	1	1	10	0
IPI00294443	2	CLIC5 Isoform 1 of Chloride intracellular channel protein 5	21	0	9	0

IPI Acc. No.	nPeptides	Protein	nBigT	nSmallT	nBigP	nSmallP
IPI00031564	7	C7orf24 Uncharacterized protein C7orf24	17	1	9	0
IPI00871140	10	NP Purine nucleoside phosphorylase	16	1	9	0
IPI00291510	9	IMPDH2 Inosine-5'-monophosphate dehydrogenase 2	15	1	9	0
IPI00792186	5	ABCF1 ATP-binding cassette, sub-family F (GCN20), member 1	15	0	9	0
IPI00029744	4	SSBP1 Single-stranded DNA-binding protein, mitochondrial precursor	15	3	9	0
IPI00166680	2	MINK1 Isoform 3 of Misshapen-like kinase 1	15	1	9	0
IPI00152441	3	HM13 Isoform 1 of Minor histocompatibility antigen H13	14	0	9	0
IPI00303318	8	FAM49B Protein FAM49B	13	2	9	0
IPI00643166	4	PGM3 Isoform 2 of Phosphoacetylglucosamine mutase	13	0	9	0
IPI00641950	14	GNB2L1 Lung cancer oncogene 7	12	2	9	0
IPI00293464	11	DDB1 DNA damage-binding protein 1	12	1	9	0
IPI00783097	10	GARS Glycyl-tRNA synthetase	12	2	9	0
IPI00219617	4	PRPS2 Isoform 1 of Ribose-phosphate pyrophosphokinase 2	12	1	9	0

IPI Acc. No.	nPeptides	Protein	nBigT	nSmallT	nBigP	nSmallP
IPI00018415	7	TM9SF2 Transmembrane 9 superfamily member 2 precursor	12	1	9	0
IPI00550032	4	LOC653232;RPL15 Ribosomal protein L15 pseudogene 3	12	0	9	0
IPI00013933	65	DSP Isoform DPI of Desmoplakin	11	2	9	0
IPI00009032	15	SSB Lupus La protein	11	1	9	0
IPI00792100	10	C14orf166 CLE	11	1	9	0
IPI00010491	3	RAB27B Ras-related protein Rab-27B	11	2	9	0
IPI00302850	3	SNRPD1 Small nuclear ribonucleoprotein Sm D1	11	1	9	0
IPI00644712	20	XRCC6 ATP-dependent DNA helicase 2 subunit 1	10	1	9	0
IPI00025874	20	RPN1 Dolichyl-diphosphooligosaccharide--protein glycosyltransferase 67 kDa subunit precursor	10	2	9	0
IPI00744889	9	CDH1 E-cadherin	10	2	9	0
IPI00002372	8	ABCD3 Isoform 1 of ATP-binding cassette sub-family D member 3	10	4	9	0
IPI00289601	4	HDAC2 histone deacetylase 2	10	2	9	0
IPI00328867	5	SRC Isoform 2 of Proto-oncogene tyrosine-protein kinase Src	10	0	9	0
IPI00872474	3	LYN LYN protein	10	0	9	0

IPI Acc. No.	nPeptides	Protein	nBigT	nSmallT	nBigP	nSmallP
		(Fragment)				
IPI00031804	8	VDAC3 Isoform 1 of Voltage-dependent anion-selective channel protein 3	10	2	9	0
IPI00654777	7	EIF3F Eukaryotic translation initiation factor 3 subunit 5	10	0	9	0
IPI00440703	6	GSTK1 GSTK1 protein	10	3	9	0
IPI00885106	3	TM9SF4 Isoform 2 of Transmembrane 9 superfamily member 4 precursor	10	0	9	0
IPI00029267	2	SNRPB2 U2 small nuclear ribonucleoprotein B"	10	2	9	0
IPI00414717	12	GLG1 golgi apparatus protein 1	9	1	9	0
IPI00216293	11	TST Thiosulfate sulfurtransferase	9	6	9	0
IPI00029629	11	TRIM25 Tripartite motif-containing protein 25	9	2	9	0
IPI00219147	4	CSDA Isoform 2 of DNA-binding protein A	9	1	9	0
IPI00219953	9	CMPK1 cytidine monophosphate (UMP-CMP) kinase 1, cytosolic	9	3	9	0
IPI00827508	7	RPL10A 25 kDa protein	9	2	9	0
IPI00014053	6	TOMM40 Isoform 1 of Mitochondrial import receptor subunit TOM40 homolog	9	0	9	0

IPI Acc. No.	nPeptides	Protein	nBigT	nSmallT	nBigP	nSmallP
IPI00299048	20	IQGAP2 Isoform 1 of Ras GTPase-activating-like protein IQGAP2	8	2	9	0
IPI00242956	70	FCGBP IgGFc-binding protein precursor	8	9	9	2
IPI00291467	18	SLC25A6 ADP/ATP translocase 3	8	2	9	0
IPI00290566	22	TCP1 T-complex protein 1 subunit alpha	8	2	9	0
IPI00470502	14	PPA2 Isoform 2 of Inorganic pyrophosphatase 2, mitochondrial precursor	8	1	9	0
IPI00003482	12	DECR1 2,4-dienoyl-CoA reductase, mitochondrial precursor	8	4	9	0
IPI00456750	10	FAM129B Niban-like protein 1	8	3	9	0
IPI00303158	10	CMAS Isoform 1 of N-acylneuraminate cytidylyltransferase	8	2	9	0
IPI00001466	8	EML4 Echinoderm microtubule-associated protein-like 4	8	2	9	0
IPI00105598	7	PSMD11 Proteasome 26S non-ATPase subunit 11 variant (Fragment)	8	1	9	0
IPI00011916	4	JTV1 Multisynthetase complex auxiliary component p38	8	1	9	0
IPI00100460	3	DARS2 Aspartyl-tRNA synthetase, mitochondrial precursor	8	1	9	0
IPI00030706	3	AHSA1 Activator of 90	8	1	9	0

IPI Acc. No.	nPeptides	Protein	nBigT	nSmallT	nBigP	nSmallP
		kDa heat shock protein ATPase homolog 1				
IPI00874145	3	DKC1 Uncharacterized protein DKC1 (Fragment)	8	0	9	0
IPI00177817	24	ATP2A2 Isoform SERCA2A of Sarcoplasmic/endoplasmic reticulum calcium ATPase 2	7	2	9	0
IPI00409717	13	EIF4A2 Isoform 2 of Eukaryotic initiation factor 4A-II	7	1	9	0
IPI00719752	15	EIF3B Isoform 2 of Eukaryotic translation initiation factor 3 subunit B	7	2	9	0
IPI00001091	12	AFG3L2 AFG3-like protein 2	7	7	9	0
IPI00026530	10	LMAN1 Protein ERGIC-53 precursor	7	2	9	0
IPI00007676	9	HSD17B12 Estradiol 17-beta-dehydrogenase 12	7	2	9	0
IPI00291930	8	CLINT1 Isoform 1 of Clathrin interactor 1	7	2	9	0
IPI00293853	7	GPA33 Cell surface A33 antigen precursor	7	6	9	0
IPI00797738	5	COX6B1 12 kDa protein	7	2	9	0
IPI00885058	3	MBOAT7 Isoform 2 of Membrane-bound O-acyltransferase domain-containing protein 7	7	5	9	0

IPI Acc. No.	nPeptides	Protein	nBigT	nSmallT	nBigP	nSmallP
IPI00455383	68	CLTC Isoform 2 of Clathrin heavy chain 1	6	2	9	0
IPI00102864	12	HK2 Hexokinase-2	6	2	9	0
IPI00018931	16	VPS35 Vacuolar protein sorting-associated protein 35	6	0	9	1
IPI00010157	8	MAT2A S-adenosylmethionine synthetase isoform type-2	6	1	9	1
IPI00328715	6	MTDH Protein LYRIC	6	1	9	0
IPI00171573	5	CCDC109A Isoform 1 of Coiled-coil domain-containing protein 109A	6	0	9	0
IPI00411426	5	VPS26A Vacuolar protein sorting-associated protein 26A	6	1	9	0
IPI00017767	2	MGST2 Microsomal glutathione S-transferase 2	6	2	9	0
IPI00556311	2	DUOX2 Dual oxidase 2 variant (Fragment)	6	0	9	0
IPI00011201	10	ME2 NAD-dependent malic enzyme, mitochondrial precursor	5	3	9	0
IPI00329672	5	MYO1E Myosin-Ie	5	1	9	0
IPI00020928	4	TFAM Transcription factor A, mitochondrial precursor	5	4	9	0
IPI00009104	12	RUVBL2 RuvB-like 2	4	0	9	0
IPI00744115	12	PCCA propionyl-Coenzyme A carboxylase, alpha polypeptide precursor	4	3	9	0

IPI Acc. No.	nPeptides	Protein	nBigT	nSmallT	nBigP	nSmallP
IPI00219029	10	GOT1 Aspartate aminotransferase, cytoplasmic	4	3	9	0
IPI00029264	7	CYC1 Cytochrome c1, heme protein, mitochondrial precursor	4	2	9	0
IPI00027448	5	ATP5L ATP synthase subunit g, mitochondrial	4	2	9	0
IPI00215920	2	ARF6 ADP-ribosylation factor 6	4	0	9	0
IPI00220244	1	TRIM23 Isoform Beta of GTP-binding protein ARD-1	4	0	9	0
IPI00006674	7	ABCC3 Isoform 3 of Canalicular multispecific organic anion transporter 2	3	4	9	0
IPI00843876	7	TNPO1 Transportin-1	3	0	9	0
IPI00023001	3	C3orf28 E2-induced gene 5 protein	3	5	9	0
IPI00019038	4	LYZ Lysozyme C precursor	17	2	8	0
IPI00022462	14	TFRC Transferrin receptor protein 1	16	0	8	0
IPI00554722	7	LOC442497;SLC3A2 solute carrier family 3 (activators of dibasic and neutral amino acid transport), member 2 isoform e	16	1	8	0
IPI00010341	6	PRG2 Bone marrow proteoglycan precursor	16	7	8	0
IPI00010320	2	CBX1 Chromobox protein homolog 1	16	1	8	0

IPI Acc. No.	nPeptides	Protein	nBigT	nSmallT	nBigP	nSmallP
IPI00006690	27	EPX Eosinophil peroxidase precursor	15	8	8	0
IPI00099110	4	DMBT1 Isoform 1 of Deleted in malignant brain tumors 1 protein precursor	15	0	8	1
IPI00008240	8	MARS Methionyl-tRNA synthetase, cytoplasmic	14	1	8	0
IPI00414676	41	HSP90AB1 Heat shock protein HSP 90-beta	13	2	8	0
IPI00140420	27	SND1 Staphylococcal nuclease domain-containing protein 1	13	2	8	0
IPI00807557	14	PA2G4 PA2G4 protein (Fragment)	13	2	8	0
IPI00215879	6	SFRS6 Isoform SRP55-3 of Splicing factor, arginine/serine-rich 6	13	1	8	0
IPI00301936	8	ELAVL1 ELAV-like protein 1	13	1	8	0
IPI00025039	7	FBL rRNA 2'-O-methyltransferase fibrillarin	13	2	8	0
IPI00025427	4	RNASE3 Eosinophil cationic protein precursor	13	6	8	1
IPI00873179	4	CLC Uncharacterized protein CLC (Fragment)	13	6	8	0
IPI00655650	3	LOC728937;RPS26 40S ribosomal protein S26	13	1	8	0
IPI00888987	2	LOC345630 similar to hCG1641252	13	1	8	0
IPI00026952	13	PKP3 Plakophilin-3	12	1	8	0
IPI00009328	12	EIF4A3 Eukaryotic	12	2	8	0

