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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
21 June 2012 (21.06.2012)

WIPO | PCT

(10) International Publication Number
WO 2012/079977 A1

(51) International Patent Classification:
G01N 33/574 (2006.01)

(21) International Application Number:
PCT/EP2011/071335

(22) International Filing Date:
30 November 2011 (30.11.2011)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
10194773.7 13 December 2010 (13.12.2010) EP

(71) Applicant *(for all designated States except US)*: **UNIVERSITE DE LIEGE** [BE/BE]; Patent Department, avenue Pré-Aily, 4, B-4031 Angleur (BE).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **CASTRONOVO, Vincent** [BE/BE]; rue Darchis, 17, B-4000 Liege (BE). **TURTOI, Andrei** [DE/BE]; avenue d'Esneux, 64, B-4130 Tilff (BE).

(74) Common Representative: **UNIVERSITE DE LIEGE**; Patent Department, to the attention of N.Antheunis, avenue Pré-Aily, 4, B-4031 Angleur (BE).

(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, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

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

Declarations under Rule 4.17:

— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

Published:

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

[Continued on next page]

(54) **Title**: BIOMARKERS, USES OF BIOMARKERS AND A METHOD OF IDENTIFYING BIOMARKERS

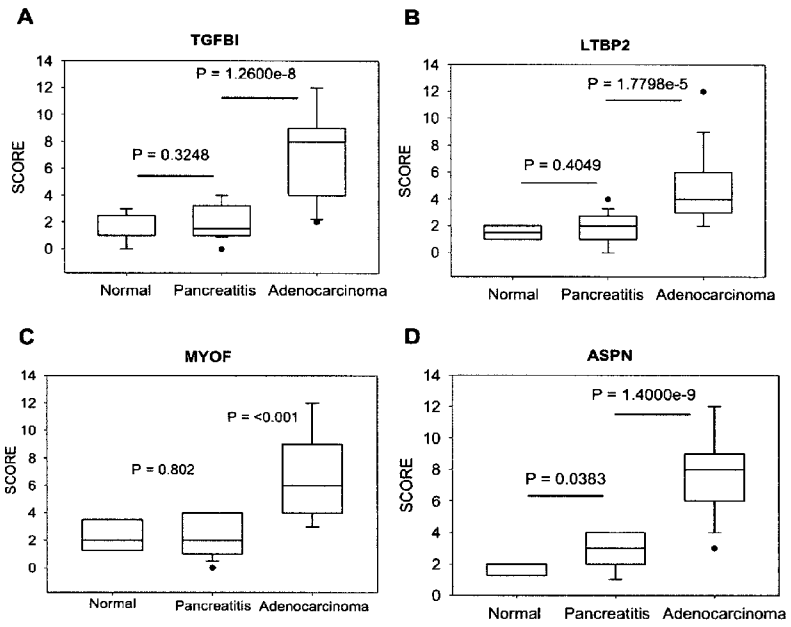


Figure 8

(57) **Abstract**: The invention relates to the use of one or more of protein myoferlin, protein latent-transforming growth factor beta binding protein2, protein transforming growth factor beta induced protein ig-h3, protein asporin,protein tenascin, protein periostin, protein galectin,protein fibronectin, protein prolargin, protein protein-glutamine gamma-glutamyl transferase 2, protein agrin, protein adipocyte enhancer-binding protein 1, protein annexin A6, protein laminin alpha-2 subunit, protein-laminin subunit alpha-4, protein mimecan, protein Ras-related protein Rap-2b,protein collagen alpha-1 (XIV) chain, protein collagen alpha-3 (VI) chain, protein latent-transforming growth factor beta binding protein1, protein V type proton ATPase catalytic subunit A, protein laminin alpha-4 and protein transmembrane protein 62 as a biomarker for pancreatic cancer, or a predisposition thereto. The invention also relates to ligands directed to the above-mentioned biomarkers, for use in the therapeutic and/or prophylactic treatment of a pancreatic cancer.

WO 2012/079977 A1

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

BIOMARKERS, USES OF BIOMARKERS AND A METHOD OF IDENTIFYING BIOMARKERS

This invention relates to novel biomarkers, to an *in vitro* method for identifying accessible biomarkers for specific diseases, in particular cancer, and to uses of the biomarkers.

A major step in many aspects of research related to diseases such as cancer is the identification of specific and sensitive biomarkers suitable for the development of effective and improved diagnostic, prognostic and therapeutic modalities. An aim of the present invention is to provide novel biomarkers for use as novel diagnostic and/or prognostic markers and/or for use in the development of novel therapeutics.

Whilst mass spectrometry, shot gun proteomics and DNA/RNA microarray analyses have resulted in an increasing list of reported potential tumor biomarkers, very few have found their way into the clinical validation phase and even fewer are used as reliable therapeutic targets or diagnostic markers (Clin.Chem. 48 (2002, whole issue) 1145–1375; Clin. Biochem. 37 (2004, whole issue) 503–647; J. Proteome Res. 4 (2005, whole issue) 1043–1456).

It is well-known and generally accepted that there is no correlation between mRNA and protein abundances in the cell, making microarray technologies unreliable for the identification of novel clinically useful biomarkers (Drug Discovery Today 7 (2002) S197-S203, page s200, col 1 and 2; FEBS Letters 583 (2009) 3966-3973; Biotechnol Genet Eng Rev 25 (2008) 77-92).

Hence, Gaspar N et al. (Molecular Pharmacology 72, 152-161 (2007)) describes TGF beta-induced latent-transforming growth factor beta binding protein 2 (LTBP-2) mRNA up regulation in pancreatic cancer cell lines, but provides no proof of a simultaneous up regulation of the LTBP-2 protein.

In the same way, Badea L et al (Hepato-Gastroenterology 55, 2015-2026 (2008, supplementary tables 1 and 3)) identify Asporin and Myoferlin (FER1L3) as overexpressed in pancreatic ductal adenocarcinoma at the mRNA level, without providing evidence of an overexpression of the corresponding proteins.

Finally, Schneider D et al (Biochimica and Biophysica Acta 1588, 1-6 (2002, figure 3) reports transforming growth factor beta induced protein ig-h3 upregulation in pancreatic cancer at the mRNA level, but not at the protein level.

An aim of the present invention is to provide a method for the identification of potential biomarkers which is targeted to the proteins most likely to be useful and reliable for diagnostic and therapeutic purposes. The biomarkers of interest to this invention are accessible biomarkers, that is, biomarkers found in the extracellular matrix or on the outer cell membrane. These proteins are of particular interest because they are reachable by systemically delivered specific agents, such as, antibodies which may be radiolabelled or linked to pharmacological compounds.

A major limitation in the identification of valid biomarkers is the scarcity of the tissues from which they need to be recovered. This is particularly true for human pathological tissues, such as cancer lesions, which are available for analysis only in very small amounts, making these samples very precious. Currently available proteomic methods that enable the analysis of accessible proteins from such limited quantity of tissue leave substantial room for improvement. The techniques available attempt to tackle the problem primarily by exploiting the physical location of the biomarkers of interest (Celis, J.E. et al. Mol. Cell. Proteomics **3**, 327-344 (2004)) and are to the less extent focused on the chemical propriety. To this end, the use of chemically modified biotin that labels accessible proteins through their free amine groups, combined with streptavidin affinity chromatography has been demonstrated to be a powerful method (Castronovo, V. et al. Proteomics **7**, 1188-1196 (2007)). Nevertheless, accessible proteins that do not bear such free amine groups will be omitted from further study. In the method of this invention the fact that most of the outer membrane bound and the extracellular proteins are glycoproteins is exploited.

According to a first aspect the invention provides the use of one or more of myoferlin, CD276, macrophage mannose receptor 2 (CD280), transmembrane 9 superfamily member 3 (EP70-P-iso), EMILIN1, adipocyte enhancer-binding protein 1, agrin, , collagen alpha-1(XII) chain, laminin alpha-5, leucine-rich

repeat-containing protein 15, nectin-like protein 2, olfactomedin-like 1, inositol monophosphatase 3 and transforming growth factor beta induced protein ig-h3 (IHC) as a biomarker for breast cancer, or a predisposition thereto. Preferably the invention provides the use of one or more of adipocyte enhancer-binding protein 1, agrin, laminin alpha-5, transforming growth factor beta induced protein ig-h3 (IHC), collagen alpha-1(XII) chain, macrophage mannose receptor 2 (CD280), and nectin-like protein 2 as a biomarker for breast cancer, or a predisposition thereto. Preferably the invention provides the use of one or more of agrin, laminin alpha-5, transforming growth factor beta induced protein ig-h3 (IHC), collagen alpha-1(XII) chain, macrophage mannose receptor 2 (CD280) and nectin-like protein 2 as a biomarker for breast cancer, or a predisposition thereto. Preferably the protein for use as a biomarker is one or more of agrin, laminin alpha-5, myoferlin and olfactomedin-like 1. Preferably the protein for use as a biomarker is adipocyte enhancer-binding protein 1. Preferably the protein for use as a biomarker is agrin. Preferably the protein for use as a biomarker is CD276. Preferably the protein for use as a biomarker is collagen alpha-1(XII) chain. Preferably the protein for use as a biomarker is laminin alpha-5. Preferably the protein for use as a biomarker is leucine-rich repeat-containing protein 15. Preferably the protein for use as a biomarker is macrophage mannose receptor 2 (CD280). Preferably the protein for use as a biomarker is myoferlin. Preferably the protein for use as a biomarker is nectin-like protein 2. Preferably the protein for use as a biomarker is olfactomedin-like 1. Preferably the protein for use as a biomarker is transmembrane 9 superfamily member 3 (EP70-P-iso). Preferably the protein for use as a biomarker is inositol monophosphatase 3. Preferably the protein for use as a biomarker is transforming growth factor beta induced protein ig-h3 (IHC). Preferably the protein for use as a biomarker is EMILIN1.

Preferably the invention provides the use of adipocyte enhancer-binding protein 1 as a therapeutic biomarker for breast cancer.

According to a second aspect the invention provides the use of one or more of protein myoferlin, protein latent-transforming growth factor beta binding protein 2, protein transforming growth factor beta induced protein ig-h3, protein asporin, protein tenascin, protein periostin, protein galectin, protein fibronectin,

protein prolargin, protein protein-glutamine gamma-glutamyl transferase 2, protein agrin, protein adipocyte enhancer-binding protein 1, protein annexin A6, protein laminin alpha-2 subunit, protein laminin subunit alpha-4, protein mimecan, protein Ras-related protein Rap-2b, protein collagen alpha-1 (XIV) chain, protein collagen alpha-3(VI) chain, protein latent-transforming growth factor beta binding protein 1, protein V type proton ATPase catalytic subunit A, protein laminin alpha-4 and protein transmembrane protein 62 as a biomarker for pancreatic cancer, or a predisposition thereto. Preferably the invention provides the use of one or more protein myoferlin, protein latent-transforming growth factor beta binding protein 2, protein transforming growth factor beta induced protein ig-h3, protein asporin, protein tenascin, protein periostin, protein galectin, protein fibronectin, protein prolargin, protein protein-glutamine gamma-glutamyl transferase 2 and protein agrin as a biomarker for pancreatic cancer, or a predisposition thereto. Preferably the protein for use as a biomarker is protein myoferlin. Preferably the protein for use as a biomarker is protein protein latent-transforming growth factor beta binding protein 2. Preferably the protein for use as a biomarker is protein transforming growth factor beta induced protein ig-h3. Preferably the protein for use as a biomarker is protein aspirin. Preferably the protein for use as a biomarker is protein tenascin. Preferably the protein for use as a biomarker is protein periostin. Preferably the protein for use as a biomarker is protein galectin. Preferably the protein for use as a biomarker is protein fibronectin. Preferably the protein for use as a biomarker is protein prolargin. Preferably the protein for use as a biomarker is protein protein-glutamine gamma-glutamyl transferase 2. Preferably the protein for use as a biomarker is protein agrin.

According to a third aspect the invention provides the use of one or more of CD97, prolargin, Thy-1 membrane glycoprotein, transforming growth factor beta induced protein ig-h3 (IHC), a growth hormone inducible transmembrane protein, biglycan, EPCAM, latent-transforming growth factor beta binding protein 2 and EMILIN2 as a biomarker for colorectal carcinoma (CRC) liver metastasis, or a predisposition thereto. Preferably the invention provides the use of one or more of transforming growth factor beta induced protein ig-h3 (IHC) and CD97 as a biomarker for colorectal carcinoma (CRC) liver metastasis, or a predisposition thereto. Preferably the protein for use as a biomarker is one or more of CD97,

Thy-1 membrane glycoprotein, transforming growth factor beta induced protein ig-h3 (IHC), a growth hormone inducible transmembrane protein, biglycan, EPCAM and EMILIN2. Preferably the protein for use as a biomarker is CD97. Preferably the protein for use as a biomarker is Thy-1 membrane glycoprotein. Preferably the protein for use as a biomarker is transforming growth factor beta induced protein ig-h3 (IHC). Preferably the protein for use as a biomarker is a growth hormone inducible transmembrane protein. Preferably the protein for use as a biomarker is prolargin. Preferably the protein for use as a biomarker is biglycan. Preferably the protein for use as a biomarker is EPCAM. Preferably the protein for use as a biomarker is latent-transforming growth factor beta binding protein 1. Preferably the protein for use as a biomarker is EMILIN2.

Preferably the invention provides the use of one or more of latent-transforming growth factor beta binding protein 2 and prolargin as a therapeutic biomarker for colorectal carcinoma (CRC) liver metastasis.

Reference to CRC liver metastasis relates to cases where the primary tumour is a colorectal carcinoma which gives rise to a liver metastasis.

According to a fourth aspect the invention provides the use of one or more of adipocyte enhancer binding protein 1 (Uniprot accession no:Q8IUX7), agrin (Uniprot accession no:O00468), CD276 (Uniprot accession no:Q5ZPR3), olfactomedin-like 1, myoferlin (Uniprot accession no:Q9NZM1), laminin alpha-5 (Uniprot accession no:O15230), transforming growth factor beta induced protein ig-h3 (Uniprot accession no:Q15582), leucine-rich repeat-containing protein 15 (Uniprot accession no:Q8TF66), transmembrane 9 superfamily member 3 (EP70-P-iso) (Uniprot accession no:Q9HD45), inositol monophosphatase 3 (Uniprot accession no:Q9NX62), EMILIN 1 (Uniprot accession no:Q9Y6C2), collagen alpha-1(XII) chain (Uniprot accession no:Q99715), macrophage mannose receptor 2 (CD280) (Uniprot accession no:Q9UBG0), nectin-like protein 2 (Uniprot accession no:Q9BY67), annexin A6 (Uniprot accession no:P08133), laminin alpha-4 (Uniprot accession no:Q16363), collagen A1 (XIV) (Uniprot accession no:Q05707), collagen A3 (VI) (Uniprot accession no:P12111), latent-transforming growth factor beta binding protein-1 (Uniprot accession

no:Q14766), latent-transforming growth factor beta binding protein-2 (Uniprot accession no:Q14767), asporin (Uniprot accession no:Q9BXN1), laminin alpha-2 (Uniprot accession no:P24043), Ras-related protein Rap-2b (Uniprot accession no:Q9Y3L5), V type proton ATPase catalytic subunit A (Uniprot accession no:P38606), protein-glutamine gamma-glutamyl transferase 2 (Uniprot accession no:P21980), transmembrane protein 62 (Uniprot accession no:Q0P6H9), mimecan (Uniprot accession no:P20774), prolargin (Uniprot accession no:P51888), biglycan (Uniprot accession no:p2181010), EPCAM (Uniprot accession no:P16422), CD97 (Uniprot accession no:P48960), Thy-1 membrane glycoprotein (Uniprot accession no:P04216) and a growth hormone inducible transmembrane protein as biomarker for cancer, or a predisposition thereto. Preferably the invention provides the use of one or more of olfactomedin-like 1, myoferlin, leucine-rich repeat-containing protein 15, transmembrane 9 superfamily member 3 (EP70-P-iso), inositol monophosphatase 3, EMILIN 1, asporin, Ras-related protein Rap-2b, V type proton ATPase catalytic subunit A, transmembrane protein 62, biglycan, EPCAM, growth hormone inducible transmembrane protein and EMILIN 2 as biomarker for cancer, or a predisposition thereto. Preferably the protein for use as a biomarker is adipocyte enhancer binding protein 1. Preferably the protein for use as a biomarker is agrin. Preferably the protein for use as a biomarker is CD276. Preferably the protein for use as a biomarker is olfactomedin-like 1. Preferably the protein for use as a biomarker is myoferlin. Preferably the protein for use as a biomarker is laminin alpha-5. Preferably the protein for use as a biomarker is transforming growth factor beta induced protein ig-h3. Preferably the protein for use as a biomarker is leucine-rich repeat-containing protein 15. Preferably the protein for use as a biomarker is transmembrane 9 superfamily member 3 (EP70-P-iso). Preferably the protein for use as a biomarker is inositol monophosphatase 3. Preferably the protein for use as a biomarker is EMILIN 1. Preferably the protein for use as a biomarker is EMILIN 2. Preferably the protein for use as a biomarker is collagen alpha-1(XII) chain. Preferably the protein for use as a biomarker is macrophage mannose receptor 2 (CD280). Preferably the protein for use as a biomarker is nectin-like protein 2. Preferably the protein for use as a biomarker is annexin A6. Preferably the protein for use as a biomarker is laminin alpha-4. Preferably the protein for use as a biomarker is collagen A1 (XIV). Preferably the protein for use as a

biomarker is collagen A3 (VI). Preferably the protein for use as a biomarker is latent-transforming growth factor beta binding protein-1. Preferably the protein for use as a biomarker is latent-transforming growth factor beta binding protein-2. Preferably the protein for use as a biomarker is asporin. laminin alpha-2. Preferably the protein for use as a biomarker is Ras-related protein Rap-2b. Preferably the protein for use as a biomarker is V type proton ATPase catalytic subunit A. Preferably the protein for use as a biomarker is protein-glutamine gamma-glutamyl transferase 2. Preferably the protein for use as a biomarker is transmembrane protein 62. Preferably the protein for use as a biomarker is mimecan. Preferably the protein for use as a biomarker is prolargin. Preferably the protein for use as a biomarker is biglycan. Preferably the protein for use as a biomarker is EPCAM. Preferably the protein for use as a biomarker is CD97. Preferably the protein for use as a biomarker is Thy-1 membrane glycoprotein. Preferably the protein for use as a biomarker is a growth hormone inducible transmembrane protein.

The proteins and/or peptides referred to in the first, second, third and fourth aspects of the invention are referred to herein as "the biomarker" or "the biomarkers". For example, reference to the biomarkers referred to in the first aspect of the invention is intended to refer to all the proteins recited in the first aspect of the invention.

According to yet another aspect, the invention provides a method for determining the breast cancer status of a subject comprising the steps of:

(a) providing a sample of material from a subject;

(b) determining the level in the sample of one or more of the biomarkers referred to in the first aspect of the invention; and

(c) comparing the level determined in (b) with one or more reference values from the same or a different subject.

According to yet another aspect, the invention provides a method for determining the pancreatic cancer status of a subject comprising the steps of:

(a) providing a sample of material from a subject;

(b) determining the level in the sample of one or more of the biomarkers referred to in the second aspect of the invention; and

(c) comparing the level determined in (b) with one or more reference values from the same or a different subject.

According to yet another aspect, the invention provides a method for determining the CRC liver metastasis status of a subject comprising the steps of:

(a) providing a sample of material from a subject;

(b) determining the level in the sample of one or more of the biomarkers referred to in the third aspect of the invention; and

(c) comparing the level determined in (b) with one or more reference values from the same or a different subject.

According to yet another aspect, the invention provides a method for determining the cancer status of a subject comprising the steps of:

(a) providing a sample of material from a subject;

(b) determining the level in the sample of one or more of the biomarkers referred to in the fourth aspect of the invention; and

(c) comparing the level determined in (b) with one or more reference values from the same or a different subject.

The samples referred to in step (a) of the preceding aspects of the invention may be obtained from a subject. Preferably the step of obtaining the sample does not form part of the method of the invention.

The method of determining a cancer status or a metastasis status of the invention may be used together with an assessment of clinical symptoms.

The phrase “cancer status" or “metastasis status" includes any distinguishable manifestation of the specific cancer or metastasis to which the biomarkers refer, for example, breast cancer, pancreatic cancer or CRC liver metastasis. For example, cancer status includes, without limitation, the presence or absence of

cancer, the risk of developing cancer, the stage of the cancer, the progression of the cancer (e.g. progress of cancer or remission of cancer over time) and the effectiveness or response of a subject to treatment for a cancer.

The method of the invention may be used, for example, for any one or more of the following: to diagnose cancer in a subject; to assess the chance of a subject developing cancer; to advise on the prognosis for a subject with cancer; to monitor disease progression; and to monitor effectiveness or response of a subject to a treatment.

The biomarkers of the invention may be used as one or more of diagnostic biomarkers, prognostic biomarkers and therapeutic biomarkers. In a preferred embodiment, biomarkers which are not explicitly stated to be therapeutic biomarkers are preferably at least one of diagnostic or prognostic biomarkers, or both.

Preferably the method allows the diagnosis of cancer in a subject from the analysis of the level of the one or more biomarkers in a sample provided by/or obtained from the subject.

The method of the invention may also or alternatively be used to develop therapies targeted to the biomarkers. Reference herein to therapeutic biomarkers is intended to refer to biomarkers which can be used as targets for therapy, this may be in addition to or an alternative to the biomarkers being used for diagnostic and/or prognostic purposes. Preferably a therapeutic biomarker represents a protein which can be targeted by a ligand to deliver a drug, for example, the therapeutic biomarker may be recognised by an antibody-drug conjugate wherein the antibody recognises the therapeutic biomarker and delivers the conjugated drug.

The sample material obtained from the subject may comprise whole blood, blood serum, blood plasma, urine, mucous or tissue. Preferably, the sample is a sample of tissues, preferably a sample of tissues suspected to be cancerous, for example a sample of breast tissue, pancreatic tissue or liver tissue, such as a biopsy sample.

Preferably determination of the level of a biomarker in a sample comprises the detection of a polypeptide with at least 65% sequence identity, more preferably at least 70%, 75%, 80%, 85%, 90%, 95% or greater sequence identity, to the published sequence of the biomarker.

Proteins are biochemical compounds comprising one or more polypeptides. A polypeptide is a single linear polymer chain of amino acids bonded together by peptide bonds between a carboxyl and an amino groups of adjacent amino acid residues. Proteins frequently exist in the body, and in samples derived there from, in a plurality of different forms. These forms can result from either or both of pre- and post-translational modification. When detecting or determining the level of a protein in a sample, the ability to differentiate between different forms of a protein depends upon the nature of the difference and the method used to detect or measure the protein level. For example, an immunoassay using a monoclonal antibody will detect all forms of a protein containing the epitope to which the antibody is raised and will not distinguish between different forms. However, a sandwich immunoassay that uses two antibodies directed against different epitopes on a protein will distinguish between forms of the protein that contain both epitopes and those that contain only one of the epitopes.

In a method of the invention the assay method used to determine the level of the biomarker protein preferably detects all forms of the specific biomarker protein. Preferably at least all biologically active forms of the biomarker protein are detected. Preferably, all forms of the biomarker protein with at least 75% or more, preferably at least 80%, 85%, 90% or 95% or more, sequence identity with the published amino acid sequence of the biomarker will be detected in the method of the invention.

Methods of measuring polypeptide/protein/nucleic acid identity are well known in the art. For example the UWGCG Package provides the BESTFIT program which can be used to calculate identity (e.g. used on its default settings) (Devereux et al (1984) Nucleic Acids Research 12, p387-395). The PILEUP and BLAST algorithms can also be used to calculate homology or line up sequences (typically

on their default settings), for example as described in Altschul S.F. (1993) J Mol Evol 36:290-300; Altschul, S, F et al (1990) J Mol Biol 215:403-10. Software for performing BLAST analysis is publicly available through the National Centre for Biotechnology Information (http://www.ncbi.nlm.nih.gov/).

The level of the biomarker present in the sample may be determined by any suitable assay which may comprise the use of any of the group comprising enzyme assays, immunoassays, spectrometry, mass spectrometry, Matrix Assisted Laser Desorption/Ionization Time-of-Flight (MALDI-TOF) Mass Spectrometry, microscopy, northern blot, western blot, Southern blot, isoelectric focusing, SDS-PAGE, PCR, quantitative RT-PCR, gel electrophoresis, protein microarray, DNA microarray, and antibody microarray, or combinations thereof.

If antibodies are used one or more antibodies may be synthetic, monoclonal, polyclonal, bispecific, chimeric or humanised. A chimeric antibody includes portions derived from different animals. Humanised antibodies are antibodies from non-human species having one or more complementary determining regions from the non-human species and a framework region from a human immunoglobulin molecule. Chimeric and humanised antibodies can be produced by recombinant techniques well known in the art.

The one or more antibodies may comprise a tag or a label selected from the group comprising a radioactive, a fluorescent, a chemiluminescent, a dye, an enzyme, or a histidine tag or label, or any other suitable label or tag known in the art.

Preferably the reference value, to which the determined levels of the biomarker are compared, is the level of the same protein observed in one or more normal samples. A normal sample preferably refers to a sample of tissue taken from a normal/non diseased location corresponding to the potentially pathologic/diseased tissue being tested. The corresponding normal/non-diseased tissue may be taken from the same or a different individual to the suspected pathologic/diseased tissue. When it is taken from the same individual it could for example be taken during a medical treatment such as drug treatment, radiation treatment or surgical treatment. Alternatively, the normal sample may be a sample of normal/non-

diseased tissue which is a different tissue to the suspected pathologic/diseased tissue, again this may be taken from the same or a different individual. For example, for a patient with suspected breast cancer, a sample may be taken of the suspected cancer and the normal sample may be taken from non-diseased breast tissue from the same or a different individual.

Alternatively, the reference value may be a previous value obtained for a specific subject. This kind of reference value may be used if the method is to be used to monitor progression of a cancer or to monitor the response of a subject to a particular treatment.

When the determined level of the biomarker is compared with a reference value, an increase or a decrease in the level of the biomarker may be indicative of the cancer status of the subject.

More specifically an increase in the level of the biomarker may be indicative, or diagnostic, of cancer.

Alternatively, a decrease in the level of the biomarker compared to a previously measured level in the same individual may be indicative that a particular therapy has been, or is being, effective.

The method of the invention may also be used to monitor cancer progression and/or to monitor the efficacy of treatments administered to a subject. This may be achieved by analysing samples taken from a subject at various time points following initial diagnosis, and monitoring the changes in the levels of the biomarker, and comparing these levels to normal and/or reference values. In this case reference levels may include the initial levels of the biomarker in the subject, or the level of the biomarker in the subject when they were last tested, or both.

Preferably the method of the invention is carried out *in vitro*.

The subject may be a mammal, and is preferably a human, but may alternatively be a monkey, ape, cat, dog, cow, horse, rabbit or rodent.

According to a yet further aspect the invention provides a method for detecting the presence or absence of a beneficial response in a patient after administration of a therapy, comprising:

(a) providing a biological sample from a patient;

(b) measuring in the sample the level of expression of a biomarker;

(c) comparing levels in (b) to a control value for levels of the biomarker;

(d) determining whether or not the difference in levels between the sample and control reflects a beneficial response in the patient

wherein the biomarker is selected from those identified in the first, second, third or fourth aspect of the invention.

In this aspect of the invention the patient may suffer from one or more of breast cancer, pancreatic cancer, CRC liver metastasis or any other cancer.

This method of the invention may allow the determination of whether or not a particular therapeutic agent is efficacious for the treatment of a particular condition.

Preferably, the control value is the level of a particular biomarker in the patient prior to the administration of a particular therapy, or during the administration period. Preferably, if the level in (b) is lower than the control value this is indicative that the therapy administered is efficacious.

In all aspects of the invention the step of providing a sample preferably does not include the step of obtaining/taking the sample from an individual/patient/subject.

According to another aspect of the invention, there is provided the use of a ligand, preferably a high affinity ligand, directed to a biomarker according to the invention, for the manufacture of a medicament for the therapeutic and/or prophylactic treatment of a human or animal disease. Preferably the medicament

is based on the use of specific ligands, preferably antibodies or antibody-drug conjugates.

In the context of the present invention, the term ligand relates to antibodies, antibody fragments, drugs, prodrugs, ligands, biotin, and derivatives and conjugates thereof, preferably conjugates of antibodies or antibody fragments with drugs or prodrugs. Preferably, the ligand is directed to an extracellular domain of the biomarker. By extracellular domain of the biomarker, one means the domain of the biomarker situated outside the cancer cell. For example, the extracellular domain of the biomarker may represent a portion or the totality of the biomarker. By portion one means 1%, 2%, 3%, 5%, 7%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98% or 99% or a fraction thereof of the totality of the biomarker.

According to a yet further aspect the invention provides a ligand, preferably a high affinity ligand, directed to a biomarker according to the invention, for use in the therapeutic and/or prophylactic treatment of a human or animal disease. Preferably the ligand is an antibody or antibody-drug conjugate.

According to a still further aspect of the invention, there is provided a method of therapeutic and/or prophylactic treatment of a human or animal disease comprising administering a therapeutically or prophylactically effective amount of a ligand, preferably a high affinity ligand, directed to a biomarker according to the invention. Preferably the ligand is an antibody or an antibody-drug conjugate.

According to another aspect of the invention, there is provided a kit for use in determining the cancer status or metastasis status of a subject, wherein the kit comprises at least one agent for determining the level of one or more biomarkers according to the invention in a sample provided by the subject.

The agent may be an enzyme, a nucleic acid, a protein probe, a metabolite, a ligand such as an antibody, or any other suitable composition.

The kit may comprise one or more capture agents for capturing the biomarker, the level of which is to be determined. The capture agent may be one or more antibodies.

The kit may comprise instructions for suitable operational parameters in the form of a label or separate insert. The instructions may inform a consumer about how to collect the sample, and/or how to wash the capture agent.

According to a yet further aspect, the invention provides the use of the determination of the levels of one or more biomarkers according to the invention as a means of assessing the cancer status in an individual.

According to another aspect the invention provides a method of treating cancer in a subject comprising administering to the subject an agent capable of modulating the level of a biomarker according to the invention in the subject.

The agent may act at the transcriptional, translation and/or post translational level.

According to a further aspect the invention provides a method of identifying compounds for treating cancer comprising screening for one or more compounds that modulate the level of one or more biomarkers according to the invention.

According to another aspect of the invention, there is provided an *in vitro* method for the analysis of accessible proteins in a sample comprising:
providing a sample;
labelling accessible proteins in the sample with a labelling reagent;
isolating and recovering proteins labelled with the labelling reagent from the sample;
digesting the remaining, non-labelled, proteins in the sample;
isolating and recovering glycopeptides from the digested sample;
recovering non-glycosylated peptides that were not recovered in the isolation of labelled proteins or glycopeptides;
analysing the recovered proteins/peptides.

Surprisingly the combination of steps in the method of the invention significantly increases the potential to identify accessible proteins from a sample, compared to performing any of the steps alone.

The method of the invention provides an improved method for the identification of accessible biomarkers from pathologic/diseased tissue samples, such a tissue biopsies. The accessible biomarkers/proteins may be membrane bound proteins or proteins secreted into the extracellular space/matrix.

Preferably the accessible proteins can be reached by systemically delivered agents, for example, systemically delivered antibodies. Preferably this is achieved because the accessible proteins are more exposed to the interstitial fluid than other proteins.

The method of the invention preferably allows the number of biomarkers identified in a precious sample, such as a biopsy, to be maximised whilst still only considering accessible proteins.

By analysing the accessible proteins the method allows biomarkers to be identified which can be used for a number of purposes, including one or more of the following, disease diagnosis, disease prognosis, monitoring disease progression, monitoring the efficacy of a treatment, and for targeting potential therapies to cells or tissues expressing the biomarker. Examples of potential therapies include antibodies, antibody fragments, drugs, prodrugs or other ligands which may be targeted to a specific biomarker.

Preferably the method of the invention may be used to screen for/identify biomarkers for specific diseases wherein the biomarkers are accessible from the extracellular space in a tissue or cell sample.

The method of the invention preferably allows a minimal amount of material to be used, making it suitable for the quantitative analysis of scarce biopsy tissue. The method may allow the number of identified valuable biomarkers to be maximized from a precious sample, whilst preserving the quality of these proteins

and thus allowing potentially accessible and hence relevant biomarkers to be identified.

The labelling reagent may be any reagent capable of labelling a protein via a characteristic functional group, wherein the labelling reagent then allows the enrichment of labelled proteins. An example of a suitable labelling reagent is a reactive biotin, preferably a biotin reactive ester derivative, which can be used to label primary amines on proteins/peptides in a sample. In the method of the invention the sample is preferably a tissue sample in which the cells are substantially intact and have substantially intact membranes, such that when a labelling reagent is added it cannot penetrate the cells and thus substantially only extracellular or membrane bound proteins are labelled.

If biotin is used as the labelling reagent the high affinity of biotin for avidin may then be exploited to isolate and recover biotin labelled proteins. Other suitable reagents for labelling accessible proteins include a reagent which binds to sugar moieties on proteins or peptides and magnetic beads attached to a functional group that will react with sugar moieties or primary amines.

The method of the invention further includes enriching for accessible proteins by recovering glycosylated peptides. Glycosylated peptides are often membrane associated or secreted, and hence are accessible.

Glycosylated is used herein to refer to proteins or peptides which have an attached saccharide group. The terms glycoprotein/glycopeptide are used interchangeable herein with glycosylated protein/glycosylated peptide, and are intended to have the same meaning.

The sample may be a cell or tissue sample. Preferably the sample is a tissue sample. The tissue sample may be a pathologic tissue sample, that is, from a diseased tissue. Alternatively the tissue sample may be a sample of normal, non diseased, tissue. Where the tissue sample is a pathologic tissue sample the tissue may be selected from tumour tissue, inflamed tissue, atheromatatic tissue, and

tissue resulting from a degenerative, metabolic or genetic disease. Preferably the sample is a tissue biopsy.

Reference herein to a normal/non-diseased tissue refers to either normal/non-diseased tissue corresponding to the pathologic/diseased tissue from the same individual or from a different individual. The two tissue samples, normal and pathologic, may be taken from the same location, for example both liver or both breast, or from different locations, in the same or different individuals. For example, the pathologic tissue may be breast tissue and the normal tissue may be a sample of non-diseased breast tissue from the same individual, or a sample of non-diseased breast tissue from a different individual, or a sample of a different tissue from the same or a different individual. Preferably the normal/non-diseased tissue is a sample of the same type of tissue as the pathologic/diseased tissue.

The term "biomarker" is used herein with reference to the method of the invention to refer to any protein or peptide (which may be modified, for example glycosylated) whose level of expression in pathologic tissue allows the tissue to be distinguished from a normal tissue.

Preferably proteins in the sample are labelled with a labelling reagent by immersing the sample in a solution comprising the labelling reagent. Preferably the immersion allows accessible proteins to be labelled with the labelling reagent. The labelling reagent, such as a biotin derivative, is preferably able to covalently link onto accessible primary amines in a protein. Thus only proteins with accessible primary amines will be labelled. Preferably the sample is immersed in a solution of reactive biotin ester derivatives in order to label the accessible proteins.

The sample immersed in the labelling reagent containing solution is preferably a native sample wherein only accessible proteins are labelled, and non-accessible proteins are not, or are essentially not, labelled.

A native sample means that the sample has not been denatured or fixed prior to immersion in the labelling reagent containing solution. Preferably upon

immersion of the sample in the labelling reagent containing solution many of the accessible proteins become labelled with the labelling reagent. However, as not all proteins are able to be labelled, further protein/peptide recovery steps are included in the method of the invention to maximise the biomarkers identified.