IPI Acc. No.	nPeptides	Protein	nBigT	nSmallT	nBigP	nSmallP
		initiation factor 4A-III				
IPI00759824	6	ANP32B Isoform 2 of Acidic leucine-rich nuclear phosphoprotein 32 family member B	12	2	8	1
IPI00029048	9	TTLL12 Tubulin--tyrosine ligase-like protein 12	12	3	8	0
IPI00029764	6	SF3A3 Splicing factor 3A subunit 3	12	1	8	0
IPI00303954	4	CYB5B cytochrome b5 outer mitochondrial membrane precursor	12	2	8	0
IPI00007402	4	IPO7 Importin-7	12	0	8	0
IPI00012382	2	SNRPA U1 small nuclear ribonucleoprotein A	12	2	8	0
IPI00446377	2	ENG cDNA FLJ41744 fis, clone HSYRA2005496, highly similar to ENDOGLIN	12	0	8	0
IPI00186290	45	EEF2 Elongation factor 2	11	2	8	0
IPI00304596	18	NONO Non-POU domain-containing octamer-binding protein	11	2	8	0
IPI00220219	17	COPB2 Coatomer subunit beta'	11	2	8	0
IPI00003918	16	RPL4 60S ribosomal protein L4	11	2	8	0
IPI00020632	14	ASS1 Argininosuccinate synthase	11	3	8	0
IPI00396661	5	CYP2S1 Isoform 1 of Cytochrome P450 2S1	11	0	8	0

IPI Acc. No.	nPeptides	Protein	nBigT	nSmallT	nBigP	nSmallP
IPI00005589	5	hCG_18290 Uncharacterized protein ENSP00000275524	11	1	8	0
IPI00178440	4	EEF1B2 Elongation factor 1-beta	11	1	8	0
IPI00894416	4	BZW2 47 kDa protein	11	0	8	0
IPI00793862	3	SHMT1 Serine hydroxymethyltransferase	11	0	8	0
IPI00029750	3	RPS24 Isoform 1 of 40S ribosomal protein S24	11	1	8	0
IPI00873680	3	EIF4E Uncharacterized protein EIF4E (Fragment)	11	2	8	0
IPI00784154	45	HSPD1 60 kDa heat shock protein, mitochondrial precursor	10	3	8	0
IPI00219005	15	FKBP4 FK506-binding protein 4	10	2	8	0
IPI00003519	13	EFTUD2 116 kDa U5 small nuclear ribonucleoprotein component	10	1	8	0
IPI00013774	4	HDAC1 Histone deacetylase 1	10	3	8	0
IPI00428967	6	TICAM2;TMED7 Toll-like receptor adapter molecule 2	10	1	8	0
IPI00219445	5	PSME3 Isoform 2 of Proteasome activator complex subunit 3	10	1	8	0
IPI00029731	4	RPL35A 60S ribosomal protein L35a	10	2	8	0
IPI00025329	3	RPL19 60S ribosomal	10	4	8	0

IPI Acc. No.	nPeptides	Protein	nBigT	nSmallT	nBigP	nSmallP
		protein L19				
IPI00455757	3	Similar to 60S ribosomal protein L35	10	2	8	0
IPI00009407	2	DAD1 Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit DAD1	10	0	8	0
IPI00010951	33	EPPK1 Epiplakin	9	1	8	0
IPI00658109	16	LOC100133623;CKMT1B;CKMT1A Creatine kinase, ubiquitous mitochondrial precursor	9	9	8	0
IPI00293721	9	AKR7A3 Aflatoxin B1 aldehyde reductase member 3	9	4	8	0
IPI00642211	15	RNPEP Aminopeptidase B	9	2	8	0
IPI00299608	14	PSMD1 Isoform 1 of 26S proteasome non-ATPase regulatory subunit 1	9	2	8	0
IPI00017726	13	HSD17B10 Isoform 1 of 3-hydroxyacyl-CoA dehydrogenase type-2	9	2	8	0
IPI00479722	13	PSME1 Proteasome activator complex subunit 1	9	1	8	0
IPI00009750	12	LGALS4 Galectin-4	9	6	8	0
IPI00334190	10	STOML2 Stomatin-like protein 2	9	2	8	0
IPI00853220	10	SEC31A Isoform 6 of Protein transport protein Sec31A	9	2	8	0

IPI Acc. No.	nPeptides	Protein	nBigT	nSmallT	nBigP	nSmallP
IPI00030131	9	TMPO Isoform Beta of Lamina-associated polypeptide 2, isoforms beta/gamma	9	2	8	0
IPI00640155	8	PSMB8 proteasome beta 8 subunit isoform E2 proprotein	9	3	8	0
IPI00221222	7	SUB1 Activated RNA polymerase II transcriptional coactivator p15	9	2	8	0
IPI00465361	6	RPL13 60S ribosomal protein L13	9	2	8	0
IPI00005537	6	MRPL12 39S ribosomal protein L12, mitochondrial precursor	9	4	8	1
IPI00293564	5	HMGCL Hydroxymethylglutaryl-CoA lyase, mitochondrial precursor	9	3	8	0
IPI00785096	4	BZW1 Isoform 1 of Basic leucine zipper and W2 domain-containing protein 1	9	1	8	0
IPI00100656	4	GPSN2 Isoform 1 of Synaptic glycoprotein SC2	9	0	8	1
IPI00009922	4	C14orf156 SRA stem-loop-interacting RNA-binding protein, mitochondrial precursor	9	1	8	0
IPI00007188	17	SLC25A5 ADP/ATP translocase 2	8	2	8	0
IPI00017375	9	SEC23A Protein	8	1	8	0

IPI Acc. No.	nPeptides	Protein	nBigT	nSmallT	nBigP	nSmallP
		transport protein Sec23A				
IPI00873762	3	TAF15 65 kDa protein	8	0	8	0
IPI00885213	7	TXNRD1 Isoform 6 of Thioredoxin reductase 1, cytoplasmic	8	1	8	3
IPI00329352	6	NOMO1;NOMO3 Nodal modulator 1 precursor	8	3	8	0
IPI00029628	3	RCN2 Reticulocalbin-2 precursor	8	0	8	0
IPI00797230	2	RPL8 32 kDa protein	8	3	8	0
IPI00220835	2	SEC61B Protein transport protein Sec61 subunit beta	8	2	8	0
IPI00784366	13	AP2B1 Isoform 2 of AP-2 complex subunit beta-1	7	0	8	0
IPI00220994	5	H2AFY2 Core histone macro-H2A.2	7	1	8	0
IPI00012268	12	PSMD2 26S proteasome non-ATPase regulatory subunit 2	7	2	8	0
IPI00013068	12	EIF3E Eukaryotic translation initiation factor 3 subunit E	7	0	8	0
IPI00472054	11	FAM120A Isoform A of Constitutive coactivator of PPAR-gamma-like protein 1	7	4	8	0
IPI00383046	6	CMBL Carboxymethylenebutenolidase homolog	7	8	8	1
IPI00031820	5	FARSA Phenylalanyl-	7	2	8	0

IPI Acc. No.	nPeptides	Protein	nBigT	nSmallT	nBigP	nSmallP
		tRNA synthetase alpha chain				
IPI00004845	4	NIPSNAP3A Protein NipSnap homolog 3A	7	3	8	0
IPI00018465	24	CCT7 T-complex protein 1 subunit eta	6	2	8	0
IPI00027442	19	AARS Alanyl-tRNA synthetase, cytoplasmic	6	3	8	0
IPI00017592	9	LETM1 Leucine zipper-EF-hand-containing transmembrane protein 1, mitochondrial precursor	6	2	8	0
IPI00300050	8	HSD11B2 Corticosteroid 11-beta-dehydrogenase isozyme 2	6	13	8	2
IPI00306516	8	TIMM44 Import inner membrane translocase subunit TIM44, mitochondrial precursor	6	1	8	0
IPI00293267	5	LGALS9 Isoform Short of Galectin-9	6	3	8	0
IPI00021258	7	ARFIP1 Isoform B of Arfaptin-1	6	2	8	1
IPI00303568	5	PTGES2 Prostaglandin E synthase 2	6	1	8	0
IPI00003870	5	CLPP Putative ATP-dependent Clp protease proteolytic subunit, mitochondrial precursor	6	2	8	0
IPI00843910	4	FUCA1 Tissue alpha-L-fucosidase precursor	6	16	8	4
IPI00452747	4	LOC653566 Similar to	6	1	8	0

IPI Acc. No.	nPeptides	Protein	nBigT	nSmallT	nBigP	nSmallP
		Signal peptidase complex subunit 2				
IPI00304612	2	RPL13A 60S ribosomal protein L13a	6	1	8	0
IPI00871366	9	RAB1B Small GTP-binding protein	5	1	8	0
IPI00006451	9	NSF Vesicle-fusing ATPase	5	0	8	1
IPI00025239	6	NDUFS2 NADH dehydrogenase [ubiquinone] iron-sulfur protein 2, mitochondrial precursor	5	2	8	0
IPI00550644	5	LL22NC03-5H6.5 UPF0530 protein	5	3	8	0
IPI00016077	3	GBAS Protein NipSnap homolog 2	5	6	8	0
IPI00031534	2	ST6GALNAC1 Alpha-N-acetylgalactosaminide alpha-2,6-sialyltransferase 1	5	5	8	0
IPI00025796	9	NDUFS3 NADH dehydrogenase [ubiquinone] iron-sulfur protein 3, mitochondrial precursor	4	1	8	0
IPI00794899	3	37 kDa protein	4	4	8	0
IPI00029054	4	NT5C2 Cytosolic purine 5'-nucleotidase	4	0	8	0
IPI00002255	4	LRBA Lipopolysaccharide-responsive and beige-like anchor protein	4	3	8	0
IPI00253050	2	L1TD1 LINE-1 type	4	0	8	0

IPI Acc. No.	nPeptides	Protein	nBigT	nSmallT	nBigP	nSmallP
		transposase domain-containing protein 1				
IPI00465179	8	PFKM cDNA FLJ44241 fis, clone THYMU3008436, highly similar to 6-phosphofructokinase, muscle type	3	1	8	0
IPI00294187	9	PADI2 Protein-arginine deiminase type-2	3	13	8	2
IPI00102581	7	SULT1B1 Sulfotransferase family cytosolic 1B member 1	3	9	8	0
IPI00023647	4	UBA6 Isoform 1 of Ubiquitin-like modifier-activating enzyme 6	3	2	8	0
IPI00030320	5	DDX6 Probable ATP-dependent RNA helicase DDX6	2	1	8	0
IPI00220740	8	NPM1 Isoform 2 of Nucleophosmin	17	2	7	0
IPI00386755	7	ERO1L ERO1-like protein alpha precursor	17	1	7	1
IPI00295542	13	NUCB1 Nucleobindin-1 precursor	15	3	7	0
IPI00293655	13	DDX1 ATP-dependent RNA helicase DDX1	14	1	7	0
IPI00007084	11	SLC25A13 Mitochondrial aspartate-glutamate carrier protein	14	2	7	0
IPI00007427	8	AGR2 AGR2	14	5	7	0
IPI00152409	7	AGR3 Anterior gradient protein 3 homolog precursor	14	4	7	0

IPI Acc. No.	nPeptides	Protein	nBigT	nSmallT	nBigP	nSmallP
IPI00004573	27	PIGR Polymeric immunoglobulin receptor precursor	13	5	7	1
IPI00021290	15	ACLY ATP-citrate synthase	13	1	7	0
IPI00456919	14	HUWE1 Isoform 1 of E3 ubiquitin-protein ligase HUWE1	13	0	7	0
IPI00015018	13	PPA1 Inorganic pyrophosphatase	13	2	7	0
IPI00218493	8	HPRT1 Hypoxanthine-guanine phosphoribosyltransferase	13	1	7	0
IPI00876962	6	INF2 Isoform 2 of Inverted formin-2	13	2	7	0
IPI00479997	5	STMN1 Stathmin	13	1	7	0
IPI00791426	5	RPL24 13 kDa protein	13	1	7	0
IPI00215719	4	RPL18 60S ribosomal protein L18	13	1	7	0
IPI00334713	9	HNRNPAB Isoform 3 of Heterogeneous nuclear ribonucleoprotein A/B	12	3	7	0
IPI00012442	8	G3BP1 Ras GTPase-activating protein-binding protein 1	12	0	7	0
IPI00301311	8	SET Isoform 2 of Protein SET	12	2	7	0
IPI00011913	6	HNRNPA0 Heterogeneous nuclear ribonucleoprotein A0	12	2	7	0
IPI00413673	5	BCLAF1 Isoform 4 of Bcl-2-associated	12	3	7	0