After the labelling step the sample is preferably subjected to a solubilisation step wherein only the soluble fraction is retained and used in the rest of the method. Preferably the insoluble fraction is discarded.

If the proteins have been labelled with biotin, or a biotin derivative, the labelled proteins may be recovered using streptavidin, preferably using streptavidin beads. Biotinylated proteins/peptides are easily purified due to the highly specific interaction between biotin and streptavidin. This interaction is strong even in the presence of lysis buffers containing strong detergents, thus minimising non-specific binding during purification. The biotinylated proteins may be eluted from the streptavidin beads using DTT. The eluted protein may then be digested, for example with trypsin, to provide peptide fragments for analysis by mass spectrometry.

Preferably the non-labelled proteins are digested using any proteinase (like trypsin) to produce shorter peptides before the recovery of glycosylated peptides.

Glycoproteins/glycopeptides, preferably glycopeptides, within the remaining proteins/peptide may be isolated by first oxidising sugar moieties on the proteins/peptides. The sugar moieties may be oxidised by using a periodate solution, and then isolating any proteins/peptides with oxidised sugar moieties, for example by using a hydrazide resin. The bound proteins/peptides may be released from the hydrazide using PNGase F. Other methods to isolate glycoproteins/glycopeptides are known and may be used in the method of the invention.

The remaining non-labelled and non-glycosylated proteins/peptides, preferably peptides, are then recovered and are also analysed, for example, by mass spectrometry.

Preferably the remaining non-labelled and non-glycosylated peptides are recovered in the flow through fraction from the method used to recover glycosylated proteins/peptides.

Preferably the recovered proteins/peptides are analysed to determine their sequence and hence the identity of the protein from which they are derived. The recovered proteins/peptides may be analysed by mass spectrometry. Preferably mass spectrometry analysis provides sequence information relating to the proteins/peptides which can then be compared to protein/peptide sequence databases to allow the protein/peptide to be identified.

Preferably if mass spectrometry is used it comprises sequencing by tandem mass spectrometry.

Preferably the analysis of the non-glycosylated peptides recovered allows the identification of glycosylated peptides (recovered in the previous step of the method of the invention) which cannot be confidently assigned to a particular protein based on the analysis of the glycosylated peptide alone. More specifically, the identification of non-glysoylated peptides which are found in a glycosylated protein may allow an isolated glycosylated peptide to be more confidently assigned to a particular glycosylated protein.

Analysis of the non-glycosylated peptides may also allow non-glycosylated, but accessible, proteins to be identified.

Preferably the glycosylated and the non-glycosylated peptides are all analysed by mass spectrometry.

The method of the invention may include the further step of comparing the recovered proteins/peptides from a pathologic tissue sample to the recovered proteins/peptides from a normal tissue sample.

Preferably by comparing the differential expression of recovered proteins/peptides from a pathologic tissue sample with that of a normal sample, potential biomarkers for the disease of the pathologic tissue sample may be identified.

Preferably, by identifying recovered proteins/peptides which have a higher expression in the pathologic tissue sample compared to the normal tissue sample, one or more biomarkers for a specific disease may be identified. The identified biomarkers, due to the nature of the method of the invention, are preferably accessible *in vitro* and *in vivo* from the extracellular space.

The method of the invention may be applied to any tissue and any condition/disease, and because it considers accessible proteins it may, in addition to identifying biomarkers, allow custom therapies to be developed targeted to highly accessible targets. For example, antibody based treatments may be targeted to the biomarker.

In a preferred embodiment of the invention the labelling and recovery steps are performed using a pathologic tissue sample and a normal tissue sample and the differential expression of recovered proteins/peptides are compared. More preferably the method comprises the steps of:

providing a sample of normal tissue and a sample of pathologic tissue;

labelling the accessible proteins in each sample separately with a labelling reagent;

isolating and recovering labelled proteins from each sample;

digesting the remaining, non-labelled, proteins in each sample;

isolating and recovering glycopeptides from each digested sample;

recovering non-glycosylated peptides from each sample that were not recovered in the isolation of labelled proteins or glycopeptides;

analysing all the recovered proteins and peptides;

determining the differential expression pattern of the recovered proteins/peptides in the pathologic sample compared to the normal sample;

identifying proteins/peptides that have a higher expression in the pathologic tissue sample compared to the normal tissue sample, or which are more frequently

expressed in the pathologic tissue sample compared to the normal tissue sample. These proteins/peptides may be used as biomarkers for the pathologic tissue, preferably they are accessible for high affinity ligands from the extracellular space.

Preferably the biomarker is expressed in the pathologic tissue and not the normal tissue.

The purified labelled proteins may be cleaved after purification to smaller peptides before further analysis. The cleavage is preferably by proteolytic digestion, this may be achieved using trypsin.

According to a further aspect the invention provides a biomarker identified by the method of the invention.

According to another aspect of the invention, there is provided the use of a ligand, preferably a high affinity ligand, directed to a biomarker identified by the method of invention, for the manufacture of a medicament for the therapeutic and/or prophylactic treatment of a human or animal disease. Preferably the medicament is based on the use of specific ligands, preferably antibodies or antibody-drug conjugates.

According to a yet further aspect the invention provides a ligand, preferably a high affinity ligand, directed to a biomarker identified by the method of invention, for use in the therapeutic and/or prophylactic treatment of a human or animal disease. Preferably the medicament is based on the use of specific ligands, preferably antibodies or antibody-drug conjugates.

According to a still further aspect of the invention, there is provided a method of therapeutic and/or prophylactic treatment of a human or animal disease comprising administering a therapeutically or prophylactically effective amount of a ligand, preferably a high affinity ligand, directed to a biomarker identified by the method of invention. Preferably the ligand is an antibody or an antibody-drug conjugate.

It will be appreciated that optional features applicable to one aspect or embodiment of the invention can be used in any combination, and in any number. Moreover, they can also be used with any of the other aspects or embodiments of the invention in any combination and in any number. This includes, but is not limited to, the dependent claims from any claim being used as dependent claims for any other claim in the claims of this application.

Embodiments of the present invention will now be described herein, by way of example only, with reference to the following figures.

Figure 1 – illustrates schematically the sequential extraction of accessible proteins from tissue samples. The first step consists of biotinylation of accessible proteins and their isolation (B fraction). The second step utilizes protein digestion of non biotinylated proteins into peptides, followed by the isolation of glycopeptides (G fraction). The final step also collects the non-glycosylated peptides (R) using them to complement the sequence information of non-assigned glycopeptides (detailed in Figure 2). Additionally, the R fraction contains a significant number of accessible proteins which supplement the already identified proteins in the B and G fractions. Three internal standards (IS1, 2 and 3) were added at different manipulation steps in the method in order to monitor, recovery, reproducibility and accuracy of quantification. The composition of the respective internal standard is outlined later. All the fractions were analyzed using the 2D-nanoUPLC-MS^e system, which consisted essentially of two C18-phases run at pH 10 and pH 3 respectively.

Figures 2A and 2B – illustrate the quantitative accuracy, recovery and reproducibility of the technique described in Figure 1. Figure 2A illustrates the absolute quantitative evaluation of the internal standards spiked in the sample during the preparation process using the sequential method (see Figure 1). Protein quantity was calculated using the PLGS software, based on a previously calculated response factor. The error indicates the standard deviation of means, based on three full process

technical replicates. Figure 2B illustrates the absolute quantification of the internal standards spiked in the sample during the glycopeptides analysis alone. For comparative reasons the isolation of glycopeptides was also conducted as a standalone technique. Quantification and error bars are same as in Figure 2A.

Figures 3A and 3B – illustrate the analysis of the effective value of the combinatory method of Figure 1A with respect to the individual components alone (biotin, glycopeptides and rest) and the absolute numbers of accessible proteins identified (accessible proteins: extracellular, secreted and/or membrane). **Figure 3A** illustrates the number of accessible proteins identified in each of the steps of the combinatory method as well as the sum of all the components together. Three full process technical replicates are displayed. Notably, G indicates the average number of proteins obtained using only the data from the trapped glycoprotein fraction. However, using the in-silico method, the overall number of glycoproteins was increased; this resulted in the data marked GR. Importantly, the specific dynamic range of proteins characterizing the R (rest) fraction allowed for a significant identification of additional accessible proteins. **Figure 3B** is the same as Figure 3A conducted only for the glycopeptides isolation procedure as standalone technique.

Figure 4 – illustrates the comparisons of the identification overlap of accessible proteins in different fractions of the combinatory method of Figure 1 and glycopeptides analysis alone. The Venn diagrams display three full process replicates. Regarding the combinatory method, Figure 4 shows that there is substantial overlap among all the separate fractions. However, each fraction also contains a significant number of unique proteins. Owing to the technical process the replicates demonstrate that there is significant added value in the individual steps of the combinatory method and that this is reproducible and beyond the measured variability of the analysis. The lower part of the figure shows a comparison between the combinatory method (BGR) and glycoprotein analysis alone (GR&R).

The diagrams display only accessible proteins. For correct comparison accessible proteins found in the rest fractions are always included. It is worth noting that glycoprotein analysis as a standalone technique identifies on average 18 % of unique proteins. In contrast to this, over 50 % of the proteins identified in the combinatory method are unique.

Figure 5 – illustrates immunohistochemical validation that the proteins laminin α-5 and myoferlin are overexpressed in breast cancer tissue. Figure 5 shows representative cases from 25 tumoral and 11-13 normal individuals. Both laminin α-5 and myoferlin exhibit strong positive staining in the tumor tissue compared to the normal tissue (original magnification X400). The values in the graphs are reported as scores (NB= normal breast; DC= ductal carcinoma). The error bars indicate standard error of the mean. P-values ≤ 0.05 were considered as significant.

Figure 6A - is a list of potential breast tumor biomarkers obtained from the analysis of 3 non tumoral adjacent and 3 tumoral specimens. The selection of the potential biomarkers has been conducted with respect of the proteins presence in tumoral and absence or diminished presence in non-tumoral tissue samples. For certain proteins that have shown presence in both the tumoral and the adjacent normal tissue, semi-quantitative data (emPAI ratio) are included when these were significantly overexpressed in tumor tissue (ratio greater than 2.0). All the proteins are potentially accessible as they are located on the outer side of the cell membrane (secreted, extracellular or membrane).

Following abbreviations were used:

Subcellular location: S: secreted; E: extracellular; M: membrane; C: cell membrane; PlM: plasma membrane; SP: single pass membrane protein; MP: multipass membrane protein; LA: lipid anchor; MM: mitochondrion matrix; PM: peripheral membrane protein; Cy: cytoplasm; NM: nuclear membrane; U: cell component unknown.

Biological process: C: cell communication, ST: signal transduction; I: immune response; M: metabolism; E: energy pathway; T: transport; A:

apoptosis; CG: Cell growth and/or maintenance; U: biological process unknown.

Fraction: G - glycopeptide; B - biotinylated; N - non glycopeptides;

Figures 6B-D - list of potentially accessible modulated proteins obtained from the analysis of 5 non-tumoral adjacent and 5 tumoral individual matched specimens; Figure 6B represents proteins isoloated in the biotin fraction, Figure 6C represents proteins isoloated in the rest fracton and Figure 6D represents proteins isoloated in the glyco-fraction. The proteins were selected with respect to their presence in tumoral tissue samples and their absence or reduced presence in non-tumoral tissue samples. For certain proteins present in both the tumoral and the adjacent normal tissue, quantitative data (relative ratio of expression) are included if they were significantly overexpressed in the tumor (ratio $\geq$ 1.5). All the proteins are accessible as they are located on the outer side of the cell membrane (secreted, extracellular or membrane). Following abbreviations were used: Subcellular location: S: secreted; E: extracellular; M: membrane; Cy: cytoplasm; CJ: cell junction; CP: cell projection; CSu: cell surface; N: nucleus; G: Golgi; ER: endoplasmic reticulum; Mel: melanosome; SR: sarcoplasmic reticulum; L: lysosome;

Figure 7 – illustrates the accessible differentially expressed proteins identified as overexpressed in pancreas tumor (n = 3) compared to normal tissue (n = 3). A protein is considered overexpressed when: i) it was detected only in the tumoral condition, ii) found in both conditions but more often identified in tumoral samples than normal samples, iii) found with a fold change ratio of the emPAI tumor/normal > 2. The following proteins were identified as multiple isoforms/subunits: [1] Calpain-1 and Calpain-2 catalytic subunit; [2] Guanine nucleotide-binding protein G (i) subunit alpha-2, (olf) subunit alpha, (k) subunit alpha, (o) subunit alpha, (q) subunit alpha, (t) subunit alpha-1and (t) subunit alpha-2; [3] Collagen alpha-2(IV) chain and alpha-1(XIV) chain; [4] Fibulin-1 and -2; [5] Galectin-3 and -4; [6] HLA class II histocompatibility antigen DQ alpha 2 chain and

DR alpha chain; [7] Thrombospondin-2 and -1; [8] Laminin subunit beta-2 and alpha-1; [9] Tenascin and isoform X.

Potential subcellular location: C = Cell membrane. CS = Cell Surface. E = Extracellular. N = Nucleus. Cy = Cytoplasm. S = Secreted. ER = Endoplasmic Reticulum. G = Golgi; Biological process: U = Unknown. R = Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism. Cc; St = Cell communication, Signal transduction. GM = Cell growth and/or maintenance. P = Cell proliferation. I = Immune response. M; Ep = Metabolism, Energy pathways. Pm = Protein metabolism. T = Transport;

Figure 8 – illustrates immunohistochemical evaluation (Box-Plot) of the identified modulated proteins in pancreas non-cancerous tissue (normal pancreas and pancreatitis) and adenocarcinoma. **(A)** TGFBI: compared to pancreatitis, the positivity in the tumor was significant ($P \leq 0.0001$); the expression level of TGFBI in normal and pancreatitis tissues was not significantly different. **(B)** LTBP2: statistical testing indicated significant difference between adenocarcinoma and normal/ pancreatitis tissues ($P \leq 0.0001$); LTBP2 expression level in the normal tissue was not significantly different compared to pancreatitis. **(C)** MYOF: statistical testing indicated significant difference between adenocarcinoma and both inflammatory and normal pancreas tissues ($P \leq 0.002$); the expression level of MYOF in normal and pancreatitis tissues was not significantly different. **(D)** For ASPN, significant positivity was observed in the tumoral with respect to pancreatitis tissue ($P \leq 0.0001$); pancreatitis was only somewhat positive in comparison to the normal pancreas, however the difference was statistically significant ($P \leq 0.05$). Statistical evaluation was conducted using Wilcoxon rank sum test.

Figure 9 – illustrates immunohistochemistry performed on the normal tissue TMA using antibodies against ASPN, TGFBI and LTBP2 targets. Here displayed are examples of the most important tissues scored. Brain, intestine, endometrium and prostate tissues were scored negative for all the antigens investigated. TGFBI expression was observed in liver,

however was of weak intensity. Kidney cortex and esophagus resulted in a weak expression of ASPN (score 2 to 4). TGFBi expression was found in the cortex of adrenal gland and prostate ducts (score 4 to 6). A weak expression of LTBP2 was observed in medulla of the adrenal gland and thong epithelium (score2).

Figure 10 – illustrates validation of ASPN, LGAL3, LTBP2, MYOF, POSTN, TNXB (TNC) and TGFBI proteins in pancreas adenocarcinoma using Western blot analysis. Each lane indicates one individual patient where PN refers to normal pancreas tissue, PI to inflammatory tissue and PT to tumoral pancreas. For normalization purposes, following the transfer Ponceau red staining of the WB membrane was performed (results not shown).

Figure 11 – illustrates MRM based validation of FN1, PRELP, TGM2 and AGRN modulated proteins. The relative quantification is calculated as a mean value with a corresponding standard deviation.

Figure 12 – illustrates assessment of migration ability (chemotaxis/ haptotaxis) of BXPC3 pancreas carcinoma cells following the depletion of (A) TGFBI and (B) MYOF proteins by siRNA. Error bars indicate standard deviation of means calculated from three independent replicates. Significance was evaluated using two-sided Student's t-test (where *** indicate p value < 0.01).

Figure 13 – illustrates the accessible differentially expressed proteins identified in liver tumor (meta) tissue (n = 3) compared to normal tissue (n = 3).

Figure 14 – illustrates immunohistochemical evaluation (Box-Plot) of the identified modulated proteins in colorectal carcinoma liver metastases. **(A)** AEBP1: compared to normal hepatocytes, the positivity in the tumor stroma was significant ($P \leq 0.01$), the tumoral cells were negative; **(B)**

EMILIN1: statistical testing indicated significant difference between tumor stroma and normal liver tissue ($P \leq 0.01$), tumor cells were negative; **(C)** LTBP2: statistical testing indicated significant difference between hepatocytes and tumor stroma ($P \leq 0.001$); **(D)** POSTN: statistical testing indicated significant difference between tumor stroma and normal liver tissue ($P \leq 0.01$); low positivity was detectable in the tumoral cells, however not statistically different when compared to the normal hepatocytes. **(E)** TGFBI: compared to hepatocytes, the positivity in the tumor stroma and cells was significant ($P \leq 0.001$); Statistical evaluation was conducted using Wilcoxon rank sum test.

Figure 15 – illustrates validation of (A) AEBP1, (B) EMILIN, (C) LTBP2, (D) POSTN, (E) TGFBI and (F) MYOF protein expression in normal liver and colorectal carcinoma liver metastases tissues using Western blot analysis. (A-F) Represents the densitometrical analyses of Western blot performed on seven individual patients. Error bars represent standard deviation of means. Following the transfer HSC70 protein was used for normalization purposes. Within a liver metastasis lesion several zones were assessed: healthy liver tissue surrounding the metastasis (peritumoral), the rim of the metastasis (rim) and the middle of the lesion (center). All samples were compared to the normal liver tissue found far away from the lesion site (normal).

Figure 16 – illustrates assessment of migration ability (chemotaxis/ haptotaxis) of colorectal carcinoma SW1222 cells following the depletion of MYOF protein by siRNA. Error bars indicate standard deviation of means calculated from three independent replicates. Significance was evaluated using two-sided Student's t-test (where *** indicate p value < 0.01).

Figure 17 – illustrates assessment of migration ability (chemotaxis/ haptotaxis) of colorectal carcinoma SW1222 cells following the pre-incubation with a polyclonal MYOF antibody (Anti-MYOF). Error bars

indicate standard deviation of means calculated from three independent replicates. Significance was evaluated using two-sided Student's t-test (where ** indicate p value = 0.05).

Figure 18 – details accessible proteins obtained from the analysis of various different tumour tissues using the method of the invention, and details the peptide sequences found which lead to the identification of the protein biomarker.

BIOMARKERS FOR BREAST CANCER

The following study illustrates the invention, both the method of the invention and novel breast cancer biomarkers. The study isolated accessible protein biomarkers from breast cancer biopsy samples, and demonstrated their use as biomarkers for breast cancer.

All the individuals involved in the current work were informed in detail regarding the aims of the study and gave their written consent. The purpose of this project and the experiments undertaken complied with the regulations and ethical guidelines of the University of Liège, Belgium. The study is divided into two parts: i) technical (demonstrating the value, reproducibility and accuracy of the method) and ii) biological (proof of concept study for the discovery and validation of novel biomarkers). The technical part employed one "master sample" which was prepared as a pool of equal amounts of all the tissue samples involved in the MS analysis (all individuals, both tumoral and normal specimen). The biological part and specifically the discovery of modulated proteins was conducted using proteins isolated in the technical part. Finally, the validation of selected differentially expressed proteins was conducted on a separate group of breast cancer patients (25 tumoral and 13 normal individuals). All patients involved in the immunohistochemistry validation study were diagnosed with ductal breast adenocarcinoma, had clinical grades of at least 2 and presented no metastasis at the time of surgery.

Regarding the number and type of replicates involved in the study it is to be noted: i) all analyses conducted in the technical part involved respectively three full technical replicates (from tissue solubilization to MS analysis), performed on separate days and using the "master sample"; ii) the investigations conducted in the biological part, specifically discovery phase using MS, are single replicates of matched tumoral and normal samples originating from five individuals;

Tissue sample preparation. Pieces of fresh human breast cancer biopsies obtained from the Pathology Department of the University Hospital of Liège, Belgium, were immediately sliced and soaked in freshly prepared EZ-link Sulfo NHS-SS biotin (1 mg/ml, Pierce, Rockford, IL, USA) solution. Tissue samples were then snap-frozen in liquid nitrogen and pulverized using a Mikro-Dismembrator U (Braun Biotech, Melsungen, Germany). Approximately 100 mg of tissue powder was dissolved in the PBS buffer (50 mM PBS, 0.5 M NaCl, pH=7) containing protease inhibitor (PI) cocktail (Halt™, Pierce, Rockford, IL, USA), 0.5 mM oxidized glutathione (GSSG) and level 1 internal standard mix (IS1 consisting of bovine alpha-2-HS-glycoprotein, bovine casein and biotinylated chicken ovalbumin [performed by incubation of ovalbumin with EZ-link Sulfo NHS-SS biotin reagent]; ratio spike/sample – 1/200). Homogenates were than sonicated (2 x 30 s) with a 2 mm microprobe and centrifuged at 20,000 x g for 10 minutes at 4°C. Human serum albumin (HSA) and immunoglobulins (IgGs) were eliminated using Qproteome HSA and IgGs Removal Kit (Qiagen, Valencia, CA, SA). The remaining pellet was suspended in the RIPA buffer (1% Nonidet P40 (NP40), 0.5% deoxycholic acid (DOC), 0.1% SDS, 0.5 mM GSSG and PI cocktail in PBS, pH=7.0), sonicated (2 x 30 s) and centrifuged at 20,000 x g for 10 minutes (at 4°C). The sample was subjected to HSA and IgGs depletion as described above. 2 % SDS solution was added to the remaining insoluble pellet following a final re-solubilization. The sample was then centrifuged (as mentioned above) and the supernatant collected. All lysates from the three solubilization steps were finally pooled together and boiled for 5 minutes.

Isolation of biotinylated proteins. The total protein extract was mixed with 100 µL/mg Streptavidin (SA) resin (Pierce, Rockford, IL, USA) for 120 minutes under rotational conditions at room temperature. After the streptavidin

incubation, the supernatant was retained for the subsequent glycoproteomic analysis (fraction 1). The streptavidin beads were washed 4 times with 0.5 mL buffer A (1% NP40, 0.1% SDS and 0.5 mM GSSG in PBS buffer), 4 times with 0.5 mL buffer B (0.1% NP40, 1.5 M NaCl and 0.5 mM GSSG in PBS buffer), 2 times with 0.5 mL buffer C (0.1 M Na_2CO_3 and 0.5 mM GSSG in PBS buffer at pH=11.0) and finally 2 times with 0.5 mL PBS buffer at pH=7 without GSSG. The biotinylated proteins were eluted 2 times with 0.4 mL of 100 mM dithiothreitol (DTT) and incubated at 60°C for 30 minutes (fraction 2). Fraction 1 was also reduced in 100 mM DTT. Both fractions were alkylated with 150 mM iodoacetamide for 30 minutes in the absence of light. At this stage level 2 internal standard (IS2 consisting of bovine beta-lactoglobulin [1/200]) was added to both fractions. Proteins were then precipitated with 20% trichloroacetic acid (TCA) at 4°C overnight. The protein pellets (fractions 1 and 2) were solubilized in 50 mM NH_4HCO_3 and digested using trypsin (Promega, Madison, WI, USA) (1:50 protease/protein ratio) at 37°C overnight. The biotinylated peptides (fraction 1) were further processed using MS. Fraction 2 was used for the isolation of glycopeptides as described below.

Isolation of glycopeptides. The digested protein sample was acidified with 30 µL of 1 % HCl, transferred onto the C18 Sep-pak column (Waters, Milford, Massachusetts) and washed with 3 x 1 mL of 0.1 % formic acid solution. The peptides were eluted using acetonitrile (80%) and evaporated to dryness. The peptide-containing sample was dissolved in the oxidation buffer (100 mM sodium acetate, 150 mM NaCl at pH=5.5) and incubated with 10 mM sodium periodate (Pierce, Rockford, IL, USA) for 1 hour in the dark. Following this, 10 µL of 120 mM sodium sulfite was added and incubation was extended for an additional 10 min. The sample was loaded onto hydrazide resin (Bio-Rad, Hercules, CA, USA) and the glycopeptides were bound at RT overnight. The glycopeptide-free flow-through was collected for the subsequent MS analysis (non-glycosylated proteins, NGP or R). After extensive washing (2 times each with H_2O, 1.5 M NaCl, methanol, 80% ACN and 50 mM NH_4HCO_3) the hydrazide resin was loaded with 50 mM NH_4HCO_3 solution and incubated overnight with 500 units of PNGase F (New England Biolabs, Ipswich, MA, USA) at 37 °C. After the incubation period, the glycopeptide-containing flow-through (G) was collected and desiccated.

MS analysis. Peptides originating from biotinylated, glycosylated and also the non-glycopeptide fractions were desalted using C18 ZipTip pipette tips (Millipore, Billerica, MA, USA) prior to mass spectrometry analysis. Following this, 5 µg of peptide containing samples were dissolved in 18 µL of 100 mM ammonium formiate buffer. To the dissolved samples level 3 internal standard mix was added (IS3 composed of MassPREP Digestion Standard Mixture 1™ [Waters Corporation] – containing equimolar mix of yeast alcohol dehydrogenase, rabbit glycogen phosphorylase b, bovin serum albumin and yeast enolase; final concentration in 18 µL sample was adjusted to 135 fmol of yeast alcohol dehydrogenase). Of the sample prepared, 9 µL was injected corresponding to an estimated protein load of 2.5 µg. For the MS analysis the nano Aquity® UPLC (Waters) was coupled on-line with the SYNAPT G1 qTOF system (Waters). The configuration of the UPLC system was the following: trap column Symmetry® C18 5 µm, 180 µm × 20 mm (Waters), analytical column BEH® C18 1.7 µm, 75 µm × 150 mm (Waters), solvent A (0.1 % formic acid in water) and solvent B (0.1 % formic acid in ACN). The flow rate was set at 300 nL/min and the gradient had the following composition: 0 min, 97 % A; 90 min, 60 % A. The MS acquisition parameters were: data independent, alternate scanning (MS^E) mode, 50–1500 *m/z* range, ESI+, V optics, scan time 1 s, cone 30 V and lock mass [Glu1]-Fibrinopeptide B $[M+2H]^{2+}$ 785.8426 *m/z*. Raw data was processed (deconvoluted, deisotoped, protein identification, absolute and relative quantification) using ProteinLynx GlobalSERVER® (PLGS) v2.4. The processing parameters were: MS TOF resolution and the chromatographic peak width were set to automatic, low-/elevated- energy detection threshold to 250/100 counts, identification intensity threshold to 1500 counts and lock mass window to 785.8426 ± 0.35 Da. For protein identification UniProt® human database served as the reference (canonical sequence data with 20,280 enteries). Peptide modification carbamidomethylation was set as fixed and oxidation (M) as variable. In addition, for glycoprotein analysis, deamidation (N) was included as a variable modification as well. A response factor (2200) for the conversion of the peptide intensities into absolute quantities was deduced previously following a repeated injection of the alcohol dehydrogenase (yeast, SwissProt P00330) digest. This response factor was kept constant throughout the entire study.

Routinely, IS1, IS2 and IS3 were checked for the correct relationship between the spiked and the measured absolute amount as well as for the relative ratio between the compared samples. The IS tolerances for both absolute quantities and relative ratios had to be within ± 25 % deviation in order for the data set to be acceptable and included in further analysis. PLGS® software calculated score and false positive rate (FPR) for each individual protein hit. Within the present study a protein was considered as identified if the FPR was ≥ 96 % and the score ≥ 80.

Glycoprotein data analysis. Regarding the glycoproteins, the processed MS data (deconvoluted spectra) were submitted for the database search, first separately for the fraction obtained from the hydrazide beads (G) and then combined with the flow-through fraction (R). Following this, all the glycoproteins originating from the hydrazide beads were filtered out with a home-made program (ARAC v3.5, available on request). This program checked for the presence of deamidated asparagines at the consensus sequence site (NXS/T, where X can be replaced by any amino acid except proline) for each of the peptides in question. In this initial step, a certain number of glycopeptides could immediately be assigned to a respective glycoprotein (G fraction). The remaining glycopeptides were not specific enough or had lower scores so that they could not be unambiguously associated with a protein. In order to help assign these peptides to a protein they were matched with the peptides from the R fraction analysis where several non-glycosylated peptides in conjunction with the glycosylated peptides (from the G fraction) permitted significant protein identification. This new combined pool of proteomic results was named the GR fraction.

Immunohistochemical validation of selected biomarkers. The expression of laminin α-5 and myoferlin was assessed by immunohistochemistry in formalin-fixed paraffin-embedded breast tissue sections. Samples originating from 25 tumoral and 13 normal breasts were immunostained using anti-laminin α-5 (Millipore, Billerica, US, dilution 1/100) and anti-myoferlin (Abcam, Cambridge, UK, dilution 1/400). Tissue sections of 5 µm thickness were unparaffined by two baths in xylene during 5 min and hydrated in the methanol gradient (100%, 95%, 70%, 50% and H_2O). Blocking of endogenous peroxidase was performed by 30 minutes incubation with 3% H_2O_2 and 90% methanol. Antigen retrieval was

conducted in 10 mM citrate buffer (pH 6) using 95°C water bath for 40 min. Following 30 min blocking in PBS-normal serum solution (150µl normal goat serum and 20µl Tween 20 in 10ml PBS) the sections were incubated with the primary antibody overnight at 4°C. Sections were then incubated with the biotinylated secondary antibody for 30 minutes and further with avidin biotin complex kit (ABC kit) for additional 30 minutes. 3,3'-diaminobenzidine tetrachlorhydrate dihydrate (DAB) with 5% H_2O_2 was used for colorization. The slides were finally counter-stained with hematoxylin. Immunostaining was assessed by evaluating the samples for percentage of positive cells (five arbitrary units/ classes: 0 = 0%, 1 = 0-25%, 2 = 25-50%, 3 = 50-75% and 4 = 75-100%) and for staining intensity (four arbitrary units/ classes: 0 = no staining, 1 = weak, 2 = moderate and 3 = strong). The results obtained by these two scales were then multiplied together yielding a single value named score (y axis in the Figure 5). Statistical analyses were performed using unpaired two-tail Student *t* test for comparison between groups. Gaussian distribution was verified by the D'Agostino-Pearson test. Two-tail Mann-Whitney *U* test was used when Gaussian distribution and/or heteroscedasticity were not confirmed (in the case when the normal sample did not display any positive cells or/and no staining). P-values $\leq$ 0.05 were considered as significant. Statistical analyses were conducted using PRISM software (GraphPad Software, San Diego, California, version 4.0b).

RESULTS

Isolation of accessible proteins from tissue samples – the sequential method

The schematic overview of the sequential method is displayed in Figure 1. The method is composed essentially of three distinct steps: i) isolation of biotinylated proteins, ii) purification of the glycopeptides and iii) analysis of the remaining peptides. The latter fraction served for in-silico complementation of the non-assigned glycopeptides. In addition, as this fraction contained a significant number of accessible proteins, it was allowed to contribute this group of interest to the overall pool of modulated proteins. Due to the complexity of the method several rapid tools for monitoring the inter-replicate reproducibility were introduced. These consisted chiefly of: i) examining the flow-through of the biotin step for remaining biotinylated proteins and ii) measurement of the flow-through fraction after the hydrazide capture for residual glycopeptides. Further

process controls were the internal standards which were spiked at three different steps, amounting to 8 individual proteins of non-human origin. Overall it can be said that the capture of both biotinylated proteins and glycopeptides was efficient and allowed for almost perfect specificity and minor samples loss. This is particularly obvious from the data presented in the Figure 2A. Here it is noteworthy that in the biotinylated protein fraction, quantitative recoveries of biotinylated albumin and negligible contamination of fetuin was detected (IS1). As far as the other two fractions are concerned (glyco and rest), fetuin was reproducibly quantified in both fractions (owing to both glycosylated and non-glycosylated peptides) whereas casein (IS1), being a pure phosphoprotein, was only recovered in the rest fraction. Regarding the reproducibility of the protein digestion and subsequent purification steps, good recoveries of beta-lactoglobulin (IS2) in all fractions except the glycopeptide one (for it is not a glycosylated protein), indicate that the quantitative and qualitative variability introduced by these steps is within the acceptable limits.

The present study used MS^e technology to perform absolute label-free quantification. The method used highly reproducible HPLC and data independent alternate scanning. In addition high mass accuracy measurements are provided by an orthogonal time-of-flight mass spectrometer and "on the flight" acquisition of lock-mass allowing for subsequent recalibration. The data was processed based on the detection and correlation of all detectable precursor and fragment ions sharing the same chromatographic profile. Rapid alteration between low- and elevated energy states applied to the collision cell allowed for the simultaneous quantification and identification of proteins in a single experiment. IS3 was observed to spike in the sample at the very late step of the sequential method (last step before the sample is injection in the UPLC) which allowed for a good estimate of the performance of the quantification approach. As outlined in the Figure 2A/B, the quantification of four proteins demonstrated that absolute quantification is feasible and accurate. However, a slight overestimation of the protein amounts especially for albumin (bovine) and glycogen phosphorylase b (rabbit) is evident. This could be related with the presence of human homologues found abundantly in the sample. Following the rationale that this variation is well

below the 2-fold ratio chosen as the threshold for claiming differential expression of a protein, this deviation was not considered as significant.

Actual results using the method of the invention and the peptides found are given in Figure 18.

Comparison of the sequential approach with the individual methods

In order to determine the exact benefit of combining two known procedures in a new method, the sequential technique was compared with the individual method parts respectively. For this purpose it was necessary to perform the second part of the method, the isolation of the glycopeptides, as a "standalone" technique. As far as the technical aspects of the method are concerned, the direct isolation of glycosylated peptides and the analysis of the remaining rest-fraction produced similar results as the sequential method. As shown in Figure 2A and 2B the glyco-fraction recovered specifically glycosylated fetuin, whereas biotinylated ovalbumin and casein were in-addition to fetuin only present in the rest-fraction. Internal standard 2 and 3 indicate reproducible digestion, purification and MS quantification.

As far as the absolute numbers of accessible proteins is concerned, the biotinylated protein fraction in the combinatory method isolated on average 310 proteins. The number of accessible proteins in the glycopeptide fraction (both in the sequential and standalone approach) was approximately 80 which increased to 110 following the in-silico combination method. This clearly demonstrates the value of performing this operation especially for non-assigned glycopeptides (an increase of ~30%). In the rest fraction on average 200 potentially accessible proteins were confidently identified. The combination of all the method parts yielded in the sequential setting over 410 proteins (+ 30 % in comparison to the biotinylation alone) and in the glycopeptide setting alone (including the corresponding rest fraction) 210 proteins (- 50 % in comparison to the sequential method). The analytical procedure of removal of previously biotinylated proteins is clearly superior, in terms of the percentage of identified unique proteins (~45

%), to both glycopeptide and the analysis of remaining (non-glycosylated and non-bioinylated) proteins. However, the joining of the glycopeptide isolation brings additional ~36 % of unique, potentially accessible, proteins. The remaining pool of proteins (rest-fraction), due to a relatively high absolute numbers and high percentage of membrane and extracellular proteins, bears another ~25 % of unique proteins of interest.

Biomarker validation using immunohistochemistry

Using the MS-based semi-quantitative approach, laminin α-5 and myoferlin were found differentially overexpressed in the breast cancer samples. A study using 25 patients diagnosed with breast cancer, and the control group of 13 normal individuals, demonstrated that both laminin α-5 and myoferlin showed a strong positive staining in human breast cancer tissue. The corresponding normal breast tissue did not display any significant positivity for either of the antibodies used (Figure 5).