IPI Acc. No.	nPeptides	Protein	nBigT	nSmallT	nBigP	nSmallP
		transcription factor 1				
IPI00059292	4	MAGOHB Protein mago nashi homolog 2	12	0	7	1
IPI00409635	3	FAM62B Isoform 2 of Extended synaptotagmin-2	12	0	7	0
IPI00555747	9	PABPC4 Isoform 2 of Polyadenylate-binding protein 4	11	2	7	0
IPI00604620	25	NCL Isoform 1 of Nucleolin	11	3	7	0
IPI00000877	24	HYOU1 Hypoxia up-regulated protein 1 precursor	11	2	7	0
IPI00141318	24	CKAP4 Isoform 1 of Cytoskeleton-associated protein 4	11	2	7	0
IPI00217468	10	HIST1H1B Histone H1.5	11	2	7	0
IPI00797148	15	HNRNPA1 HNRPA1 protein	11	2	7	0
IPI00438229	15	TRIM28 Isoform 1 of Transcription intermediary factor 1-beta	11	1	7	0
IPI00031812	6	YBX1 Nuclease-sensitive element-binding protein 1	11	2	7	0
IPI00221092	8	RPS16 40S ribosomal protein S16	11	2	7	0
IPI00396321	8	LRRC59 Leucine-rich repeat-containing protein 59	11	3	7	0
IPI00009659	4	C20orf77	11	1	7	0

IPI Acc. No.	nPeptides	Protein	nBigT	nSmallT	nBigP	nSmallP
		Uncharacterized protein C20orf77				
IPI00157790	4	KIAA0368 KIAA0368 protein	11	0	7	0
IPI00794978	2	MRPL47 MRPL47 protein	11	1	7	0
IPI00004524	2	GCA Grancalcin	11	0	7	0
IPI00217437	7	TTBK2 Tau-tubulin kinase	10	5	7	0
IPI00880104	3	KRT74 59 kDa protein	10	3	7	0
IPI00550661	3	KRT13 Isoform 2 of Keratin, type I cytoskeletal 13	10	5	7	0
IPI00382470	44	HSP90AA1 heat shock protein 90kDa alpha (cytosolic), class A member 1 isoform 1	10	1	7	0
IPI00651677	24	DDX17 Isoform 2 of Probable ATP-dependent RNA helicase DDX17	10	2	7	0
IPI00017617	20	DDX5 Probable ATP-dependent RNA helicase DDX5	10	3	7	0
IPI00010471	29	LCP1 Plastin-2	10	1	7	1
IPI00007765	34	HSPA9 Stress-70 protein, mitochondrial precursor	10	2	7	0
IPI00298520	16	ARCN1 Putative uncharacterized protein DKFZp686M09245	10	2	7	0
IPI00217030	14	RPS4X 40S ribosomal protein S4, X isoform	10	2	7	0
IPI00000494	13	RPL5 60S ribosomal	10	2	7	0

IPI Acc. No.	nPeptides	Protein	nBigT	nSmallT	nBigP	nSmallP
		protein L5				
IPI00012585	10	HEXB Beta-hexosaminidase subunit beta precursor	10	3	7	0
IPI00026105	10	SCP2 Isoform SCPx of Non-specific lipid-transfer protein	10	4	7	0
IPI00008455	8	MYO6 Isoform 2 of Myosin-VI	10	1	7	0
IPI00399183	7	APOB48R Isoform 1 of Apolipoprotein B-100 receptor	10	5	7	1
IPI00744364	6	SFRS7 Uncharacterized protein SFRS7	10	4	7	0
IPI00465132	5	COPE Coatomer subunit epsilon	10	1	7	0
IPI00607584	5	MYBBP1A Isoform 2 of Myb-binding protein 1A	10	0	7	0
IPI00787692	4	LOC650788 similar to 40S ribosomal protein S28	10	2	7	0
IPI00014808	3	PAFAH1B3 Platelet-activating factor acetylhydrolase IB subunit gamma	10	0	7	0
IPI00019449	3	LOC100133484;HLA-DQB1;HLA-DRB1;hCG_1998957;HLA-DRB4;LOC100133811;LOC100133583;HLA-DRB3;HLA-DRB5;HLA-DQB2;HLA-	10	9	7	1

IPI Acc. No.	nPeptides	Protein	nBigT	nSmallT	nBigP	nSmallP
		DRB2;ZNF749;LOC100133661;RNASE2 Non-secretory ribonuclease precursor				
IPI00217975	31	LMNB1 Lamin-B1	9	3	7	1
IPI00455599	12	HSP90AB2P Heat shock protein 90Bb	9	2	7	0
IPI00853059	14	FUBP1 Isoform 2 of Far upstream element-binding protein 1	9	1	7	0
IPI00218852	30	VIL1 Villin-1	9	3	7	0
IPI00789551	15	MATR3 Uncharacterized protein MATR3	9	1	7	0
IPI00410693	11	SERBP1 Isoform 1 of Plasminogen activator inhibitor 1 RNA-binding protein	9	4	7	0
IPI00646899	10	RPL10 Ribosomal protein L10	9	2	7	0
IPI00220362	9	HSPE1 10 kDa heat shock protein, mitochondrial	9	3	7	0
IPI00299573	9	RPL7A 60S ribosomal protein L7a	9	2	7	0
IPI00221354	5	FUS Isoform Short of RNA-binding protein FUS	9	2	7	0
IPI00893715	7	TACSTD1 38 kDa protein	9	4	7	0
IPI00872533	6	CD2AP 76 kDa protein	9	1	7	0
IPI00006181	6	EIF3D Eukaryotic translation initiation factor 3 subunit D	9	3	7	0

IPI Acc. No.	nPeptides	Protein	nBigT	nSmallT	nBigP	nSmallP
IPI00009368	5	SFXN1 Sideroflexin-1	9	2	7	1
IPI00219516	4	GUSB Isoform Short of Beta-glucuronidase precursor	9	2	7	0
IPI00398135	4	hCG_21078 hypothetical protein LOC389435	9	2	7	0
IPI00082310	3	KHDRBS1 Isoform 3 of KH domain-containing, RNA-binding, signal transduction-associated protein 1	9	1	7	0
IPI00328268	2	EIF4G3 EIF4G3 protein	9	0	7	0
IPI00644079	21	HNRNPU heterogeneous nuclear ribonucleoprotein U isoform a	8	2	7	0
IPI00414980	9	MYO1B Isoform 2 of Myosin-Ib	8	5	7	4
IPI00646304	13	PPIB peptidylprolyl isomerase B precursor	8	3	7	0
IPI00017334	13	PHB Prohibitin	8	1	7	0
IPI00744692	13	TALDO1 Transaldolase	8	1	7	0
IPI00329633	12	TARS Threonyl-tRNA synthetase, cytoplasmic	8	1	7	0
IPI00030009	11	PAPSS2 Isoform A of Bifunctional 3'-phosphoadenosine 5'-phosphosulfate synthetase 2	8	10	7	1
IPI00027851	7	HEXA Beta-hexosaminidase subunit alpha precursor	8	3	7	0
IPI00013174	7	RBM14 Isoform 1 of	8	3	7	0

IPI Acc. No.	nPeptides	Protein	nBigT	nSmallT	nBigP	nSmallP
		RNA-binding protein 14				
IPI00024933	7	RPL12 60S ribosomal protein L12	8	2	7	0
IPI00793375	6	XPNPEP1 Xaa-Pro aminopeptidase 1	8	1	7	0
IPI00219160	5	RPL34 60S ribosomal protein L34	8	4	7	0
IPI00010949	5	SIAE Isoform 1 of Sialate O-acetylesterase precursor	8	7	7	1
IPI00005737	4	SURF4 Isoform 1 of Surfeit locus protein 4	8	1	7	0
IPI00029631	3	ERH Enhancer of rudimentary homolog	8	0	7	0
IPI00029601	15	CTTN Src substrate cortactin	7	3	7	0
IPI00030179	11	RPL7 60S ribosomal protein L7	7	3	7	0
IPI00644570	3	18 kDa protein	7	4	7	0
IPI00456887	13	HNRNPUL2 Heterogeneous nuclear ribonucleoprotein U-like protein 2	7	0	7	0
IPI00031169	11	RAB2A Ras-related protein Rab-2A	7	0	7	0
IPI00016910	11	EIF3CL;EIF3C Eukaryotic translation initiation factor 3 subunit C	7	3	7	0
IPI00031583	11	USO1 Putative uncharacterized protein DKFZp451D234	7	4	7	4
IPI00165360	9	MPST 3-mercaptopyruvate	7	3	7	0

IPI Acc. No.	nPeptides	Protein	nBigT	nSmallT	nBigP	nSmallP
		sulfurtransferase				
IPI00025019	7	PSMB1 Proteasome subunit beta type-1 precursor	7	2	7	0
IPI00335930	4	DAZAP1 Isoform 2 of DAZ-associated protein 1	7	1	7	0
IPI00015029	4	PTGES3 Prostaglandin E synthase 3	7	2	7	0
IPI00029039	4	REG3A Regenerating islet-derived protein 3 alpha precursor	7	0	7	0
IPI00061525	3	GNPNAT1 Glucosamine 6-phosphate N-acetyltransferase	7	0	7	0
IPI00026167	3	NHP2L1 NHP2-like protein 1	7	0	7	0
IPI00000811	3	PSMB6 Proteasome subunit beta type-6 precursor	7	2	7	1
IPI00790799	3	SEC11A 16 kDa protein	7	0	7	0
IPI00413654	3	SFRS5 Isoform SRP40-4 of Splicing factor, arginine/serine-rich 5	7	2	7	0
IPI00006092	3	PMM2 Phosphomannomutase 2	7	1	7	0
IPI00328840	3	THOC4 THO complex subunit 4	7	1	7	0
IPI00216237	3	RPL36 60S ribosomal protein L36	7	3	7	1
IPI00173589	2	LOC284064 similar to ribosomal protein L29	7	3	7	0
IPI00009342	64	IQGAP1 Ras GTPase-	6	2	7	0

IPI Acc. No.	nPeptides	Protein	nBigT	nSmallT	nBigP	nSmallP
		activating-like protein IQGAP1				
IPI00413947	10	AP1B1 Isoform B of AP-1 complex subunit beta-1	6	0	7	0
IPI00398798	5	H2AFV H2A histone family, member V isoform 3	6	4	7	0
IPI00303882	13	M6PRBP1 Isoform B of Mannose-6-phosphate receptor-binding protein 1	6	2	7	0
IPI00790743	5	Protein	6	5	7	0
IPI00021800	8	CASP1 Isoform Alpha of Caspase-1 precursor	6	5	7	0
IPI00060181	7	EFHD2 EF-hand domain-containing protein D2	6	0	7	1
IPI00017510	6	MT-CO2 Cytochrome c oxidase subunit 2	6	6	7	0
IPI00006443	6	CRYL1 Lambda-crystallin homolog	6	6	7	0
IPI00020956	6	HDGF Hepatoma-derived growth factor	6	2	7	0
IPI00060200	5	GALM Aldose 1-epimerase	6	6	7	0
IPI00185374	5	PSMD12 26S proteasome non-ATPase regulatory subunit 12	6	1	7	0
IPI00012340	4	SFRS9 Splicing factor, arginine/serine-rich 9	6	3	7	1
IPI00016405	4	OCIAD1 Isoform 1 of OCIA domain-containing protein 1	6	5	7	0

IPI Acc. No.	nPeptides	Protein	nBigT	nSmallT	nBigP	nSmallP
IPI00019329	3	DYNLL1 Dynein light chain 1, cytoplasmic	6	4	7	0
IPI00642816	3	SRP9;hCG_1781062 Signal recognition particle 9 kDa protein	6	2	7	0
IPI00013968	3	COX7C Cytochrome c oxidase subunit 7C, mitochondrial precursor	6	2	7	0
IPI00026570	2	COX7A2 Cytochrome c oxidase polypeptide VIIa-liver/heart, mitochondrial precursor	6	4	7	0
IPI00005719	10	RAB1A Isoform 1 of Ras-related protein Rab-1A	5	2	7	0
IPI00374686	6	Uncharacterized protein ENSP00000341227 (Fragment)	5	1	7	0
IPI00219077	15	LTA4H Isoform 1 of Leukotriene A-4 hydrolase	5	2	7	0
IPI00025366	13	CS Citrate synthase, mitochondrial precursor	5	1	7	1
IPI00290110	13	PDCD4 Programmed cell death protein 4	5	5	7	0
IPI00332828	10	CES2 carboxylesterase 2 isoform 1	5	7	7	5
IPI00182757	7	KIAA1967 Isoform 1 of Protein KIAA1967	5	3	7	0
IPI00030654	6	CPSF6 Isoform 2 of Cleavage and polyadenylation specificity factor subunit 6	5	1	7	0