BIOMARKERS FOR PANCREATIC CANCER

The method of the invention was also used to identify biomarkers for pancreatic cancer. The method used was as described above with reference to breast cancer except the samples were of pancreatic cancer tissue. Some specific methods used for validation and functional studies are described below.

MRM-based relative protein quantification

The validation of some biomarkers was conducted using MRM. Briefly, 150 µg of total protein extract (buffers and conditions are detailed in the "sample preparation" section) were spiked with 2 µg of bovine β-lactoglobulin (internal standard [IS1], used for the normalization of sample preparation) and precipitated using 20 % TCA. The pellet was partially dissolved in 150 µL of 50 mM ammonium bicarbonate buffer (pH 8) and subjected to trypsin digestion (protein/ protease, 50/1). MRM ion selection was based upon the 2D-HPLC-MS/MS run (from proteomics experiments described above), considering the merged results of tumoral and normal samples. Peptides which had Mascot® scores above 80, were unique for the given protein and their MS/MS spectra had fully assigned fragments were regarded as possible targets. Each protein of interest was targeted

using one specific peptide and several transitions (characteristic fragments). Only samples which were positive for all transitions were quantified. Furthermore, the specificity of the precursor/ product ion transitions was verified using the publicly available algorithm: *http://prowl.rockefeller.edu/prowl/pepfrag.html*.

Prior to the injection into the LC-MS system, a synthetic peptide (EEEGTTGPDR* [isotopically labelled arginine ^{13}C and ^{15}N], retention time 1.3 min [IS2]) was added to each sample (0.025 pg/µL). The known transitions (parent ion 550.05 $+2[H]^{2+}$ and fragments 185.2, 241.2 and 713.5 m/z) of this peptide were used to normalize the ionization differences between the individual injections of different samples.

Approximately 20 µg of digested samples were injected into the HPLC system (Alliance 2690 – Waters, Milford, MA, USA) coupled to the MS instrument (Quattro Ultima Platinum – Waters) set in the MRM mode. Sample separation was conducted using the 2.1 mm x 150 mm C-18 column (Polaris C-18 A media with 3 µm beads, 200 Å pore size [Varian Inc., Palo Alto, CA, USA]) run at the flow rate of 200 µL/min. The linear gradient was set as follows: t = 0 min, 100 % water (+ 0.1 % v/v acetic acid) and t = 16 min, 60 % water and 40 % acetonitrile (+ 0.1 % v/v acetic acid). The relevant parameters of the MS instrument were: capillary voltage 2.6 kV, desolvation temperature 115 °C, cone desolvation gas 630 L/h and collision cell pressure at 2 µbar. The data were collected and evaluated using the MassLynx (Waters) software version 4.1.

For the purpose of data evaluation the peak area (PA) of the extracted ion chromatogram (XIC) of specified peptides and transitions was normalized between different injections using the IS2. Only XIC above the threshold of 100 cps were considered for the analysis. Following this, the averages of the normalized PA were calculated using all the transitions for the respective peptide. Finally, a second round of the normalization was conducted using the IS1.

Cell culture and siRNA mediated knockdown.

BxPC3 (pancreas carcinoma) cells were transfected with small interfering RNA (siRNA) directed against MYOF and TGFBI at a concentration of 10 nM using lipofectamine (Lipofectamine 2000 reagent, Cat No 11668-019, Invitrogen). Sixteen hours after transfection, culture medium was replaced with corresponding fresh medium (with serum). This time point served as t=0. SiRNA directed

towards luciferase was used as a control. All siRNAs were purchased from Eurogentec (Seraing, Belgium).

Cell migration (chemotaxis/ haptotaxis) assay.

Forty eight hours following transfection, $2x10^5$ transfected BxPC3 cells were suspended in serum-free medium (0.1% BSA, 1% penicillin/streptomycin) and seeded in triplicates into the upper part of a transwell filter (diameter 6.5mm, pore size 8µm, Costar, Cambridge, MA, USA) coated with gelatine or laminin (100µg/ml). The lower compartment of transwell filter plate was filled with DMEM containing 6% serum, 1% BSA and 1% pen/strep. After 10 hours incubation at 37°C, migrating cells were fixed and stained with Diff-Quick kit (Medion Diagnostics, ref 130832, Switzerland). Pictures of each insert were taken at a 5-fold magnification and percentage of migrating cells was further quantified by densitometry using ImageJ software (NIH, USA). Three wells per condition were counted.

In this study, three normal pancreas tissues and three tumor lesions were analyzed in the proteomic discovery phase. A total of 736 and 841 proteins were identified in normal and tumoral conditions, respectively. Among these, more than 18% were found in all three normal (157 proteins) and tumoral (152 proteins) samples. Moreover, more than 29% of proteins of normal (239 proteins) and tumoral (244 proteins) samples were identified in two out of three biological replicates. This implied that 47% of the proteins found in each respective sample were also observed in at least another replicate. In total, 1139 unique proteins were identified. Their modulation was assessed by their "presence" or "absence" in tumoral with respect to normal condition. Moreover, emPAI values were used to semi quantify proteins that were found in both conditions. The emPAI value is an estimation of the protein abundance in complex mixtures. This value is calculated as a logarithm of the ratio of observed and theoretically observable peptides within the MS analysis. In the present study, a protein was considered overexpressed when it was (i) detected only in the tumoral condition, (ii) found in both conditions but more often indentified in tumoral samples, (iii) found with a fold change ratio of the emPAI (tumoral vs normal conditions) greater than two. Taking these criteria into account, 484 proteins were found as overexpressed in the pancreas tumor. Among these, 403 were exclusive to the tumoral samples

while 81 were present in both conditions. The differentially expressed proteins were analyzed using Swissprot®/Uniprot® database in order to determine their potential subcellular localization. The potential accessibility of a given protein was evaluated according to the following criteria: the protein is localized in (i) cell membrane, (ii) on cell surface or it is described as (iii) secreted or (iv) extracellular. Secreted proteins known to be found abundantly in the serum were excluded from further analysis. Among the 484 proteins overexpressed in the tumoral condition, we were able to identify 84 proteins as potentially accessible whereas 310, 18 and 72 were described as unaccessible, serum and unknown proteins, respectively. The biological profiles of the 84 differentially expressed accessible proteins were determined according to their Gene Ontology annotation. The majority of the proteins were involved in cell communication and signal transduction (30%) as well as cell growth and/or maintenance (30%). The other biological processes were less represented: protein metabolism (9%), general metabolism and energy pathways (9%) and transport (6%).

Due to the unique method, the current study brought to light a considerable number of proteins (n=84) with potential accessible characteristics (Figure 7). Transforming growth factor beta induced protein ig-h3 (TGFBI), latent-transforming growth factor beta binding protein 2 (LTBP2), myoferlin (MYOF), asporin (ASPN), tenascin (TNXB or TNC), periostin (POSTN), galectin (LGALS3), fibronectin (FN1), prolargin (PRELP), protein-glutamine gamma-glutamyltransferase 2 (TGM2) and agrin (AGRN) proteins are novel biomarkers of pancreatic cancer.

The expression of TGFBI, LTBP2, MYOF and ASPN proteins was studied using immunohistochemistry in a series of tumoral (n=34) and non-tumoral (n=28) pancreatic tissues (Figure 8). The non-tumoral tissues were further subdivided in normal pancreas and chronic pancreatitis. This served to better characterise the tumor specificity of the selected proteins. As far as TGFBI is concerned (Figure 8A) the staining was generally found in the ECM (extracellular matrix) of the ductal adenocarcinoma tissues. Tumoral epithelial cells showed moderate cytoplasmic staining. This was not observed in the normal tissue. A strong ECM expression was found in 50% of all the adenocarcinoma analyzed, while the remaining cases showed lower staining. Some ducts localized in the areas of

chronic pancreatitis were moderately positive. Normal ducts showed no expression of the protein. Altogether, the average expression of TGFBI in tumoral tissue was significantly stronger ($p \leq 0.0001$) in the tumor (score ~8) when compared to the inflammatory (score ~2) and normal (score ~2) pancreas tissue. TGFBI levels in chronic pancreatitis and normal tissues were not significantly different. LTBP2 immunoreactivity was mainly found in tumoral ductal and peri-ductal stromal tissue. More than 50% of the adenocarcinoma showed a strong expression of the protein, which was significantly higher ($p \leq 0.0001$) in the tumoral (score ~5) comparing to the inflammatory (score ~2) and normal (score ~2) tissue (Figure 8B). LTBP2 protein expression in normal and pancreatitis tissues was not significantly different. The expression pattern of MYOF was assessed as well (Figure 8C). In contrast to the TGFBI and LTBP2, MYOF expression was mainly conferred to the cell membrane and cytoplasm of the tumor cells. Notable was the difference between different tumor grades, where the staining intensity appeared stronger in grades 2 and 3 (more advanced tumors) and less in grade 1 (less advanced tumors). Normal ducts including the inflammatory ones were mainly negative. In summary, the positivity in the tumoral tissues (score ~6) was significantly different ($p \leq 0.001$) in comparison to the pancreatitis and normal pancreas tissues (score ~2). ASPN was exclusively detected in the extracellular matrix (ECM) and no staining was observed in either normal or tumoral epithelial cells. ASPN expression (Figure 8D) was significantly ($p \leq 0.0001$) stronger in cancer specimens (score ~8) in comparison with the inflammatory (score ~3) and normal (score ~2) tissues, were no or very low staining was noted. Although the inflammatory tissues showed only low positivity, the protein level of ASPN was significantly elevated ($p \leq 0.05$) with respect to the normal specimen.

Other normal tissue originating from adrenal gland, brain, endometrium, intestine, liver, placenta, prostate (both gland and ducts), pulmonary vein and artery, thymus, thyroid and umbilical cord tested negative for presence of ASPN (Figure 9). Regarding the expression of TGFBI in other normal human tissue, the tissue microarray analysis indicated that colon, brain, breast, endometrium, myenteric plexus, myometrium, prostate (gland), thymus and thyroid were negative. A fairly modest positivity (score 2-4) was measured in adrenal gland, heart, liver, oesophagus, tongue and prostate (ducts) (Figure 9). Concerning the

expression of LTBP2 in normal tissue, the immunohistochemistry experiments revealed that this protein was not present in most of the tissues (adrenal gland (cortex), brain, breast, endometrium, heart, intestine, myenteric plexus, kidney, liver, placenta, prostate (gland), thymus, thyroid, umbilical cord and vein). Weak positivity (score ≤ 2) was registered in medulla of the adrenal gland, epithelium of the thong, mucosa of the bronchia and pulmonary artery (Figure 9).

The expression of several biomarkers was confirmed by western-blotting experiments. As shown in Figure 10, TNXB (TNC), LTBP2, ASPN and LGALS3 were expressed uniquely in cancer samples. POSTN was detectable in all the tumoral samples, however the protein was found less expressed in the patient DI-PT-003 when compared to other individuals. Non-tumoral specimens were largely negative for all the examined proteins, however ASPN tested positive in one of the two pancreatitis patients (DI-PI-001). Furthermore, TGFBI was expressed in all cancer samples. MYOF was detectable in all the tumoral samples, however this protein was found less expressed in the patient DI-PT-003 when compared to other individuals. Non-tumoral specimens were largely negative for all three examined proteins, however TGFBI tested positive in one of the two pancreatitis patients (DI-PI-001).

Western-blot was used for these seven proteins because suitable antibodies were available while the four other were quantified using MRM. The latter represents targeted MS method where specific peptide fragments of the given proteins are detected and quantified. MRM was used for FN1, PRELP, TGM2 and AGRN and confirmed that all proteins were overexpressed in human pancreas cancers compared to the normal counterpart tissue (Figure 11).

Next, the focus has been placed on clarifying the functional role of TGFBI and MYOF in pancreatic cancer cells. We have primarily assessed the impact of these proteins on the migration capacity (chemotaxis/haptotaxis) of the pancreatic cancer cells as one of the key steps in cancer progression and development. For this purpose BxPC3 pancreas carcinoma cells were transfected with siRNA against TGFBI and MYOF and the alteration in migratory capacity was assessed 48h post-transfection (Figure 12A-B). SiRNA mediates degradation of the corresponding mRNA and hence leads to depletion of the target protein in the respective cells (not shown). As displayed in the Figure 12, depletion of TGFBI

and MYOF proteins lead to a significant inhibition of migration (-60% and -25% respectively) in pancreas carcinoma cells. The results demonstrate that both proteins can serve as suitable targets for therapeutic applications.

BIOMARKERS FOR CRC LIVER METASTASES

The method of the invention has also been used to identify biomarkers for liver metastases of colorectal carcinoma. The method used was as described above with reference to breast cancer except the samples were of liver cancer tissue.

For functional studies, the following methods were used:

Cell culture and siRNA mediated knockdown.

SW1222 (colorectal carcinoma) cells were transfected with small interfering RNA (siRNA) directed against MYOF at a concentration of 10 nM using lipofectamine (Lipofectamine 2000 reagent, Cat No 11668-019, Invitrogen). Sixteen hours after transfection, culture medium was replaced with corresponding fresh medium (with serum). This time point served as t=0. SiRNA directed towards luciferase was used as a control. All siRNAs were purchased from Eurogentec (Seraing, Belgium).

Cell migration (chemotaxis/ haptotaxis) assay.

Forty eight hours following transfection, $2x10^5$ transfected cells were suspended in serum-free medium (0.1% BSA, 1% penicillin/streptomycin) and seeded in triplicates into the upper part of a transwell filter (diameter 6.5mm, pore size 8µm, Costar, Cambridge, MA, USA) coated with gelatine or laminin (100µg/ml). For experiments involving blocking antibodies, non transfected cells were pre-incubated 90 minutes at 37°C in buffer (0.1% BSA, 1% penicillin/streptomycin) supplemented with a polyclonal antibody directed against MYOF or rabbit IgG (control) at a concentration of 50µg/ml. The lower compartment of transwell filter plate was filled with DMEM containing 6% serum, 1% BSA and 1% pen/strep. After 24 hours incubation at 37°C, migrating cells were fixed and stained with Diff-Quick kit (Medion Diagnostics, ref 130832, Switzerland). Pictures of each insert were taken at a 5-fold magnification and percentage of migrating cells was further quantified by densitometry using ImageJ software (NIH, USA). Three wells per condition were counted.

Figure 13 details the accessible proteins identified using liver tissue.

Adipocyte enhancer binding protein 1 (AEBP1), EMILIN1, latent transforming growth factor beta binding protein 2 (LTBP2), periostin (POSTN), transforming growth factor beta induced protein igh3 (TGFBI) and myoferlin (MYOF) proteins are novel biomarkers of colorectal carcinoma liver metastases.

In order to further validate the expression of these antigens, immunohistochemistry was performed on 14 tumoral and normal (matched) samples. Figure 14 displays the evaluation of the immunological staining. The analysis revealed that AEBP1, EMILIN1 and POSTN antigens were mostly present in the tumor stroma (all significantly up-regulated), whereas the tumor cells were predominantly negative. As further outlined in the Figure 14, TGFBI showed significant positivity in the tumor stroma (score ~8). The expression of TGFBI in tumor cells (score ~6) was less then in the stroma yet significantly higher then the normal hepatocytes (score ~2), which were predominantly negative. As far as the LTBP2 is concerned (Figure 14C), the expression of this protein was predominantly conferred to the tumor stroma (score ~4). The expression of LTBP2 in the tumoral cells and normal hepatocytes was not significantly different (both at the score ~2).

Validation of AEBP1, EMILIN1, LTBP2, POSTN, TGFBI and MYOF protein expression in normal liver and colorectal carcinoma liver metastases tissues was confirmed using Western blot analysis (Figure 15). All 6 proteins were predominantly detected in the rim of the metastasis and the middle of the lesion, but not in the normal liver tissue (Figure 15).

Next, the focus has been placed on clarifying the functional role of MYOF in colorectal carcinoma liver metastasis. The role of MYOF protein was functionally assessed in SW1222 colorectal carcinoma cells. For this purpose, and similarly to the functional analysis performed in the BXPC3 pancreas cancer cells (as described above), inhibition of migration (chemotaxis/haptotaxis) was assesses following either i) depletion of the MYOF protein using siRNA or by ii) incubating cells with a polyclonal antibody directed towards the MYOF protein. The data are outlined in the Figures 16 and 17. Inhibition of MYOF expression by siRNA as well as targeting of this protein using an anti-MYOF antibody exerted a migration-inhibitory effect on colorectal cancer cells. These data were consistent to the ones obtained in the pancreas cancer cells, suggesting the role of this protein in more then one particular malignancy.

PAGE INTENTIONALLY LEFT BLANK

CLAIMS

1. The use of one or more of protein myoferlin, protein latent-transforming growth factor beta binding protein 2, protein transforming growth factor beta induced protein ig-h3, protein asporin, protein tenascin, protein periostin, protein galectin, protein fibronectin, protein prolargin, protein protein-glutamine gamma-glutamyl transferase 2, protein agrin, protein adipocyte enhancer-binding protein 1, protein annexin A6, protein laminin alpha-2 subunit, protein laminin subunit alpha-4, protein mimecan, protein Ras-related protein Rap-2b, protein collagen alpha-1 (XIV) chain, protein collagen alpha-3(VI) chain, protein latent-transforming growth factor beta binding protein 1, protein V type proton ATPase catalytic subunit A, protein laminin alpha-4 and protein transmembrane protein 62 as a biomarker for pancreatic cancer, or a predisposition thereto.

2. The use of claim 1 wherein one or more protein myoferlin, protein latent-transforming growth factor beta binding protein 2, protein transforming growth factor beta induced protein ig-h3, protein asporin, protein tenascin, protein periostin, protein galectin, protein fibronectin, protein prolargin, protein protein-glutamine gamma-glutamyl transferase 2 and protein agrin is for use as a biomarker for pancreatic cancer, or a predisposition thereto.

3. The use of claim 1 or 2 wherein protein myoferlin is for use as a biomarker for pancreatic cancer, or a predisposition thereto.

4. The use of claim 1 or 2 wherein protein latent-transforming growth factor beta binding protein 2 is for use as a biomarker for pancreatic cancer, or a predisposition thereto.

5. The use of claim 1 or 2 wherein protein transforming growth factor beta induced protein ig-h3 is for use as a biomarker for pancreatic cancer, or a predisposition thereto.

6. The use of claim 1 or 2 wherein protein asporin is for use as a biomarker for pancreatic cancer, or a predisposition thereto.

7. The use of claim 1 or 2 wherein protein tenascin is for use as a biomarker for pancreatic cancer, or a predisposition thereto.

8. The use of claim 1 or 2 wherein protein periostin is for use as a biomarker for pancreatic cancer, or a predisposition thereto.

9. The use of claim 1 or 2 wherein protein galectin is for use as a biomarker for pancreatic cancer, or a predisposition thereto.

10. The use of claim 1 or 2 wherein protein fibronectin is for use as a biomarker for pancreatic cancer, or a predisposition thereto.

11. The use of claim 1 or 2 wherein protein prolargin is for use as a biomarker for pancreatic cancer, or a predisposition thereto.

12. The use of claim 1 or 2 wherein protein protein-glutamine gamma-glutamyl transferase 2 is for use as a biomarker for pancreatic cancer, or a predisposition thereto.

13. The use of claim 1 or 2 wherein protein agrin is for use as a biomarker for pancreatic cancer, or a predisposition thereto.

14. A method for determining the pancreatic cancer status of a subject comprising the steps of:

(a) providing a sample of material from a subject;

(b) determining the level in the sample of one or more of the biomarkers of any of claims 1 to 13; and

(c) comparing the level determined in (b) with one or more reference values from the same or a different subject.

15. The method of the claim 14 for use in any one or more of the following: to diagnose pancreatic cancer in a subject; to assess the chance of a subject developing pancreatic cancer; to advise on the prognosis for a subject with pancreatic cancer; to monitor disease progression of pancreatic cancer; and to monitor effectiveness or response of a subject to a treatment for pancreatic cancer.

16. The method of claim 14 or 15 wherein the sample is a pancreatic tissue biopsy.

17. The method of any of claims 14 to 16 wherein the reference value, to which the determined levels of the biomarker are compared, is the level of the same protein observed in one or more normal samples.

18. The method of any of claims 14 to 16 wherein the reference value is a previous value for a specific biomarker obtained for a specific subject.

19. The method of any of claims 14 to 18 wherein an increase in the level of the biomarker is indicative, or diagnostic, of pancreatic cancer.

20. The method of any of claims 14 to 18 for use in monitoring pancreatic cancer progression and/or to monitor the efficacy of a treatment administered to a subject.

21. A ligand directed to a biomarker according to any of claims 1 to 13, for use in the therapeutic and/or prophylactic treatment of a pancreatic cancer.

22. The ligand of claim 21, wherein said ligand is directed to an extracellular domain of the biomarker.

23. The ligand of claim 21 or 22, wherein the biomarker is selected from the group consisting of protein myoferlin and protein transforming growth factor beta induced protein ig-h3.

24. The ligand of claim 21 or 22, wherein the biomarker is protein myoferlin.

25. A kit for use in determining the pancreatic cancer status of a subject comprising at least one agent for determining the level of a biomarker according to any of claims 1 to 13 in a sample provided by the subject.

26. The kit of claim 25, wherein the biomarker is selected from the group consisting of protein myoferlin, protein latent-transforming growth factor beta binding protein 2, protein transforming growth factor beta induced protein ig-h3, protein asporin, protein tenascin, protein periostin, protein galectin, protein fibronectin, protein prolargin, protein protein-glutamine gamma-glutamyl transferase 2 and protein agrin.

27. The kit of claim 25, wherein the biomarker is selected from the group consisting of protein myoferlin and protein transforming growth factor beta induced protein ig-h3.

28. The kit of claim 25, wherein the biomarker is protein myoferlin.

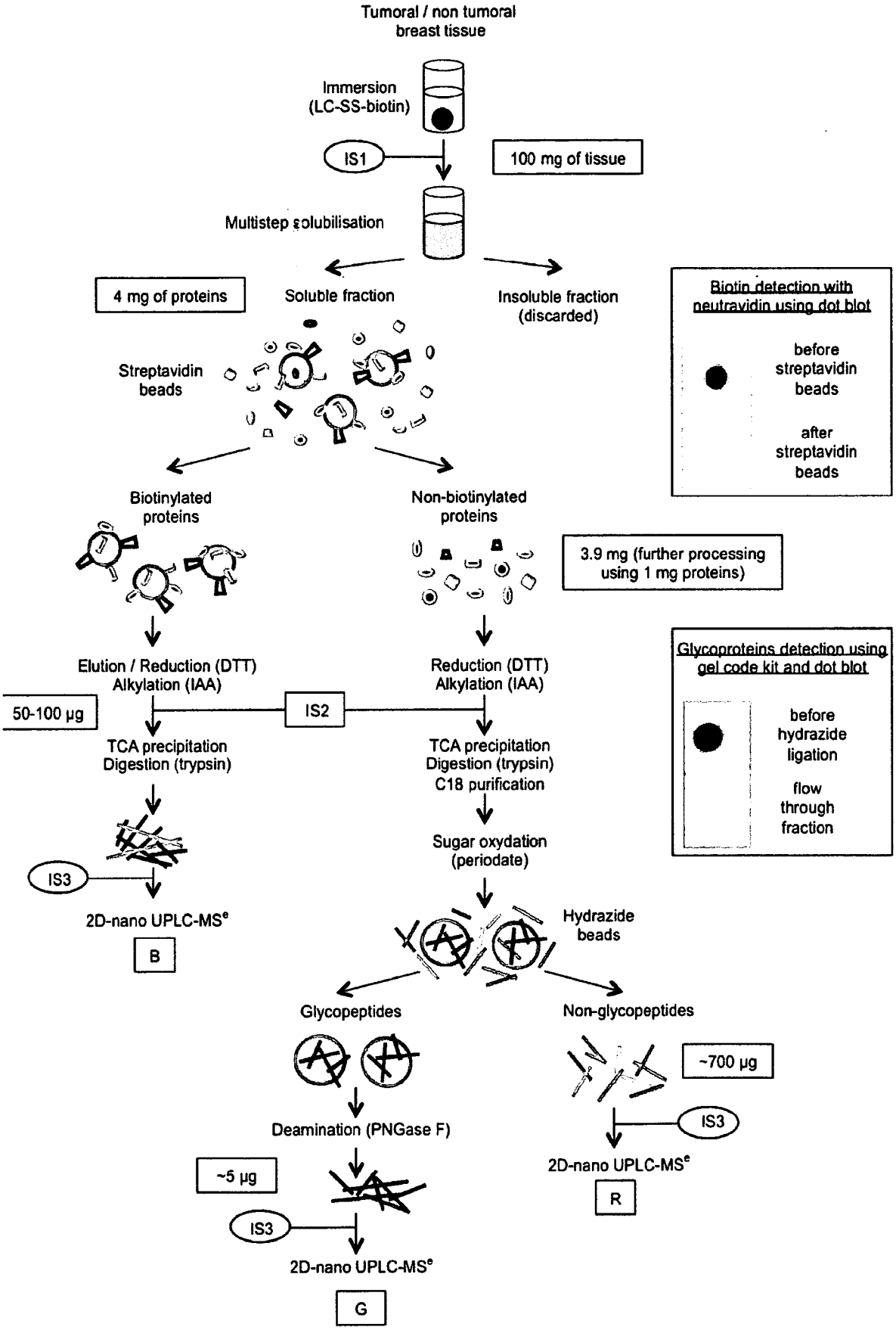

Tumoral / non tumoral breast tissue
Immersion (LC-SS-biotin)
IS1
100 mg of tissue
Multistep solubilisation
4 mg of proteins
Soluble fraction
Insoluble fraction (discarded)
Biotin detection with neutravidin using dot blot
before streptavidin beads
after streptavidin beads
Streptavidin beads
Biotinylated proteins
Non-biotinylated proteins
3.9 mg (further processing using 1 mg proteins)
Glycoproteins detection using gel code kit and dot blot
before hydrazide ligation
flow through fraction
Elution / Reduction (DTT) Alkylation (IAA)
Reduction (DTT) Alkylation (IAA)
50-100 µg
IS2
TCA precipitation Digestion (trypsin)
TCA precipitation Digestion (trypsin) C18 purification
IS3
2D-nano UPLC-MS^e
B
Sugar oxydation (periodate)
Hydrazide beads
Glycopeptides
Non-glycopeptides
~700 µg
IS3
2D-nano UPLC-MS^e
R
Deamination (PNGase F)
~5 µg
IS3
2D-nano UPLC-MS^e
G

Figure 1

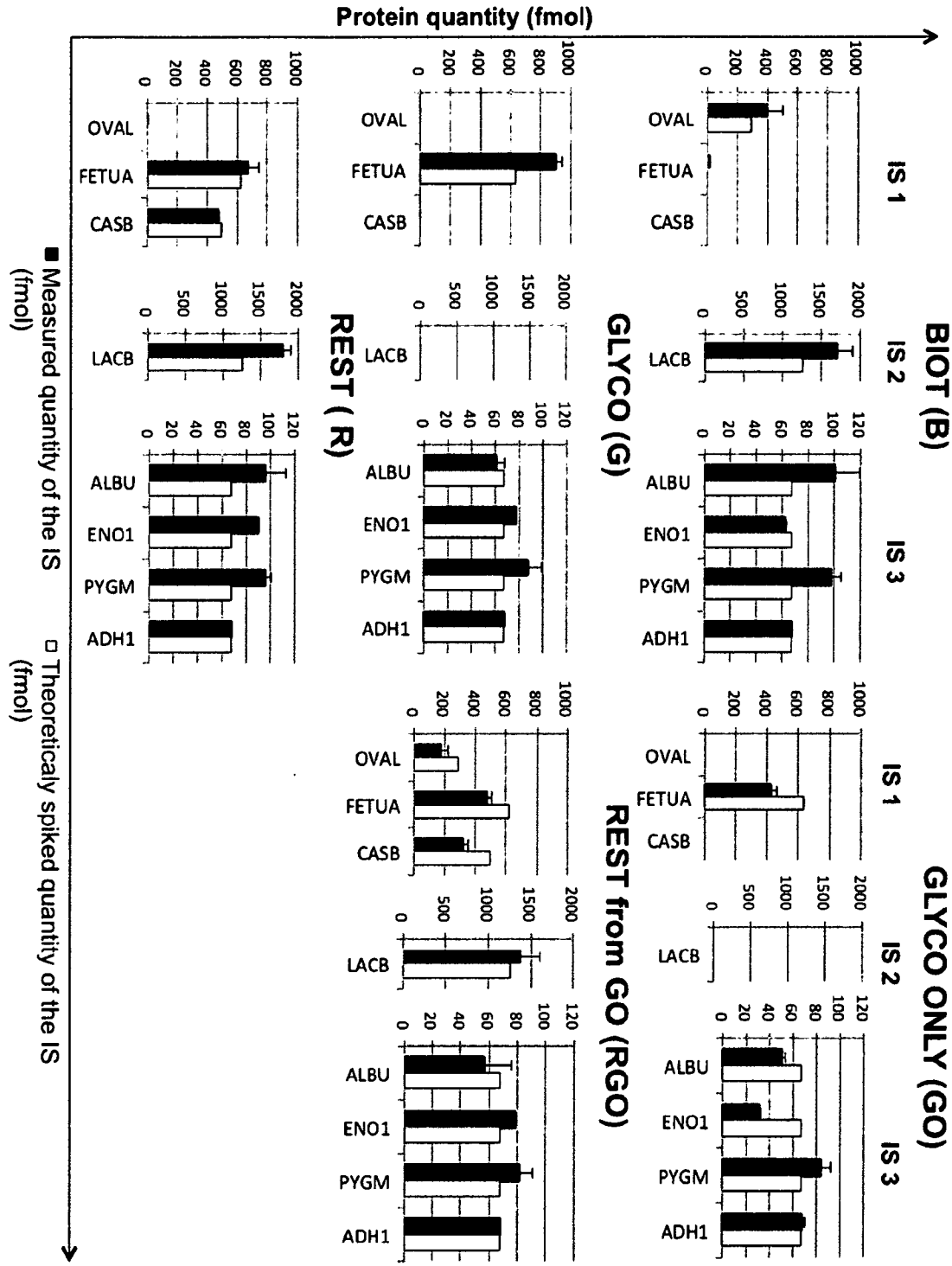

Protein quantity (fmol)
BIOT (B)
GLYCO (G)
REST (R)
GLYCO ONLY (GO)
REST from GO (RGO)
IS 1
IS 2
IS 3
OVAL
FETUA
CASB
LACB
ALBU
ENO1
PYGM
ADH1
Measured quantity of the IS (fmol)
Theoretically spiked quantity of the IS (fmol)

Figure 2

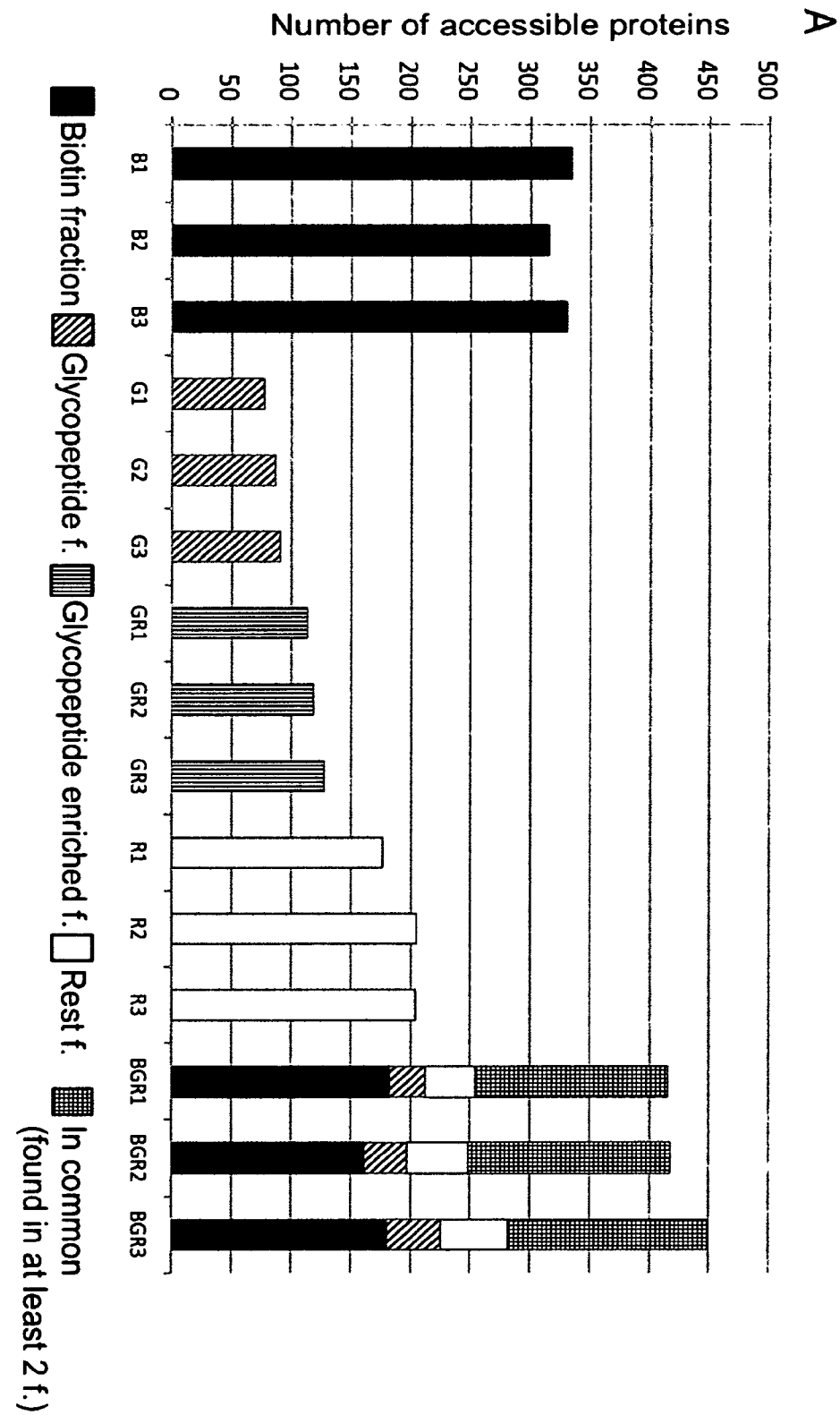

A
Number of accessible proteins
0
50
100
150
200
250
300
350
400
450
500
B1
B2
B3
G1
G2
G3
GR1
GR2
GR3
R1
R2
R3
BGR1
BGR2
BGR3
Biotin fraction
Glycopeptide f.
Glycopeptide enriched f.
Rest f.
In common
(found in at least 2 f.)