IPI Acc. No.	nPeptides	Protein	nBigT	nSmallT	nBigP	nSmallP
IPI00789848	5	IVD Isovaleryl-CoA dehydrogenase, mitochondrial precursor	5	7	7	0
IPI00024934	4	MUT Methylmalonyl-CoA mutase, mitochondrial precursor	5	0	7	0
IPI00015972	4	COX6C Cytochrome c oxidase polypeptide VIc precursor	5	6	7	0
IPI00031772	3	C15orf48 Normal mucosa of esophagus-specific gene 1 protein	5	7	7	0
IPI00019888	2	ALDH5A1 Succinate-semialdehyde dehydrogenase, mitochondrial precursor	5	0	7	0
IPI00873259	2	ATP5J2 10 kDa protein	5	2	7	0
IPI00005159	11	ACTR2 Actin-related protein 2	4	1	7	0
IPI00465256	8	AK3 GTP:AMP phosphotransferase mitochondrial	4	2	7	1
IPI00513827	8	ACADM Putative uncharacterized protein DKFZp686M24262	4	4	7	0
IPI00554811	6	ARPC4;TTLL3 Actin-related protein 2/3 complex subunit 4	4	2	7	0
IPI00874156	6	OTUB1 Isoform 1 of Ubiquitin thioesterase OTUB1	4	2	7	0
IPI00024661	4	SEC24C Protein transport protein Sec24C	4	2	7	1
IPI00028387	4	C20orf116 Isoform 1 of	4	1	7	0

IPI Acc. No.	nPeptides	Protein	nBigT	nSmallT	nBigP	nSmallP
		Uncharacterized protein C20orf116 precursor				
IPI00022277	4	CCDC56 Coiled-coil domain-containing protein 56	4	6	7	0
IPI00024742	2	UQCRQ Cytochrome b-c1 complex subunit 8	4	3	7	0
IPI00307547	2	C9orf46 Uncharacterized protein C9orf46	4	1	7	0
IPI00029133	10	ATP5F1 ATP synthase subunit b, mitochondrial precursor	3	3	7	0
IPI00651719	6	PBLD MAWD binding protein isoform b	3	2	7	1
IPI00219755	1	SPCS1 Signal peptidase complex subunit 1	3	3	7	0
IPI00880101	9	CEACAM5 Protein	29	0	6	0
IPI00302944	51	COL12A1 Isoform 4 of Collagen alpha-1(XII) chain precursor	24	3	6	0
IPI00027769	5	ELA2 Leukocyte elastase precursor	23	0	6	0
IPI00026781	46	FASN Fatty acid synthase	20	2	6	0
IPI00304754	5	FERMT1 Isoform 1 of Fermitin family homolog 1	18	3	6	0
IPI00019472	5	SLC1A5 Neutral amino acid transporter B	16	0	6	0
IPI00026833	5	ADSS Adenylosuccinate synthetase isozyme 2	16	0	6	0
IPI00026328	3	TXNDC12 Thioredoxin domain-containing	14	2	6	0

IPI Acc. No.	nPeptides	Protein	nBigT	nSmallT	nBigP	nSmallP
		protein 12 precursor				
IPI00009904	36	PDIA4 Protein disulfide-isomerase A4 precursor	13	3	6	0
IPI00216044	8	RALY Isoform 1 of RNA-binding protein Raly	13	0	6	0
IPI00215790	6	RPL38 60S ribosomal protein L38	13	2	6	0
IPI00872940	6	RPL30 Uncharacterized protein RPL30 (Fragment)	13	1	6	0
IPI00012007	11	AHCY Adenosylhomocysteinase	12	2	6	0
IPI00012069	6	NQO1 NAD	12	1	6	0
IPI00385267	5	SRPR Signal recognition particle receptor subunit alpha	12	1	6	0
IPI00856058	4	RPL31 ribosomal protein L31 isoform 3	12	3	6	0
IPI00395865	3	RBBP7 Histone-binding protein RBBP7	12	0	6	0
IPI00215734	3	PRMT1 Isoform 2 of Protein arginine N-methyltransferase 1	12	0	6	0
IPI00554648	51	KRT8 Keratin, type II cytoskeletal 8	11	3	6	0
IPI00027201	64	MUC2 Mucin-2 precursor	11	4	6	1
IPI00789401	22	PLS1 PLS1 protein	11	2	6	0
IPI00215743	48	RRBP1 Isoform 3 of Ribosome-binding protein 1	11	3	6	0

IPI Acc. No.	nPeptides	Protein	nBigT	nSmallT	nBigP	nSmallP
IPI00003881	11	HNRNPF Heterogeneous nuclear ribonucleoprotein F	11	2	6	0
IPI00410017	18	PABPC1 Isoform 2 of Polyadenylate-binding protein 1	11	2	6	0
IPI00022228	21	HDLBP Vigilin	11	2	6	0
IPI00401264	11	TXNDC4 Thioredoxin domain-containing protein 4 precursor	11	2	6	0
IPI00008438	9	RPS10 40S ribosomal protein S10	11	2	6	0
IPI00295741	9	CTSB Cathepsin B precursor	11	5	6	0
IPI00004968	7	PRPF19 Pre-mRNA-processing factor 19	11	0	6	1
IPI00014230	6	C1QBP Complement component 1 Q subcomponent-binding protein, mitochondrial precursor	11	2	6	0
IPI00646917	4	NUDT21 Cleavage and polyadenylation specificity factor subunit 5	11	0	6	0
IPI00219155	4	RPL27 60S ribosomal protein L27	11	1	6	0
IPI00218606	4	RPS23 40S ribosomal protein S23	11	5	6	0
IPI00101405	4	FDPS Farnesyl diphosphate synthase	11	2	6	0
IPI00028006	3	PSMB2 Proteasome subunit beta type-2	11	2	6	0
IPI00386662	2	FUSIP1 Isoform 4 of	11	1	6	0

IPI Acc. No.	nPeptides	Protein	nBigT	nSmallT	nBigP	nSmallP
		FUS-interacting serine-arginine-rich protein 1				
IPI00027230	43	HSP90B1 Endoplasmin precursor	10	2	6	0
IPI00027834	15	HNRNPL heterogeneous nuclear ribonucleoprotein L isoform a	10	1	6	0
IPI00219870	23	CTNND1 Isoform 1A of Catenin delta-1	10	3	6	0
IPI00383296	21	HNRNPM Isoform 2 of Heterogeneous nuclear ribonucleoprotein M	10	2	6	0
IPI00020599	19	CALR Calreticulin precursor	10	2	6	0
IPI00644989	18	PDIA6 Isoform 1 of Protein disulfide-isomerase A6 precursor	10	2	6	0
IPI00646486	9	HP1BP3 Heterochromatin protein 1, binding protein 3	10	2	6	0
IPI00010896	15	CLIC1 Chloride intracellular channel protein 1	10	2	6	0
IPI00000690	13	AIFM1 Isoform 1 of Apoptosis-inducing factor 1, mitochondrial precursor	10	4	6	0

nPeptides, number of peptides that were identified in all the samples analyzed; nBigT, number of patients in which the ratio (expression level in tumor tissue/ expression level in healthy tissue) was >3; nSmallT, number of patients in which the ratio (expression level in tumor tissue/ expression level in healthy tissue) was <1/3; nBigP, number of patients in which the ratio (expression level in polyp tissue/ expression level in healthy tissue) was >3; nSmallP, number of patients in which the ratio (expression level in polyp tissue/ expression level in healthy tissue) was <1/3.

Example 3. Identification of a group of protein markers that are more highly expressed in polyps than in advanced stages of colorectal cancer.

Table 4 lists proteins that were observed to be highly expressed in polyps, whereas in more advanced stages of colorectal cancer these proteins tended to have decreased levels of expression. Accordingly, a diagnostic array of reagents directed to detection of at least some of the proteins in this group may be used as a screening test for very early detection of colorectal cancer. Such a screening test could identify susceptible at-risk individuals, even prior to the stage at which polyp visualization is possible by endoscopic techniques.

Table 4.

Protein (IPI Acc. No.)	nBigTumor	nSmallTumor	nBigPolyp	nSmallPolyp
CPT2 Carnitine O-palmitoyltransferase 2, mitochondrial precursor (IPI00012912)	6	5	12	0
ARL1 ADP-ribosylation factor-like protein 1 (IPI00219518)	6	1	12	0
PFKL Isoform 1 of 6-phosphofructokinase, liver type (IPI00332371)	6	2	11	0
GOT2 Aspartate aminotransferase, mitochondrial precursor (IPI00018206)	6	3	11	0
AP1G1 AP-1 complex subunit gamma-1 (IPI00643591)	6	1	11	0
STRBP Isoform 2 of Spermatid perinuclear RNA-binding protein (IPI00413860)	6	0	11	0
CLCA1 Calcium-activated chloride channel regulator 1 precursor (IPI00014625)	4	13	11	1
CYFIP1 Isoform 1 of Cytoplasmic FMR1-interacting protein 1 (IPI00644231)	4	2	11	0
COQ9 Isoform 1 of Ubiquinone biosynthesis protein COQ9, mitochondrial precursor (IPI00470631)	4	3	11	0
NDUFA9 NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 9, mitochondrial precursor (IPI00003968)	3	4	11	0

Protein (IPI Acc. No.)	nBigTumor	nSmallTumor	nBigPolyp	nSmallPolyp
ALDH7A1 Similar to Antiquitin (IPI00221234)	2	5	11	0
HMGCS1 Hydroxymethylglutaryl-CoA synthase, cytoplasmic (IPI00008475)	6	4	10	0
NNT NAD(P) transhydrogenase, mitochondrial precursor (IPI00337541)	6	2	10	0
PRDX5 Uncharacterized protein PRDX5 (Fragment) (IPI00876999)	6	2	10	0
PCCB Propionyl-CoA carboxylase beta chain, mitochondrial precursor (IPI00007247)	6	5	10	0
COPZ1 Coatomer subunit zeta-1 (IPI00032851)	6	4	10	0
BAX BCL2-associated X protein isoform sigma (IPI00845474)	6	2	10	0
ACAD9 Acyl-CoA dehydrogenase family member 9, mitochondrial precursor (IPI00152981)	6	0	10	0
UBXD8 UBX domain-containing protein 8 (IPI00172656)	6	1	10	0
HMGCS2 Hydroxymethylglutaryl-CoA synthase, mitochondrial precursor (IPI00008934)	5	12	10	0
SLC25A3 cDNA FLJ90278 fis, clone NT2RP1000325, highly similar to Phosphate carrier protein, mitochondrialprecursor (IPI00790115)	5	1	10	0
SLC25A11 Mitochondrial 2-oxoglutarate/malate carrier protein (IPI00219729)	5	4	10	0
PDCD6 Programmed cell death protein 6 (IPI00025277)	5	0	10	0
UCRC Cytochrome b-c1 complex subunit 9 (IPI00554701)	4	0	10	0
DEFA6 Defensin-6 precursor (IPI00008301)	4	0	10	0
DYNC1H1 Cytoplasmic dynein 1 heavy chain 1 (IPI00456969)	3	1	10	0

Protein (IPI Acc. No.)	nBigTumor	nSmallTumor	nBigPolyp	nSmallPolyp
HK1 Isoform 2 of Hexokinase-1 (IPI00220663)	3	1	10	0
CYFIP2 Isoform 2 of Cytoplasmic FMR1-interacting protein 2 (IPI00719600)	3	2	10	0
DCI Isoform 2 of 3,2-trans-enoyl-CoA isomerase, mitochondrial precursor (IPI00398758)	3	5	10	0
CISD1 CDGSH iron sulfur domain-containing protein 1 (IPI00020510)	1	1	10	0

nBigTumor, number of patients in which the ratio (expression level in tumor tissue/ expression level in healthy tissue) was >3; nSmallTumor, number of patients in which the ratio (expression level in tumor tissue/ expression level in healthy tissue) was <1/3; nBigPolyp, number of patients in which the ratio (expression level in polyp tissue/ expression level in healthy tissue) was >3; nSmallPolyp, number of patients in which the ratio (expression level in polyp tissue/ expression level in healthy tissue) was <1/3.