Figure 3A

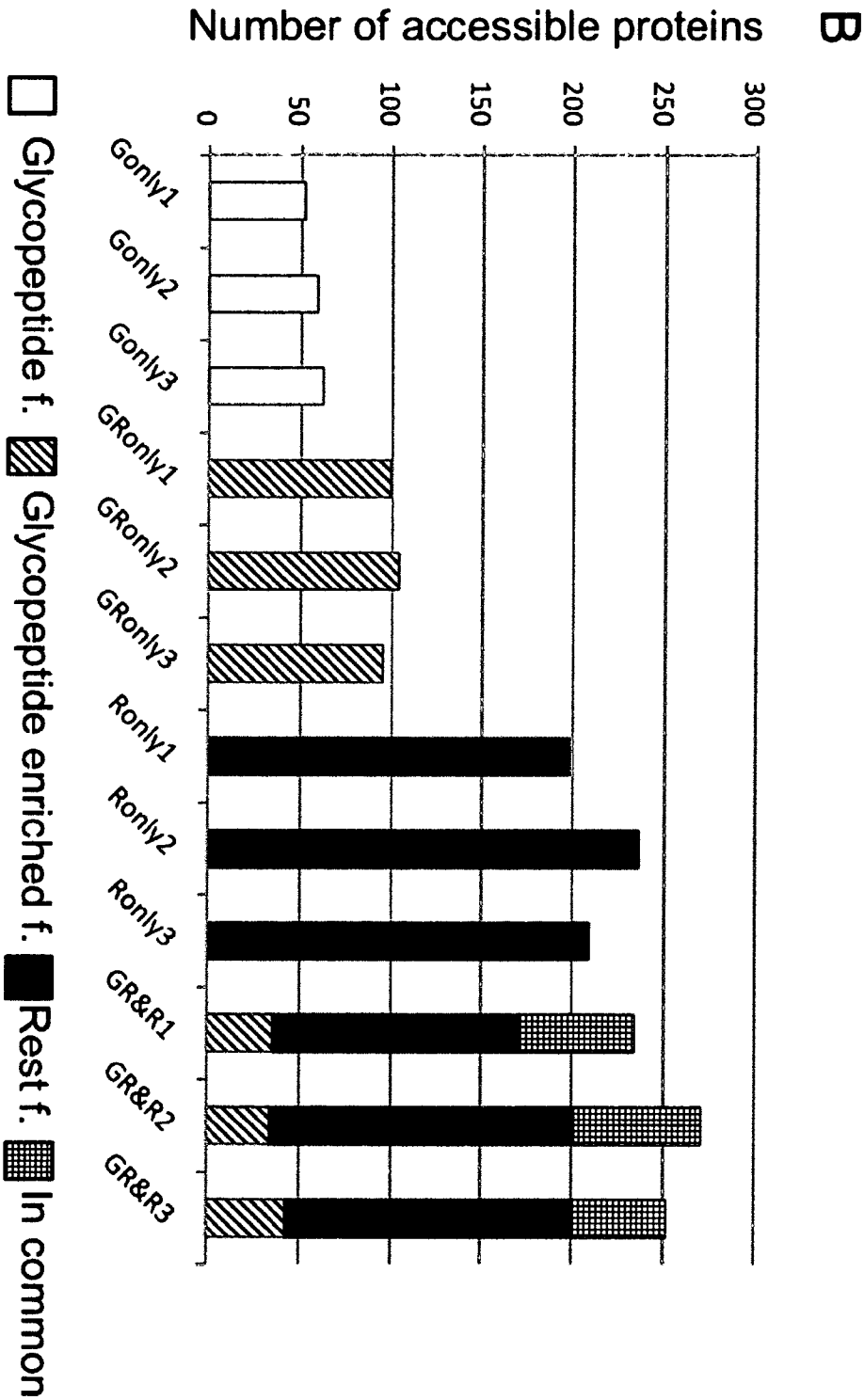

B
Number of accessible proteins
0
50
100
150
200
250
300
Gonly1
Gonly2
Gonly3
GRonly1
GRonly2
GRonly3
Ronly1
Ronly2
Ronly3
GR&R1
GR&R2
GR&R3
Glycopeptide f.
Glycopeptide enriched f.
Rest f.
In common

Figure 3B

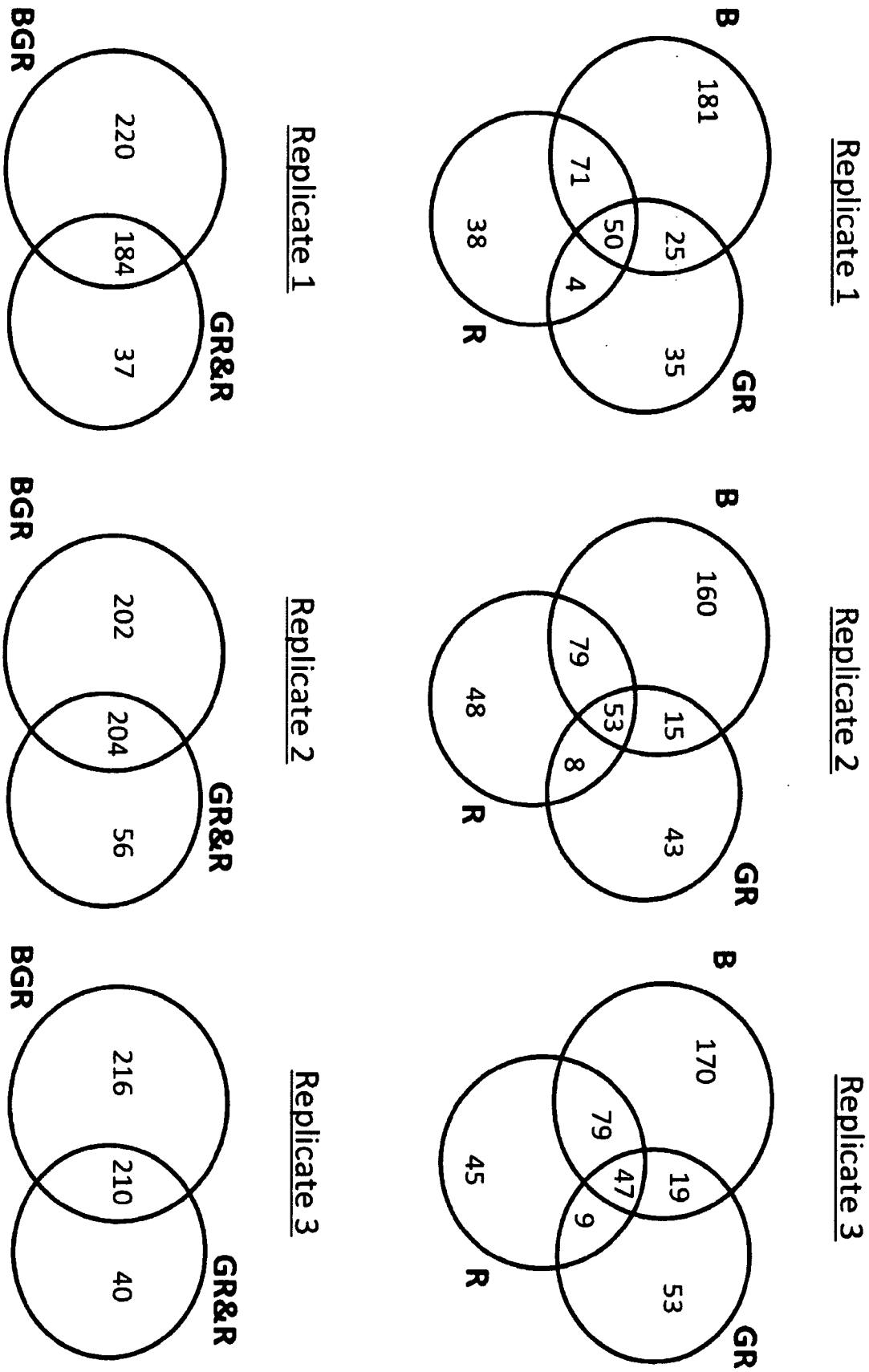
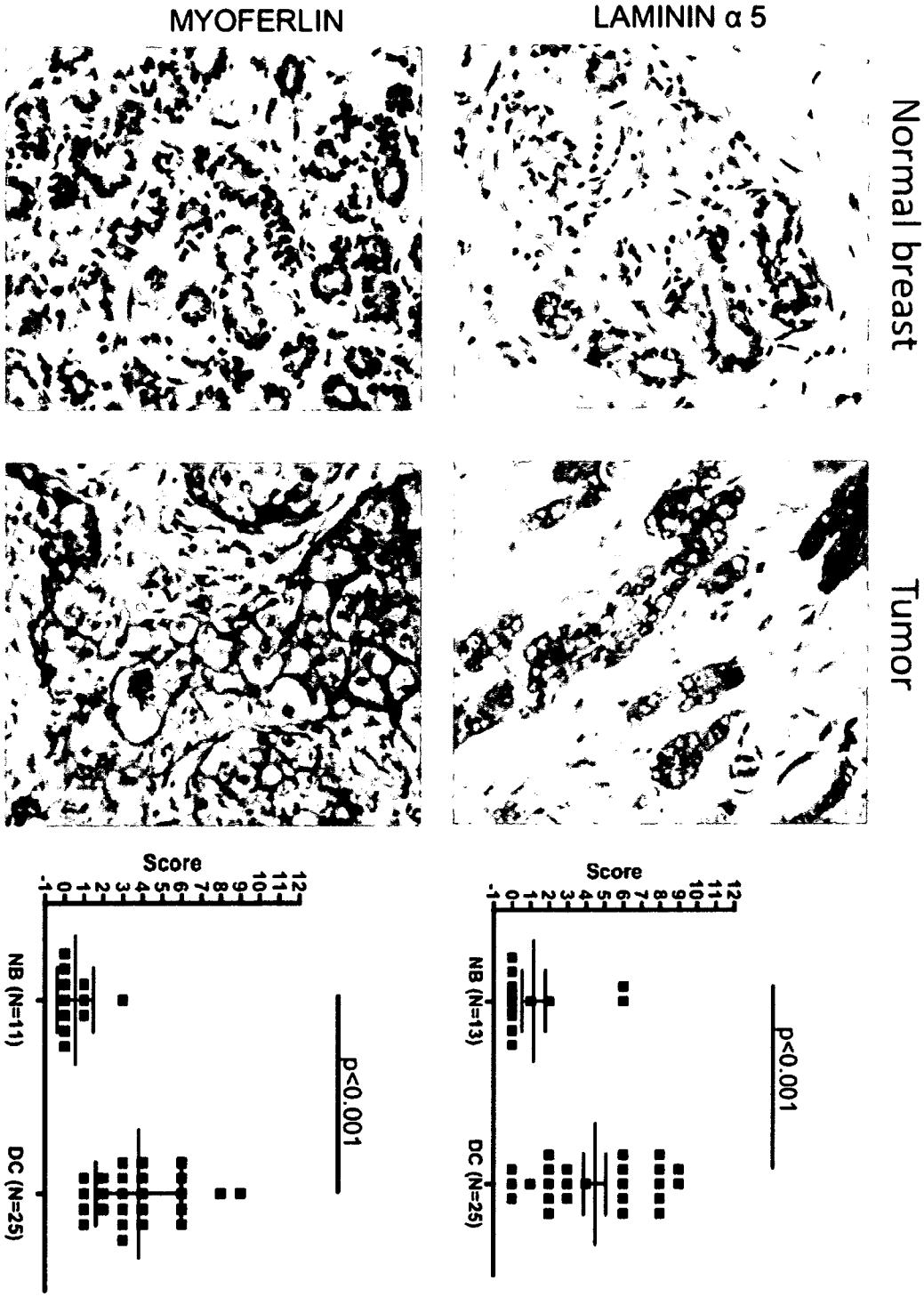

Replicate 1
B
181
71
50
25
38
4
35
R
GR
Replicate 2
B
160
79
53
15
48
8
43
R
GR
Replicate 3
B
170
79
47
19
45
9
53
R
GR
Replicate 1
BGR
220
184
37
GR&R
Replicate 2
BGR
202
204
56
GR&R
Replicate 3
BGR
216
210
40
GR&R

Figure 4

Figure 5


MYOFERLIN
LAMININ α 5
Normal breast
Tumor
Score
NB (N=11)
DC (N=25)
p<0.001
Score
NB (N=13)
DC (N=26)
p<0.001

Accession	Protein names	Gene symbol	Normal	Tumor	Ratio T/N	Subcellular locations	Biological process	Fraction
Q8IUX7	Adipocyte enhancer-binding protein 1	AEBP1	3	3	3.54	S; Cy; N	CG	G/B/N
O00468	AGRIN	AGRIN	1	3		S; E	CG	G
Q65813	ANKRD26-like family C member 1A	A26CA	3	3	1.39	PIM	U	G/B/N
P00505	Aspartate aminotransferase	AATM	0	2		C; MM	M; E	N
Q6DD88	Atlastin-3	ATLA3	0	2		I; M; MP	-	G/B/N
P07384	Calpain-1	CAN1	1	3		C; Cy	M	B/N
P17655	Calpain-2	CAN2	1	3		C; Cy	M	N
P16070	CD44	CD44	2	3	1.05	M ; SP	C ;ST	G/B
Q5ZPR3	CD276	CD276	0	2		M; SP	I	G
Q99715	Collagen alpha-1(XII) chain	COCA1	3	3	3.43	S; E	C; ST	G/B/N
P02751	Fibronectin	FINC	3	3	2.07	S; E	-	G/B/N
P19440	Gamma-glutamyl-transpeptidase 1 (CD224)	GGT1	1	2		M; SP.	M; E	G
P36268	Gamma-glutamyl-transpeptidase 2	GGT2	1	2		M; SP	-	G
P12544	Granzyme A	GRAA	0	2		S ;Cy	M	G/B
P17936	Insulin-like growth factor-binding protein 3	IBP3	0	2		S	C; ST; A	G
O15230	Laminin alpha-5	LAMA5	1	3		S; E	CG	G
Q8TF66	Leucine-rich repeat-containing protein 15	LRC15	1	2		M; SP	C; ST	G/N
Q9UBG0	Macrophage mannose receptor 2 (CD280)	MRC2	1	3		M; SP	U	G/N
Q9NZM1	Myoferlin	MYOF	0	3		C; SP; NM; Cy	CG	N
Q9BY67	Nectin-like protein 2	CADM1	0	2		C; SP	-	G
Q6UWY5	Olfactomedin-like protein 1	OLFML1	0	2		S	U	G/B
Q15063	Periostin	POSTN	3	3	2.56	S; E	C	G/B/N
P07996	Thrombospondin-1	TSP1	2	3	2.16	E	CG	G/B/N

P01033	Tissue inhibitor of metalloproteinases	TIMP1	1	3		S	CG	G	
Q9HD45	Transmembrane 9 superfamily member 3 (EP70-P-iso)	TM9S3	1	3		M; MP	T	G	
Q96M42	Uncharacterized protein C21orf129	CU129	0	3		U	U	N	
P13611	Versican (GHAP)	CSPG2	3	3	2.62	S; E	CG	G/B/N	

Figure 6A (table over previous two pages)

Accession	Protein name	Patient							
		BPSCC10/39		BPSCC10/40		BPSCC10/46		BPSCC10/47	BPSCC10/49
		Ratio T/N		Ratio T/N		Ratio T/N		Ratio T/N	Ratio T/N
Q01518	Adenylyl cyclase associated protein 1*	4.57	Tumor	2.48	Tumor	Tumor	Tumor	Tumor	Tumor
Q8IUX7	Adipocyte enhancer binding protein 1*	Tumor	Tumor	15.80	Tumor	Tumor	Tumor	Tumor	Tumor
Q8TD06	Anterior gradient protein 3 homolog	Tumor	Tumor	Tumor	Tumor	Tumor	Tumor	/	Tumor
P12830	Cadherin 1	/	Tumor	Tumor	Tumor	Tumor	Tumor	Tumor	Tumor
Q99439	Calponin 2*	Tumor	Tumor	Tumor	Tumor	Tumor	Tumor	Tumor	Tumor
Q15417	Calponin 3*	Tumor	Tumor	Tumor	Tumor	Tumor	Tumor	/	Tumor
P27797	Calreticulin*	6.69	Tumor	2.36	Tumor	Tumor	Tumor	14.73	2.08
P49747	Cartilage oligomeric matrix protein	Tumor	Tumor	Tumor	Tumor	Tumor	Tumor	/	/
P35221	Catenin alpha 1	Tumor	Tumor	Tumor	Tumor	Tumor	Tumor	Tumor	Tumor
P26232	Catenin alpha 2	Tumor	Tumor	Tumor	Tumor	Tumor	Tumor	/	Tumor
Q13740	CD166 antigen	Tumor	Tumor	Tumor	Tumor	Tumor	Tumor	/	/
Q9Y696	Chloride intracellular channel protein 4	Tumor	Tumor	/	Tumor	Tumor	Tumor	/	Tumor
P10909	Clusterin	1.43	Tumor	3.97	Tumor	2.97	Tumor	1.92	0.86
P02452	Collagen alpha 1 I chain*	0.82	Tumor	1.65	Tumor	3.60	Tumor	0.88	2.16
Q99715	Collagen alpha 1 XII chain*	Tumor	Tumor	Tumor	Tumor	Tumor	Tumor	/	/
Q96C19	EF hand domain containing protein D2	Tumor	Tumor	Tumor	Tumor	Tumor	Tumor	Tumor	/
Q9Y6C2	EMILIN 1*	Tumor	Tumor	1.92	Tumor	Tumor	Tumor	/	Tumor
P02751	Fibronectin*	Tumor	Tumor	14.73	Tumor	4.10	Tumor	Tumor	Tumor
P23142	Fibulin 1	Tumor	Tumor	2.69	Tumor	Tumor	Tumor	/	Tumor
Q08380	Galectin 3 binding protein	Tumor	Tumor	3.90	Tumor	Tumor	Tumor	4.57	0.96
P16188	HLA class I histocompatibility antigen A 30 alpha chain	/	Tumor	/	Tumor	Tumor	Tumor	/	Tumor
P13747	HLA class I histocompatibility antigen alpha chain E	/	/	Tumor	Tumor	Tumor	Tumor	Tumor	/
P17693	HLA class I histocompatibility antigen alpha chain G	/	/	Tumor	Tumor	Tumor	Tumor	/	Tumor
P18465	HLA class I histocompatibility antigen B 57 alpha chain	Tumor	Tumor	Tumor	Tumor	Tumor	Tumor	Tumor	24.05