Example 4. Cancer associated proteins identified by isotopic labeling.

The proteins identified by isotopic labeling included those listed in Table 5. All of the listed proteins were found to be present in colon cancer tissue at levels that were at least 2.5-fold greater than the level of the same protein in healthy colon tissue from the same subject, as indicated by the median ratios of the protein levels (i.e. cancerous tissue:healthy tissue) listed in the column denoted "MED".

Table 5.

Protein Name	IPI Acc. No.	nPEP	MED	#>2.5
DPEP1 Dipeptidase 1 precursor	IPI00059476	12	50	8
LCN2 Lipocalin 2	IPI00643623	6	48.75	13
FAM62B Isoform 2 of Extended synaptotagmin-2	IPI00409635	10	39.73	10
MTA2 Metastasis-associated protein MTA2	IPI00171798	8	31.45	8
S100A8 Protein S100-A8	IPI00007047	7	30.72	17
S100A9 Protein S100-A9	IPI00027462	7	28.205	17
MPO Isoform H7 of Myeloperoxidase precursor	IPI00236556	26	25.57	14
MCM2 DNA replication licensing factor MCM2	IPI00184330	12	23.37	11
FDFT1 Squalene synthetase	IPI00020944	5	23.155	7
DMBT1 Isoform 1 of Deleted in malignant brain tumors 1 protein precursor	IPI00099110	7	22.86	7

Protein Name	IPI Acc. No.	nPEP	MED	#>2.5
LTF Lactotransferrin precursor	IPI00848342	30	20.45	17
SERPINB5 Serpin B5 precursor	IPI00783625	7	20.44	8
OLFM4 Olfactomedin-4 precursor	IPI00022255	20	19.49	13
FERMT1 Isoform 1 of Fermitin family homolog 1	IPI00304754	9	16.05	14
C1R;ACYP1;C17orf13 Complement C1r subcomponent precursor	IPI00296165	5	15.61	7
PLOD3 Procollagen-lysine,2-oxoglutarate 5-dioxygenase 3 precursor	IPI00030255	10	15.17	8
CEACAM5 Carcinoembryonic antigen-related cell adhesion molecule 5 precursor	IPI00027486	9	14.12	17
PYCR1 pyrroline-5-carboxylate reductase 1 isoform 2	IPI00376503	6	12.505	12
MCM6 DNA replication licensing factor MCM6	IPI00031517	9	12.29	6
SLC1A5 Neutral amino acid transporter B	IPI00019472	8	10.03	10
THBS1 Thrombospondin-1 precursor	IPI00296099	13	8.55	14
LACTB2 Beta-lactamase-like protein 2	IPI00006952	5	8.445	8
NAT10 N-acetyltransferase 10	IPI00300127	11	7.85	10
RSL1D1 RSL1D1 protein	IPI00642046	11	7.845	12
LMO7 Isoform 3 of LIM domain only protein 7	IPI00291802	11	7.74	10
LYZ Lysozyme C precursor	IPI00019038	7	7.64	15
MCM7 Isoform 1 of DNA replication licensing factor MCM7	IPI00299904	10	7.435	7
F11R Junctional adhesion molecule A precursor	IPI00001754	5	6.775	7
MCM3 DNA replication licensing factor MCM3	IPI00013214	11	6.605	10
ATP6V1E1 vacuolar H+ ATPase E1 isoform b	IPI00719806	5	6.34	9
TNC Isoform 1 of Tenascin precursor	IPI00031008	58	6.2	12
COL12A1 Isoform 1 of Collagen alpha-1(XII) chain precursor	IPI00329573	49	6.175	17
SORD 11 kDa protein	IPI00791243	5	5.78	7
PYCR2 Pyrroline-5-carboxylate reductase 2	IPI00470610	5	5.645	9
GTF2I Isoform 2 of General transcription factor II-I	IPI00293242	10	5.515	7
DDX18 ATP-dependent RNA helicase DDX18	IPI00301323	8	5.38	8
RBM39 Isoform 2 of RNA-binding protein 39	IPI00215801	11	5.19	12
NQO1 NAD	IPI00012069	6	4.9	8

Protein Name	IPI Acc. No.	nPEP	MED	#>2.5
DNAJA3 Isoform 2 of DnaJ homolog subfamily A member 3, mitochondrial precursor	IPI00179187	7	4.73	9
NCBP1 Nuclear cap-binding protein subunit 1	IPI00019380	8	4.42	9
HSPH1 Isoform Beta of Heat shock protein 105 kDa	IPI00218993	33	4.365	13
ADSS Adenylosuccinate synthetase isozyme 2	IPI00026833	7	4.36	11
PSAT1 Isoform 1 of Phosphoserine aminotransferase	IPI00001734	9	4.295	8
ALG5 Dolichyl-phosphate beta-glucosyltransferase	IPI00002506	6	4.135	10
PCNA Proliferating cell nuclear antigen	IPI00021700	12	4.04	10
TCOF1 Isoform 2 of Treacle protein	IPI00298696	12	3.92	9
SERPINH1 Serpin H1 precursor	IPI00032140	20	3.9	16
ERO1L ERO1-like protein alpha precursor	IPI00386755	8	3.81	12
ILVBL Isoform 1 of Acetolactate synthase-like protein	IPI00554541	15	3.8	11
ANXA3 Annexin A3	IPI00024095	17	3.645	16
NAMPT Isoform 1 of Nicotinamide phosphoribosyltransferase	IPI00018873	25	3.61	12
TFRC Transferrin receptor protein 1	IPI00022462	23	3.56	9
SERPINB9 Serpin B9	IPI00032139	8	3.41	10
EIF2S2 Eukaryotic translation initiation factor 2 subunit 2	IPI00021728	8	3.38	8
SRRM2 Isoform 1 of Serine/arginine repetitive matrix protein 2	IPI00782992	8	3.365	8
ARHGEF1 Isoform 1 of Rho guanine nucleotide exchange factor 1	IPI00647786	5	3.31	7
COMT Isoform Soluble of Catechol O-methyltransferase	IPI00375513	9	3.31	10
DEK 48 kDa protein	IPI00871695	8	3.285	10
SYK Isoform Long of Tyrosine-protein kinase SYK	IPI00018597	6	3.275	7
S100A11 Protein S100-A11	IPI00013895	8	3.2	12
HSDL2 Isoform 1 of Hydroxysteroid dehydrogenase-like protein 2	IPI00414384	10	3.14	10
C7orf24 Uncharacterized protein C7orf24	IPI00031564	5	3.125	9
HM13 Isoform 1 of Minor histocompatibility antigen H13	IPI00152441	7	3.09	9
RCN1 Reticulocalbin-1 precursor	IPI00015842	14	3.07	11
DIAPH1 Diaphanous homolog 1	IPI00884341	10	3.06	8
SRM Spermidine synthase	IPI00292020	5	2.97	11

Protein Name	IPI Acc. No.	nPEP	MED	#>2.5
ATAD3A Isoform 2 of ATPase family AAA domain-containing protein 3A	IPI00295992	20	2.965	10
GPX2 Glutathione peroxidase 2	IPI00298176	9	2.945	9
LOC442497;SLC3A2 solute carrier family 3 (activators of dibasic and neutral amino acid transport), member 2 isoform e	IPI00554722	13	2.9	12
SAE1 SUMO-activating enzyme subunit 1	IPI00033130	8	2.89	9
IPO7 Importin-7	IPI00007402	11	2.85	9
SET Isoform 2 of Protein SET	IPI00301311	9	2.85	11
PAICS Multifunctional protein ADE2	IPI00217223	15	2.84	11
OCIAD2 Isoform 1 of OCIA domain-containing protein 2	IPI00555902	6	2.83	10
GLT25D1 Glycosyltransferase 25 family member 1 precursor	IPI00168262	5	2.82	7
RCC2 Protein RCC2	IPI00465044	12	2.79	10
CTSG Cathepsin G precursor	IPI00028064	10	2.77	12
CHD4 Isoform 2 of Chromodomain-helicase-DNA-binding protein 4	IPI00455210	10	2.765	8
SSBP1 Single-stranded DNA-binding protein, mitochondrial precursor	IPI00029744	6	2.765	10
ACOT7 Isoform 1 of Cytosolic acyl coenzyme A thioester hydrolase	IPI00010415	6	2.76	8
AK3 GTP:AMP phosphotransferase mitochondrial	IPI00465256	13	2.72	9
GCA Grancalcin	IPI00004524	6	2.71	9
ACIN1 Isoform 1 of Apoptotic chromatin condensation inducer in the nucleus	IPI00007334	5	2.69	9
TM9SF4 Isoform 2 of Transmembrane 9 superfamily member 4 precursor	IPI00885106	8	2.69	8
CAD Putative uncharacterized protein CAD	IPI00893035	17	2.675	9
FASN Fatty acid synthase	IPI00026781	74	2.66	11
TJP2 Isoform A1 of Tight junction protein ZO-2	IPI00003843	16	2.66	8
GLRX3 Glutaredoxin-3	IPI00008552	6	2.595	7
RCC1 regulator of chromosome condensation 1 isoform b	IPI00787306	5	2.59	7
TOP1 DNA topoisomerase 1	IPI00413611	10	2.57	9
GGH Gamma-glutamyl hydrolase precursor	IPI00023728	11	2.565	10
PUF60 Isoform 5 of Poly	IPI00856076	10	2.56	9
TPR nuclear pore complex-associated	IPI00742682	16	2.545	9

Protein Name	IPI Acc. No.	nPEP	MED	#>2.5
protein TPR				
GMPS GMP synthase	IPI00029079	11	2.525	10
PKP3 Plakophilin-3	IPI00026952	19	2.52	10
LOC731605 similar to BCL2-associated transcription factor 1 isoform 2	IPI00886854	7	2.515	9
SLC2A1 Solute carrier family 2, facilitated glucose transporter member 1	IPI00220194	8	2.515	7
C8orf55 Uncharacterized protein C8orf55 precursor	IPI00171421	7	2.51	8
HCFC1 Uncharacterized protein HCFC1	IPI00641743	7	2.51	10

IPI Acc. No, accession number in IPI database; nPEP, number of identified peptides in all the samples, MED, median of the ratio tumor/healthy from all the patients; #>2.5, number of patients in which tumor/healthy >2.5.

Example 5. Cancer associated proteins identified by multidimensional chromatography.

The proteins identified on the basis of the peptides detected by the multidimensional chromatography technique include those listed in Table 6, and reflects the abundance of these proteins in cancerous tissues. None of the proteins listed in Table 6 were identified by the isotopic labeling technique. The proteins listed in Tables 3 and 6 appear to be specifically expressed in colorectal cancer tissue, as they were substantially undetectable in all healthy colorectal tissues obtained from the colorectal cancer patients.

Table 6.

Protein Name	IPI Acc. No.	AVG	SD	No. of tumors
ADAMDEC1 ADAM DEC1 precursor	4480	50	0	
AMACR Alpha-methylacyl-CoA racemase	847727	50	0	7
AMACR;C1QTNF3 alpha-methylacyl-CoA racemase isoform 1	5918	50	0	7
ARID1A Isoform 1 of AT-rich interactive domain-containing protein 1A	643722	50	0	
CEBPZ CCAAT/enhancer-binding protein zeta	306723	50	0	
COL5A1 Collagen alpha-1(V) chain precursor	844090	50	0	
EFEMP2 Mutant p53 binding protein 1 variant (Fragment)	556657	50	0	
FAM84B Protein FAM84B	64666	50	0	
FKBP10 FK506-binding protein 10 precursor	303300	50	0	8

Protein Name	IPI Acc. No.	AVG	SD	No. of tumors
FKBP9 FK506-binding protein 9 precursor	182126	50	0	9
GPRC5A Retinoic acid-induced protein 3	22624	50	0	8
KPNA2 Karyopherin alpha 2	789457	50	0	
MMP1 Interstitial collagenase precursor	8561	50	0	
PNMA5 Paraneoplastic antigen-like protein 5	514588	50	0	
POLR1C Isoform 1 of DNA-directed RNA polymerases I and III subunit RPAC1	5179	50	0	9
SPARC SPARC precursor	14572	50	0	5
UBAP2 Ubiquitin-associated protein 2	171127	50	0	5
UCK2 Isoform 1 of Uridine-cytidine kinase 2	65671	50	0	
WDR74 Isoform 1 of WD repeat-containing protein 74	18192	50	0	

Example 6. Cancer associated proteins as potential therapeutic targets.

Table 7 lists proteins that are considered potential targets for development of cytotoxic reagents specifically directed to these proteins, for example, specific antibody or antibody fragments conjugated to toxic moieties for targeted elimination of cancer cells. As well as being highly expressed in early stage polyps and in tumors, these proteins are generally exposed, and considered vulnerable to attack by targeted cytotoxic reagents.

Table 7.

Protein (IPI Acc. No.)	nBigTumor	nSmallTumor	nBigPolyp	nSmallPolyp	Ontology
DEFA3 Neutrophil defensin 3 precursor (IPI00021827)	26	2	5	0	EX
ELA2 Leukocyte elastase precursor (IPI00027769)	23	0	6	0	EX
LYZ Lysozyme C precursor (IPI00019038)	17	2	8	0	EX
LOC442497;SLC3A2 solute carrier family 3 (activators of dibasic and neutral amino acid transport), member 2 isoform e (IPI00554722)	16	1	8	0	PM

Protein (IPI Acc. No.)	nBigTumor	nSmallTumor	nBigPolyp	nSmallPolyp	Ontology
SLC1A5 Neutral amino acid transporter B (IPI00019472)	16	0	6	0	PM
DMBT1 Isoform 1 of Deleted in malignant brain tumors 1 protein precursor (IPI00099110)	15	0	8	1	EX AM
NUCB1 Nucleobindin-1 precursor (IPI00295542)	15	3	7	0	EX AM
SLC12A2 Isoform 1 of Solute carrier family 12 member 2 (IPI00022649)	14	2	16	0	PM
GGH Gamma-glutamyl hydrolase precursor (IPI00023728)	14	1	12	0	EX
AGR3 Anterior gradient protein 3 homolog precursor (IPI00152409)	14	4	7	0	EX
MARCKSL1 MARCKS-related protein (IPI00641181)	13	0	11	0	PM
TM9SF2 Transmembrane 9 superfamily member 2 precursor (IPI00018415)	12	1	9	0	PM
SYK Isoform Long of Tyrosine-protein kinase SYK (IPI00018597)	11	0	10	0	PM
GCA Grancalcin (IPI00004524)	11	0	7	0	PM
HDLBP Vigilin (IPI00022228)	11	2	6	0	PM
C1QBP Complement component 1 Q subcomponent-binding protein, mitochondrial precursor (IPI00014230)	11	2	6	0	PM
KIAA0152 Uncharacterized protein KIAA0152 precursor (IPI00029046)	10	0	13	0	PM

Protein (IPI Acc. No.)	nBigTumor	nSmallTumor	nBigPolyp	nSmallPolyp	Ontology
CLIC1 Chloride intracellular channel protein 1 (IPI00010896)	10	2	6	0	PM

nBigTumor, number of patients in which the ratio (expression level in tumor tissue/ expression level in healthy tissue) was >3; nSmallTumor, number of patients in which the ratio (expression level in tumor tissue/ expression level in healthy tissue) was <1/3; nBigPolyp, number of patients in which the ratio (expression level in polyp tissue/ expression level in healthy tissue) was >3; nSmallPolyp, number of patients in which the ratio (expression level in polyp tissue/ expression level in healthy tissue) was <1/3; EX, extracellular region; PM, plasma membrane; AM, additional membrane.

References cited

Aebersold, R. and M. Mann (2003). "Mass spectrometry-based proteomics." Nature 422(6928): 198-207.

Beer, I., E. et al. (2004). "Improving large-scale proteomics by clustering of mass spectrometry data." Proteomics 4(4): 950-60.

Eng, J., et al. (1994). "An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database." J. Amer. Soc. Mass. Spect. 5: 976-989.

Hsu, J. L., et al. (2003). "Stable-isotope dimethyl labeling for quantitative proteomics." Anal Chem 75(24): 6843-52.

Ishihama, Y., et al. (2002). "Microcolumns with self-assembled particle frits for proteomics." J Chromatogr A 979(1-2): 233-9.

Link, A. J. (2002). "Multidimensional peptide separations in proteomics." Trends Biotechnol 20(12 Suppl): S8-13.

Link, A. J. et al. (1999). "Direct analysis of protein complexes using mass spectrometry." Nat. Biotechnol. 17(7): 676-82.

Ong, S. E. and M. Mann (2005). "Mass spectrometry-based proteomics turns quantitative." Nat Chem Biol 1(5): 252-62.

Perkins, D. N. et al. (1999). "Probability-based protein identification by searching sequence databases using mass spectrometry data." Electrophoresis 20(18): 3551-67.

Regnier, F. E. and S. Julka (2006). "Primary amine coding as a path to comparative proteomics." Proteomics 6(14): 3968-79.

Ross, P. L. et al. (2004). "Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents." Mol Cell Proteomics 3(12): 1154-69.

The foregoing description of the specific embodiments will so fully reveal the general nature of the invention that others can, by applying current knowledge, readily modify and/or adapt for various applications such specific embodiments without undue experimentation and without departing from the brand concept, and, therefore, such adaptations and modifications should and are intended to be comprehended within the meaning and range of equivalents of the disclosed embodiments. It is to be understood that the phraseology or terminology employed herein is for the purpose of description and not of limitation. The means, materials, and steps for carrying out various disclosed functions may take a variety of alternative forms without departing from the invention.

CLAIMS

1. A method of detecting cancer in a subject, the method comprising:

(i) detecting in a biological sample from the subject at least one protein selected from the group consisting of KIAA0152 (SEQ ID NO:1), NAMPT (SEQ ID NO:2), PYCR1 (SEQ ID NO:3), GPX2 (SEQ ID NO:4), PRKDC (SEQ ID NO:5), ALDH18A1 (SEQ ID NO:6), OCIAD2 (SEQ ID NO:7), GCS1 (SEQ ID NO:8), GMDS (SEQ ID NO:9), ARF4 (SEQ ID NO:10), ARF5 (SEQ ID NO:11), LRPPRC (SEQ ID NO:12), CTNNB1 (SEQ ID NO:13), ARF3 (SEQ ID NO:14), GCN1L1 (SEQ ID NO:15), BDH1 (SEQ ID NO:16), RPL9 (SEQ ID NO:17), UGCGL1 (SEQ ID NO:18), FAM3D (SEQ ID NO:19) and CCT4 (SEQ ID NO:20), so as to determine the level of the at least one protein; and

(ii) comparing the level determined in (i) to a reference level of the same at least one protein, wherein detection of a level of said at least one protein in the biological sample which is significantly different from the reference level, is indicative of cancer in the subject.

2. The method according to claim 1, wherein the biological sample is selected from the group consisting of blood, serum, nipple aspirate fluid, lymph node aspirate, a biopsy sample, a tumor sample, a tissue sample, mucosal fluid, cervical wash, lacrimal duct fluid, urine, saliva, pleural effusion and sputum.

3. The method according to claim 2, wherein the tissue sample comprises gastrointestinal tissue.

4. The method according to claim 1, wherein the cancer is colorectal cancer.

5. The method according to claim 4, wherein the colorectal cancer is selected from the group consisting of pre-cancerous polyps, early stage colorectal cancer and advanced stage colorectal cancer.

6. The method according to claim 4, comprising detecting at least one of KIAA0152 (SEQ ID NO:1) and NAMPT (SEQ ID NO:2).

7. The method according to claim 4, further comprising detecting at least one protein selected from the group consisting of CPT2 (SEQ ID NO:21), ARL1 (SEQ ID NO:22), PFKL (SEQ ID NO:23), GOT2 (SEQ ID NO:24), AP1G1 (SEQ ID NO:25), STRBP (SEQ ID NO:26), CLCA1 (SEQ ID NO:27), CYFIP1 (SEQ ID

NO:28), COQ9 (SEQ ID NO:29), NDUFA9 (SEQ ID NO:30), ALDH7A1 (SEQ ID NO:31), HMGCS1 (SEQ ID NO:32), NNT (SEQ ID NO:33), PRDX5 (SEQ ID NO:34), PCCB (SEQ ID NO:35), COPZ1 (SEQ ID NO:36), BAX (SEQ ID NO:37), ACAD9 (SEQ ID NO:38), UBXD8 (SEQ ID NO:39), HMGCS2 (SEQ ID NO:40), SLC25A3 (SEQ ID NO:41), SLC25A11 (SEQ ID NO:42), PDCD6 (SEQ ID NO:43), UCRC (SEQ ID NO:44), DEFA6 (SEQ ID NO:45), DYNC1H1 (SEQ ID NO:46), HK1 (SEQ ID NO:47), CYFIP2 (SEQ ID NO:48), DCI (SEQ ID NO:49) and CISD1 (SEQ ID NO:50).

8. The method according to claim 7, wherein the colorectal cancer comprises pre-cancerous polyps or early stage colorectal cancer.

9. The method according to claim 3, wherein the gastrointestinal tissue is obtained from the subject by a procedure selected from the group consisting of biopsy, flexible endoscopy, double balloon endoscopy and surgical colorectal re-sectioning.

10. The method according to claim 3, wherein the gastrointestinal tissue is assessed *in vivo*.

11. The method according to claim 10, comprising contacting gastrointestinal tissue with at least one of a pharmaceutical composition and an endoscopy apparatus.

12. The method according to claim 11, comprising administering the pharmaceutical composition to the subject by a route selected from the group consisting of oral, parenteral, subcutaneous, intramuscular, intrathoracic and intraarticular.

13. The method according to any one of claims 1, 7, 9 or 10, comprising use of at least one reagent suitable for detecting the level of said at least one protein.

14. The method according to claim 11, wherein the pharmaceutical composition or endoscopy apparatus comprise at least one reagent suitable for detecting the level of at said at least one protein.

15. The method according to claim 14, wherein said at least one reagent is suitable for detecting the level of at least one of KIAA0152 (SEQ ID NO:1) and NAMPT (SEQ ID NO:2).

16. The method according to claim 13, wherein said at least one reagent specifically interacts with said at least one protein or with nucleic acid encoding said at least

one protein or a fragment thereof.

17. The method according to claim 13, wherein said reagent is selected from an antibody, an antibody mimetic and a nucleic acid.

18. The method according to claim 14, wherein the at least one reagent comprises a multiplicity of reagents, wherein each reagent of the multiplicity has specifically for a distinct protein.

19. The method according to claim 16, wherein said at least one reagent comprises a detectable label.

20. The method according to claim 1, comprising use of an assay system selected from the group consisting of an immunoassay, a nucleic acid hybridization assay, a binding assay, an array, a phage display library and a combination thereof.

21. The method according to claim 3, comprising contacting gastrointestinal tissue with at least one of a pharmaceutical composition and an endoscopy apparatus, and further comprising use of an external monitoring system, wherein the external monitoring system is configured to display the level of the at least one protein detected.

22. The method according to claim 4, wherein the level of the at least one protein in the sample is increased by at least at least 3-fold relative to the reference level, and wherein the reference level is representative of the level of the same protein in non-diseased tissue.

23. A method for determining the stage of a cancerous or pre-cancerous growth in a subject, the method comprising:

(i) detecting in a test sample from the subject at least one protein selected from the group consisting of KIAA0152 (SEQ ID NO:1), NAMPT (SEQ ID NO:2), PYCR1 (SEQ ID NO:3), GPX2 (SEQ ID NO:4), PRKDC (SEQ ID NO:5), ALDH18A1 (SEQ ID NO:6), OCIAD2 (SEQ ID NO:7), GCS1 (SEQ ID NO:8), GMDS (SEQ ID NO:9), ARF4 (SEQ ID NO:10), ARF5 (SEQ ID NO:11), LRPPRC (SEQ ID NO:12), CTNNB1 (SEQ ID NO:13), ARF3 (SEQ ID NO:14), GCN1L1 (SEQ ID NO:15), BDH1 (SEQ ID NO:16), RPL9 (SEQ ID NO:17), UGCGL1 (SEQ ID NO:18), FAM3D (SEQ ID NO:19) and CCT4 (SEQ ID NO:20), so as to determine the level of the at least one protein; and

(ii) comparing the level determined in (i) to a reference level of the same at least one protein; wherein the level detected in the test sample is indicative of the stage of the growth.

24. The method according to claim 23, wherein the cancerous or pre-cancerous growth is in gastrointestinal tissue.

25. The method according to claim 24, wherein the growth is selected from the group consisting of pre-cancerous polyps, early stage colorectal cancer and advanced stage colorectal cancer.

26. The method according to claim 24, comprising detecting at least one of KIAA0152 (SEQ ID NO:1) and NAMPT (SEQ ID NO:2).

27. The method according to claim 26, further comprising detecting at least one protein selected from the group consisting of CPT2 (SEQ ID NO:21), ARL1 (SEQ ID NO:22), PFKL (SEQ ID NO:23), GOT2 (SEQ ID NO:24), AP1G1 (SEQ ID NO:25), STRBP (SEQ ID NO:26), CLCA1 (SEQ ID NO:27), CYFIP1 (SEQ ID NO:28), COQ9 (SEQ ID NO:29), NDUFA9 (SEQ ID NO:30), ALDH7A1 (SEQ ID NO:31), HMGCS1 (SEQ ID NO:32), NNT (SEQ ID NO:33), PRDX5 (SEQ ID NO:34), PCCB (SEQ ID NO:35), COPZ1 (SEQ ID NO:36), BAX (SEQ ID NO:37), ACAD9 (SEQ ID NO:38), UBXD8 (SEQ ID NO:39), HMGCS2 (SEQ ID NO:40), SLC25A3 (SEQ ID NO:41), SLC25A11 (SEQ ID NO:42), PDCD6 (SEQ ID NO:43), UCRC (SEQ ID NO:44), DEFA6 (SEQ ID NO:45), DYNC1H1 (SEQ ID NO:46), HK1 (SEQ ID NO:47), CYFIP2 (SEQ ID NO:48), DCI (SEQ ID NO:49) and CISD1 (SEQ ID NO:50).

28. The method according to claim 27, wherein the growth comprises pre-cancerous polyps or early stage colorectal cancer.

29. The method according to claim 24, wherein the growth is obtained from the subject by a procedure selected from the group consisting of biopsy, flexible endoscopy, double balloon endoscopy and surgical colorectal re-sectioning.

30. The method according to claim 24, wherein the gastrointestinal tissue is accessed *in vivo*.

31. The method according to claim 30, comprising contacting gastrointestinal tissue with at least one of a pharmaceutical composition and an endoscopy apparatus.

32. The method according to claim 31, comprising administering the pharmaceutical composition to the subject by a route selected from the group consisting of oral, parenteral, subcutaneous, intramuscular, intrathoracic and intraarticular.

33. The method according to any one of claims 23 or 27, comprising use of at least one reagent suitable for detecting the level of said at least one protein.

34. The method according to claim 31, wherein the pharmaceutical composition or endoscopy apparatus comprise at least one reagent suitable for detecting the level of at said at least one protein.

35. The method according to claim 34, wherein said at least one reagent is suitable for detecting the level of at least one of KIAA0152 (SEQ ID NO:1) and NAMPT (SEQ ID NO:2).

36. The method according to claim 23, wherein said at least one reagent specifically interacts with said at least one protein or with nucleic acid encoding said at least one protein or a fragment thereof, and wherein said reagent is selected from an antibody, an antibody mimetic and a nucleic acid.

37. The method according to claim 33, comprising use of an assay system selected from the group consisting of an immunoassay, a nucleic acid hybridization assay, a binding assay, an array, a phage display library and a combination thereof.

38. The method according to claim 24, wherein the level of the at least one protein in the sample is increased by at least at least 3-fold relative to the reference level, and wherein the reference level is representative of the level of the same protein in non-diseased gastrointestinal tissue.

39. An antigen composition comprising at least one protein selected from the group consisting of KIAA0152 (SEQ ID NO:1), NAMPT (SEQ ID NO:2), PYCR1 (SEQ ID NO:3), GPX2 (SEQ ID NO:4), PRKDC (SEQ ID NO:5), ALDH18A1 (SEQ ID NO:6), OCIAD2 (SEQ ID NO:7), GCS1 (SEQ ID NO:8), GMDS (SEQ ID NO:9), ARF4 (SEQ ID NO:10), ARF5 (SEQ ID NO:11), LRPPRC (SEQ ID NO:12), CTNNB1 (SEQ ID NO:13), ARF3 (SEQ ID NO:14), GCN1L1 (SEQ ID NO:15), BDH1 (SEQ ID NO:16), RPL9 (SEQ ID NO:17), UGCGL1 (SEQ ID NO:18), FAM3D (SEQ ID NO:19), CCT4 (SEQ ID NO:20), CPT2 (SEQ ID NO:21), ARL1 (SEQ ID NO:22), PFKL (SEQ ID NO:23), GOT2 (SEQ ID NO:24), AP1G1 (SEQ ID NO:25), STRBP (SEQ ID NO:26), CLCA1 (SEQ ID

NO:27), CYFIP1 (SEQ ID NO:28), COQ9 (SEQ ID NO:29), NDUFA9 (SEQ ID NO:30), ALDH7A1 (SEQ ID NO:31), HMGCS1 (SEQ ID NO:32), NNT (SEQ ID NO:33), PRDX5 (SEQ ID NO:34), PCCB (SEQ ID NO:35), COPZ1 (SEQ ID NO:36), BAX (SEQ ID NO:37), ACAD9 (SEQ ID NO:38), UBXD8 (SEQ ID NO:39), HMGCS2 (SEQ ID NO:40), SLC25A3 (SEQ ID NO:41), SLC25A11 (SEQ ID NO:42), PDCD6 (SEQ ID NO:43), UCRC (SEQ ID NO:44), DEFA6 (SEQ ID NO:45), DYNC1H1 (SEQ ID NO:46), HK1 (SEQ ID NO:47), CYFIP2 (SEQ ID NO:48), DCI (SEQ ID NO:49), CISD1 (SEQ ID NO:50), LYZ (SEQ ID NO:54), LOC442497;SLC3A2 (SEQ ID NO:61), DMBT1 (SEQ ID NO:84), NUCB1 (SEQ ID NO:.85), GGH (SEQ ID NO:86), AGR3 (SEQ ID NO:87), TM9SF2 (SEQ ID NO:88), SYK (SEQ ID NO:89), GCA (SEQ ID NO:90), HDLBP (SEQ ID NO:91), C1QBP (SEQ ID NO:92) and CLIC1 (SEQ ID NO:93), or an immunogenic fragment of said at least one protein.

40. The composition according to claim 39, wherein the at least one protein is selected from the group consisting of KIAA0152 (SEQ ID NO:1), LYZ (SEQ ID NO:54), LOC442497;SLC3A2 (SEQ ID NO:61), DMBT1 (SEQ ID NO:84), NUCB1 (SEQ ID NO:.85), GGH (SEQ ID NO:86), AGR3 (SEQ ID NO:87), TM9SF2 (SEQ ID NO:88), SYK (SEQ ID NO:89), GCA (SEQ ID NO:90), HDLBP (SEQ ID NO:91), C1QBP (SEQ ID NO:92) and CLIC1 (SEQ ID NO:93), or an immunogenic fragment thereof.

41. The composition according to claim 39, wherein the at least one protein is KIAA0152 (SEQ ID NO:1).

42. A method of preventing or treating cancer, comprising administering to a subject in need thereof, an effective amount of the composition according to any one of claims 39-41.

43. The method according to claim 42, wherein the cancer is colorectal cancer.

44. A kit for detecting, staging or prognosing cancer in a biological sample from a subject, the kit comprising (i) at least one reagent suitable for detecting the level of at least one protein selected from the group consisting of KIAA0152 (SEQ ID NO:1), NAMPT (SEQ ID NO:2), PYCR1 (SEQ ID NO:3), GPX2 (SEQ ID NO:4), PRKDC (SEQ ID NO:5), ALDH18A1 (SEQ ID NO:6), OCIAD2 (SEQ ID NO:7), GCS1 (SEQ ID NO:8), GMDS (SEQ ID NO:9), ARF4 (SEQ ID NO:10),

ARF5 (SEQ ID NO:11), LRPPRC (SEQ ID NO:12), CTNNB1 (SEQ ID NO:13), ARF3 (SEQ ID NO:14), GCN1L1 (SEQ ID NO:15), BDH1 (SEQ ID NO:16), RPL9 (SEQ ID NO:17), UGCGL1 (SEQ ID NO:18), FAM3D (SEQ ID NO:19), CCT4 (SEQ ID NO:20), CPT2 (SEQ ID NO:21), ARL1 (SEQ ID NO:22), PFKL (SEQ ID NO:23), GOT2 (SEQ ID NO:24), AP1G1 (SEQ ID NO:25), STRBP (SEQ ID NO:26), CLCA1 (SEQ ID NO:27), CYFIP1 (SEQ ID NO:28), COQ9 (SEQ ID NO:29), NDUFA9 (SEQ ID NO:30), ALDH7A1 (SEQ ID NO:31), HMGCS1 (SEQ ID NO:32), NNT (SEQ ID NO:33), PRDX5 (SEQ ID NO:34), PCCB (SEQ ID NO:35), COPZ1 (SEQ ID NO:36), BAX (SEQ ID NO:37), ACAD9 (SEQ ID NO:38), UBXD8 (SEQ ID NO:39), HMGCS2 (SEQ ID NO:40), SLC25A3 (SEQ ID NO:41), SLC25A11 (SEQ ID NO:42), PDCD6 (SEQ ID NO:43), UCRC (SEQ ID NO:44), DEFA6 (SEQ ID NO:45), DYNC1H1 (SEQ ID NO:46), HK1 (SEQ ID NO:47), CYFIP2 (SEQ ID NO:48), DCI (SEQ ID NO:49) and CISD1 (SEQ ID NO:50); (ii) a control biological sample for comparison to a biological sample from a subject, and (iii) instructional material comprising a protocol suitable for use in detection and quantification of the at least one protein in a biological sample from a subject.

45. The kit according to claim 44, comprising at least one reagent suitable for detecting the level of at least one protein selected from the group consisting of CPT2 (SEQ ID NO:21), ARL1 (SEQ ID NO:22), PFKL (SEQ ID NO:23), GOT2 (SEQ ID NO:24), AP1G1 (SEQ ID NO:25), STRBP (SEQ ID NO:26), CLCA1 (SEQ ID NO:27), CYFIP1 (SEQ ID NO:28), COQ9 (SEQ ID NO:29), NDUFA9 (SEQ ID NO:30), ALDH7A1 (SEQ ID NO:31), HMGCS1 (SEQ ID NO:32), NNT (SEQ ID NO:33), PRDX5 (SEQ ID NO:34), PCCB (SEQ ID NO:35), COPZ1 (SEQ ID NO:36), BAX (SEQ ID NO:37), ACAD9 (SEQ ID NO:38), UBXD8 (SEQ ID NO:39), HMGCS2 (SEQ ID NO:40), SLC25A3 (SEQ ID NO:41), SLC25A11 (SEQ ID NO:42), PDCD6 (SEQ ID NO:43), UCRC (SEQ ID NO:44), DEFA6 (SEQ ID NO:45), DYNC1H1 (SEQ ID NO:46), HK1 (SEQ ID NO:47), CYFIP2 (SEQ ID NO:48), DCI (SEQ ID NO:49) and CISD1 (SEQ ID NO:50).

46. The kit according to claim 44, comprising a reagent suitable for detecting the level of KIAA0152 (SEQ ID NO:1).

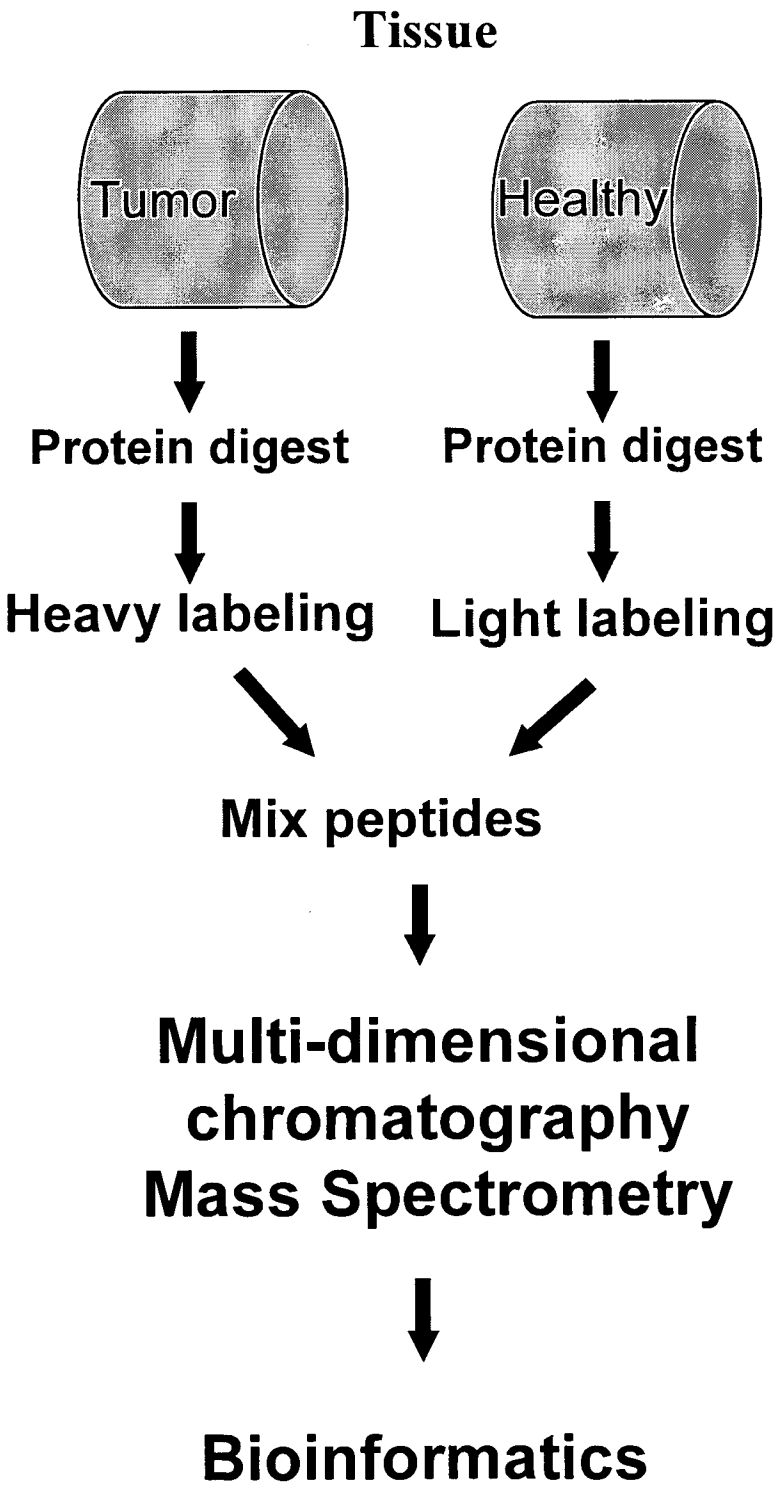
Tissue
Tumor
Healthy
Protein digest
Protein digest
Heavy labeling
Light labeling
Mix peptides
Multi-dimensional chromatography Mass Spectrometry
Bioinformatics

FIGURE 1

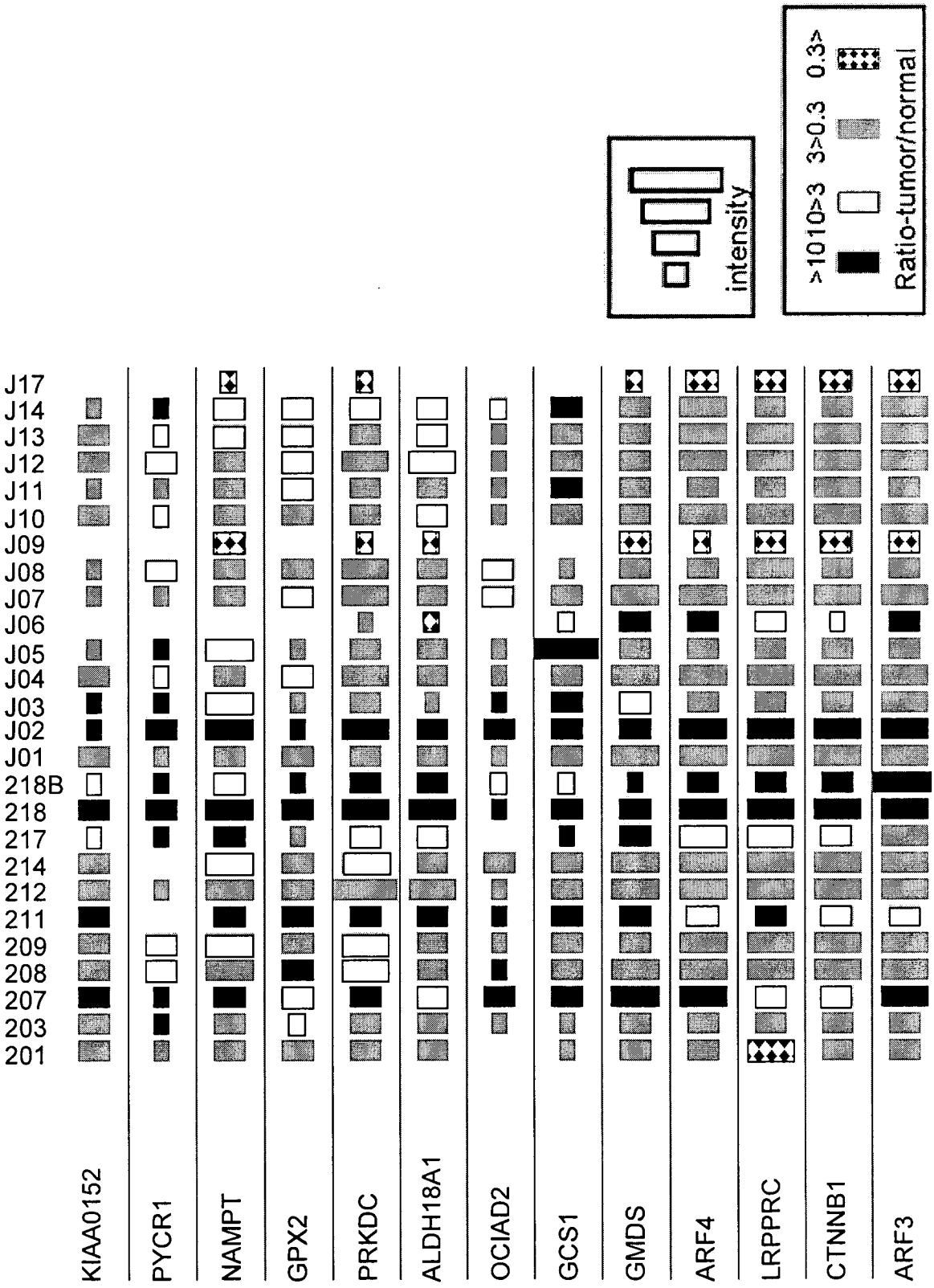
intensity
>10 10>3 3>0.3 0.3>
Ratio-tumor/normal
J17
J14
J13
J12
J11
J10
J09
J08
J07
J06
J05
J04
J03
J02
J01
218B
218
217
214
212
211
209
208
207
203
201
KIAA0152
PYCR1
NAMPT
GPX2
PRKDC
ALDH18A1
OCIAD2
GCS1
GMDS
ARF4
LRPPRC
CTNNB1
ARF3

FIGURE 2

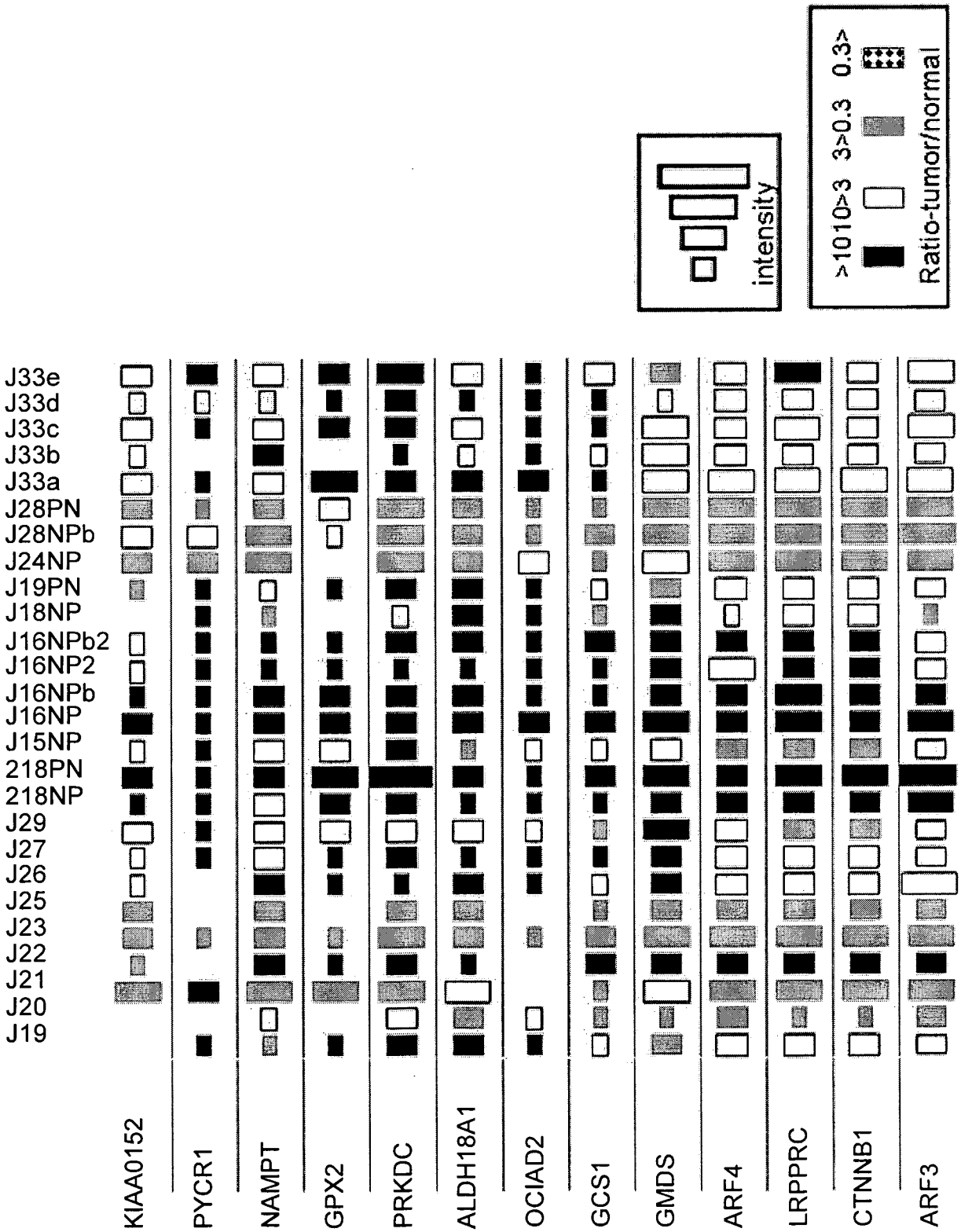

FIGURE 2 CONTINUED