P30499	HLA class I histocompatibility antigen Cw 1 alpha chain	Tumor	/	Tumor	Tumor	/
Q07000	HLA class I histocompatibility antigen Cw 15 alpha chain	/	Tumor	Tumor	Tumor	Tumor
Q29960	HLA class I histocompatibility antigen Cw 16 alpha chain	Tumor	/	Tumor	Tumor	/
P30504	HLA class I histocompatibility antigen Cw 4 alpha chain	Tumor	Tumor	Tumor	Tumor	Tumor
P01906	HLA class II histocompatibility antigen DQ alpha 2 chain	/	Tumor	Tumor	Tumor	Tumor
P01903	HLA class II histocompatibility antigen DR alpha chain	/	Tumor	/	13.60	2.32
P05556	Integrin beta 1	1.51	1.54	Tumor	Tumor	1.46
Q86UP2	Kinectin	Tumor	Tumor	Tumor	/	Tumor
P33241	Lymphocyte specific protein 1	/	Tumor	Tumor	Tumor	/
P14174	Macrophage migration inhibitory factor	Tumor	Tumor	Tumor	/	2.05
O15173	Membrane associated progesterone receptor component 2	Tumor	1.67	Tumor	/	/
P26038	Moesin	Tumor	Tumor	Tumor	Tumor	1.06
P29966	Myristoylated alanine rich C kinase substrate*	Tumor	Tumor	/	Tumor	Normal
O14745	Na H exchange regulatory cofactor NHE RF1*	Tumor	Tumor	Tumor	Tumor	/
Q15063	Periostin*	4.71	24.78	5.75	Tumor	Tumor
P36955	Pigment epithelium derived factor*	Tumor	Tumor	1.30	/	/
Q15084	Protein disulfide isomerase A6*	/	/	Tumor	Tumor	Tumor
P07237	Protein disulfide isomerase*	7.92	3.86	14.15	12.94	2.36
P01893	Putative HLA class I histocompatibility antigen alpha chain H	Tumor	/	Tumor	Tumor	/
P62834	Ras related protein Rap 1A	Tumor	Tumor	/	Tumor	/
P61224	Ras related protein Rap 1b	Tumor	Tumor	/	Tumor	Normal
Q6FHJ7	Secreted frizzled related protein 4	Tumor	Tumor	Tumor	/	/

P05026	Sodium potassium transporting ATPase subunit beta 1	Tumor	Tumor	Tumor	/	/	/
P24821	Tenascin*	Tumor	Tumor	Tumor	Tumor	/	/
P10599	Thioredoxin	2.56	Tumor	Tumor	6.11	2.14	/
P35442	Thrombospondin 2*	Tumor	Tumor	Tumor	/	/	/
P37802	Transgelin 2*	3.71	3.56	9.21	13.07	2.10	Tumor
P13611	Versican core protein*	Tumor	Tumor	Tumor	/	Tumor	/
Q9P0L0	Vesicle associated membrane protein associated protein A*	Tumor	/	Tumor	Tumor	/	/
P21796	Voltage dependent anion selective channel protein 1*	Tumor	Tumor	Tumor	/	/	/

Figure 6B (table over previous 3 pages)

Accession	Protein name	Patient				
		BPSCC10/39	BPSCC10/40	BPSCC10/46	BPSCC10/47	BPSCC10/49
		Ratio T/N	Ratio T/N	Ratio T/N	Ratio T/N	Ratio T/N
Q01518	Adenylyl cyclase associated protein 1*	2.44	2.00	Tumor	3.13	1.84
Q8IUX7	Adipocyte enhancer binding protein 1*	Tumor	Tumor	Tumor	Tumor	/
P80723	Brain acid soluble protein 1	Tumor	Tumor	Tumor	Tumor	/
Q99439	Calponin 2*	Tumor	Tumor	Tumor	/	/
Q15417	Calponin 3*	Tumor	Tumor	Tumor	/	/
P27797	Calreticulin*	Tumor	4.02	Tumor	Tumor	Tumor
O43852	Calumenin	/	Tumor	Tumor	Tumor	Tumor
P13987	CD59 glycoprotein	2.23	/	Tumor	/	2.08
P02452	Collagen alpha 1 I chain*	1.51	2.72	1.63	2.16	1.57
Q99715	Collagen alpha 1 XII chain*	Tumor	23.40	Tumor	Tumor	Tumor
P08123	Collagen alpha 2 I chain	2.75	2.42	0.93	2.18	1.88
Q13561	Dynactin subunit 2	Tumor	Tumor	Tumor	Tumor	/
Q9Y6C2	EMILIN 1*	Tumor	Tumor	Tumor	Tumor	/
P02751	Fibronectin*	18.17	18.42	2.27	4.39	2.77
P09382	Galectin 1	3.16	7.56	Tumor	15.33	1.42
P13746	HLA class I histocompatibility antigen A 11 alpha chain	/	Tumor	Tumor	Tumor	/
P01903	HLA class II histocompatibility antigen DR alpha chain	/	Tumor	Tumor	Tumor	/
P01911	HLA class II histocompatibility antigen DRB1 15 beta chain	/	Tumor	Tumor	Tumor	/
P29966	Myristoylated alanine rich C kinase substrate*	Tumor	Tumor	Tumor	Tumor	/
O14745	Na H exchange regulatory cofactor NHE RF1*	Tumor	Tumor	Tumor	Tumor	/
Q15063	Periostin*	Tumor	Tumor	Tumor	Tumor	Tumor
P36955	Pigment epithelium derived factor*	Tumor	Tumor	6.23	/	/
P13796	Plastin 2	Tumor	1.22	/	Tumor	Tumor

SUBSTITUTE SHEET (RULE 26)

Q15084	Protein disulfide isomerase A6*	2.92	4.16	Tumor	8.00	2.08
P07237	Protein disulfide isomerase*	4.53	Tumor	Tumor	13.20	2.27
Q92928	Putative Ras related protein Rab 1C	/	Tumor	Tumor	Tumor	/
P50395	Rab GDP dissociation inhibitor beta	Tumor	1.48	Tumor	/	Tumor
P46940	Ras GTPase activating like protein IQGAP1	2.64	3.98	Tumor	Tumor	Normal
P61106	Ras related protein Rab 14	/	Tumor	Tumor	Tumor	/
P08134	Rho related GTP binding protein RhoC	/	Tumor	Tumor	Tumor	/
P05023	Sodium potassium transporting ATPase subunit alpha 1	Tumor	1.06	Tumor	/	Tumor
P24821	Tenascin*	Tumor	Tumor	Tumor	Tumor	/
P07996	Thrombospondin 1*	Tumor	Tumor	Tumor	/	/
P35442	Thrombospondin 2*	Tumor	Tumor	Tumor	/	/
Q15582	Transforming growth factor beta induced protein ig h3	Tumor	2.70	/	Tumor	/
P61586	Transforming protein RhoA	Tumor	3.36	Tumor	3.19	/
P37802	Transgelin 2*	Tumor	7.34	Tumor	Tumor	/
P13611	Versican core protein*	Tumor	8.20	Tumor	Tumor	Tumor
Q9P0L0	Vesicle associated membrane protein associated protein A*	Tumor	/	Tumor	Tumor	Tumor
P18206	Vinculin	0.68	1.78	Tumor	2.03	1.09
P21796	Voltage dependent anion selective channel protein 1*	/	Tumor	Tumor	Tumor	Tumor

Figure 6C (table over previous two pages)

SUBSTITUTE SHEET (RULE 26)

Accession	Protein name	Patient					
		BPSCC10/39	BPSCC10/40	BPSCC10/46	BPSCC10/47	BPSCC10/49	
		Ratio T/N	Ratio T/N	Ratio T/N	Ratio T/N	Ratio T/N	
P21810	Biglycan	1.6	4.66	7.17	2.89	3.42	
Q10589	Bone marrow stromal antigen 2	Normal	Tumor	Tumor	Tumor	/	
Q9UBG0	C type mannose receptor 2	1.11	1.68	/	Tumor	Tumor	
Q5ZPR3	CD276 antigen	3.00	Tumor	Tumor	Tumor	/	
Q9BY67	Cell adhesion molecule 1	Tumor	7.77	Tumor	Tumor	/	
Q99715	Collagen alpha 1 XII chain*	Tumor	22.87	2.97	1.42	Tumor	
P0C0L5	Complement C4 B	Tumor	Tumor	Tumor	Tumor	/	
Q9Y6C2	EMILIN 1*	1.82	2.89	Tumor	Tumor	2.86	
Q08380	Galectin 3 binding protein	0.87	4.39	2.03	3.42	0.44	
P02790	Hemopexin	1.12	2.51	Tumor	Tumor	7.24	
P12314	High affinity immunoglobulin gamma Fc receptor I	/	Tumor	Tumor	Tumor	Normal	
P01903	HLA class II histocompatibility antigen DR alpha chain	3.63	15.33	Tumor	Tumor	2.36	
Q9NX62	Inositol monophosphatase 3	Tumor	Tumor	Tumor	Tumor	Tumor	
P06756	Integrin alpha V	1.62	4.10	11.13	5.10	/	
P05107	Integrin beta 2	/	Tumor	Tumor	Tumor	/	
P11279	Lysosome associated membrane glycoprotein 1	2.46	2.72	2.12	4.18	1.02	
P13473	Lysosome associated membrane glycoprotein 2	Tumor	3.67	11.13	19.69	2.77	
P01033	Metalloproteinase inhibitor 1	0.25	Tumor	Tumor	Tumor	0.89	
Q15063	Periostin*	3.74	Tumor	Tumor	Tumor	Tumor	
P36955	Pigment epithelium derived factor*	1.92	5.87	1.58	1.02	2.08	
P05154	Plasma serine protease inhibitor	Tumor	Tumor	2.23	0.94	1.48	
P24821	Tenascin*	Tumor	Tumor	Tumor	Tumor	Tumor	
P07996	Thrombospondin 1*	/	Tumor	Tumor	Tumor	/	
P04216	Thy 1 membrane glycoprotein	1.23	8.85	8.58	5.37	2.72	

P02786 Q9HD45	Transferrin receptor protein 1 Transmembrane 9 superfamily member 3	Tumor 1.73	/ 2.80	Tumor Tumor	Tumor /	Tumor /

Figure 6D (table over previous two pages)

Accession	Gene	Protein name	N	T	emPAI T/N	Subcellular location	Biological process	Biological function
P02751	FN1	Fibronectin	1	3	1.19	E	GM	angiogenesis; cell migration; collagen binding
P20774	OGN	Mimecan	2	3	2.19	E	Cc; St	growth factor activity
Q15063	POSTN	Periostin	2	3	5.63	E	Cc; St	cell adhesion; heparin binding; protein binding
P50395	GDI2	Rab GDP dissociation inhibitor beta	2	3	2.06	Cy; C	T	protein binding; protein transport; signal transduction
P12111	COL6A3	Collagen alpha-3(VI) chain	3	3	2.33	E	GM	cell adhesion
Q9BXN1	ASPN	Asporin	0	2		E	GM	bone mineralization; protein binding
P07384	CAPN1	Calpain-1[1]	0	2		Cy; C	PM	cell proliferation; proteolysis
P23946	CMA1	Chymase	0	2		S	PM	proteolysis; regulation of inflammatory response
P08572	COL4A2	Collagen[3]	0	2		E	GM	angiogenesis; collagen type IV; protein binding
P16444	DPEP1	Dipeptidase 1	0	2		C	PM	protein binding; proteolysis; zinc ion binding
P08294	SOD3	Extracellular superoxide dismutase	0	2		E	M; Ep	heparin binding; oxidation reduction
P17931	LGALS3	Galectin[5]	0	2		N; CS ; Cy; C	R	sugar binding
P22352	GPX3	Glutathione peroxidase 3	0	2		S	M; Ep	glutathione peroxidase activity; transcription factor binding
P04899	GNAI2	Guanine nucleotide-binding protein G[2]	0	2		C; Cy; ER ; G	Cc; St	protein binding; signal transducer

Q00839	HNRNPU	Heterogeneous nuclear ribonucleoprotein U	0	2		N; CS	R	protein binding
Q16363	LAMA4	Laminin subunit alpha-4	0	2		E	GM	cell adhesion; cell migration
Q14767	LTBP2	Latent-transforming growth factor beta-binding protein 2	0	2		S	GM	calcium ion binding; growth factor binding
Q6UWY5	OLFML1	Olfactomedin-like protein 1	0	2		S	U	Unknown
P60903	S100A10	Protein S100-A10	0	2		C; N; Cy	Cc; St	calcium ion binding; receptor binding; signal transduction
Accession	**Gene**	**Protein name**	**N**	**T**	**emPAI T/N**	**Subcellular location**	**Biological process**	**Biological function**
Q13733	ATP1A4	Sodium/potassium-transporting ATPase subunit alpha-4	0	2		C	T	regulation of cellular pH
P35442	THBS2	Thrombospondin[7]	0	2		E	GM	calcium ion binding; cell adhesion; heparin binding; protein binding
Q15582	TGFBI	Transforming growth factor-beta-induced protein ig-h3	0	2		E	Cc; St	cell adhesion; integrin binding
P38606	ATP6V1A	V-type proton ATPase catalytic subunit A	0	2		C	M; Ep	proton-transporting V-type ATPase
Q01518	CAP1	Adenylyl cyclase-associated protein 1	1	2	3.08	C	GM	signal transduction
P46940	IQGAP1	Ras GTPase-activating-like protein IQGAP1	1	2	2.75	C	Cc; St	protein phosphatase binding
P07585	DCN	Decorin	2	2	2.17	E	GM	unknown
P51888	PRELP	Prolargin	2	2	10.36	E	GM	protein binding
P21980	TGM2	Protein-glutamine gamma-glutamyltransferase 2	2	2	8.50	E; C; Cy; N	M; Ep	calcium ion binding; cell adhesion
P08253	MMP2	72 kDa type IV collagenase	0	1		E	PM	calcium ion binding; collagen catabolic;

Q10588	BST1	ADP-ribosyl cyclase 2	0	1	C	I	proteolysis
O00468	AGRN	Agrin	0	1	E	GM	unknown laminin binding; signal transduction
P05186	ALPL	Alkaline phosphatase, tissue-nonspecific isozyme	0	1	C	M; Ep	alkaline phosphatase activity
P04745	AMY1A	Alpha-amylase 1	0	1	S	M	protein binding
P01019	AGT	Angiotensinogen	0	1	S	Cc; St	blood vessel remodeling
P20073	ANXA7	Annexin A7	0	1	N; C; ER; Cy	T	calcium ion binding
P80723	BASPN1	Brain acid soluble protein 1	0	1	C	R	promoter binding
P16870	CPE	Carboxypeptidase E	0	1	S	PM	protein modification process
P49961	ENTPD1	Ectonucleoside triphosphate diphosphohydrolase 1	0	1	C	M; Ep	calcium ion binding; cell adhesion
Q12805	EFEMP1	EGF-containing fibulin-like extracellular matrix protein 1	0	1	S	GM	calcium ion binding; protein binding
Q9Y6I3	EPN1	Epsin-1	0	1	Cy; C; N	Cc; St	Lipid binding
Accession	**Gene**	**Protein name**	**N**	**T**	**emPAI T/N**	**Subcellular location**	**Biological process**	**Biological function**
P23142	FBLN1	Fibulin[4]	0	1		E	GM	calcium ion binding; protein binding
P07093	SERPINE2	Glia-derived nexin	0	1		E	PM	Heparin binding
P01906	HLA-DQA2	HLA class II histocompatibility antigen[6]	0	1		C	I	immune response
P56199	ITGA1	Integrin alpha-1	0	1		C	Cc; St	calcium ion binding; cell-matrix adhesion
P24043	LAMA2	Laminin[8]	0	1		E	GM	cell adhesion
Q13449	LSAMP	Limbic system-associated membrane protein	0	1		C	Cc; St	cell adhesion; protein binding
Q93052	LPP	Lipoma-preferred partner	0	1		N; Cy; C	GM	cell adhesion

P01033	TIMP1	Metalloproteinase inhibitor 1	0	1	S	GM	protein binding	
P55083	MFAP4	Microfibril-associated glycoprotein 4	0	1	E	GM	calcium ion binding; cell adhesion; receptor binding; signal transduction	
Q9NZM1	MYOF	Myoferlin	0	1	C	GM	protein binding	
P14543	NID1	Nidogen-1	0	1	E	GM	protein binding	
Q9NRN5	OLFML3	Olfactomedin-like protein 3	0	1	S	GM	multicelular organismal development	
P54315	PNLIPRP1	Pancreatic lipase-related protein 1	0	1	S	M; Ep	lipid catabolic process	
Q13492	PICALM	Phosphatidylinositol-binding clathrin assembly protein	0	1	C	U	unknown	
Q7Z5L7	PODN	Podocan	0	1	E	U	protein binding	
P54707	ATP12A	Potassium-transporting ATPase alpha chain 2	0	1	C	T	proton transport	
Q07954	LRP1	Prolow-density lipoprotein receptor-related protein 1	0	1	C	Cc; St	apolipoprotein bindin	
P61224	RAP1B	Ras-related protein Rap-1b	0	1	C	Cc; St	protein binding	
O43566	RGS14	Regulator of G-protein signaling 14	0	1	C	Cc; St	signal transduction	
Q9UHD8	SEPT9	Septin-9	0	1	C	P	protein binding polysaccharide binding; protein binding	
Q9UGT4	SUSD2	Sushi domain-containing protein 2	0	1	C	U	unknown	

Accession	Gene	Protein name	N	T	emPAI T/N	Subcellular location	Biological process	Biological function
O60784	TOM1	Target of Myb protein 1	0	1		Cy; C	Cc; St	protein binding
P24821	TNC	Tenascin[9]	0	1		E	Cc; St	cell adhesion; receptor binding
Q0P6H9	TMEM62	Transmembrane protein 62	0	1		C	U	unknown

P45974	USP5	Ubiquitin carboxyl-terminal hydrolase 5	0	1		E	PM	T	protein binding; ubiquitin thiolesterase activity
P02730	SLC4A1	Band 3 anion transport protein	1	1	2.00	C			anion transport
Q92597	NDRG1	Protein NDRG1	1	1	2.00	Cy; C; N	Cc; St		protein binding; response to metal ion

Figure 7 (table over previous 5 pages)

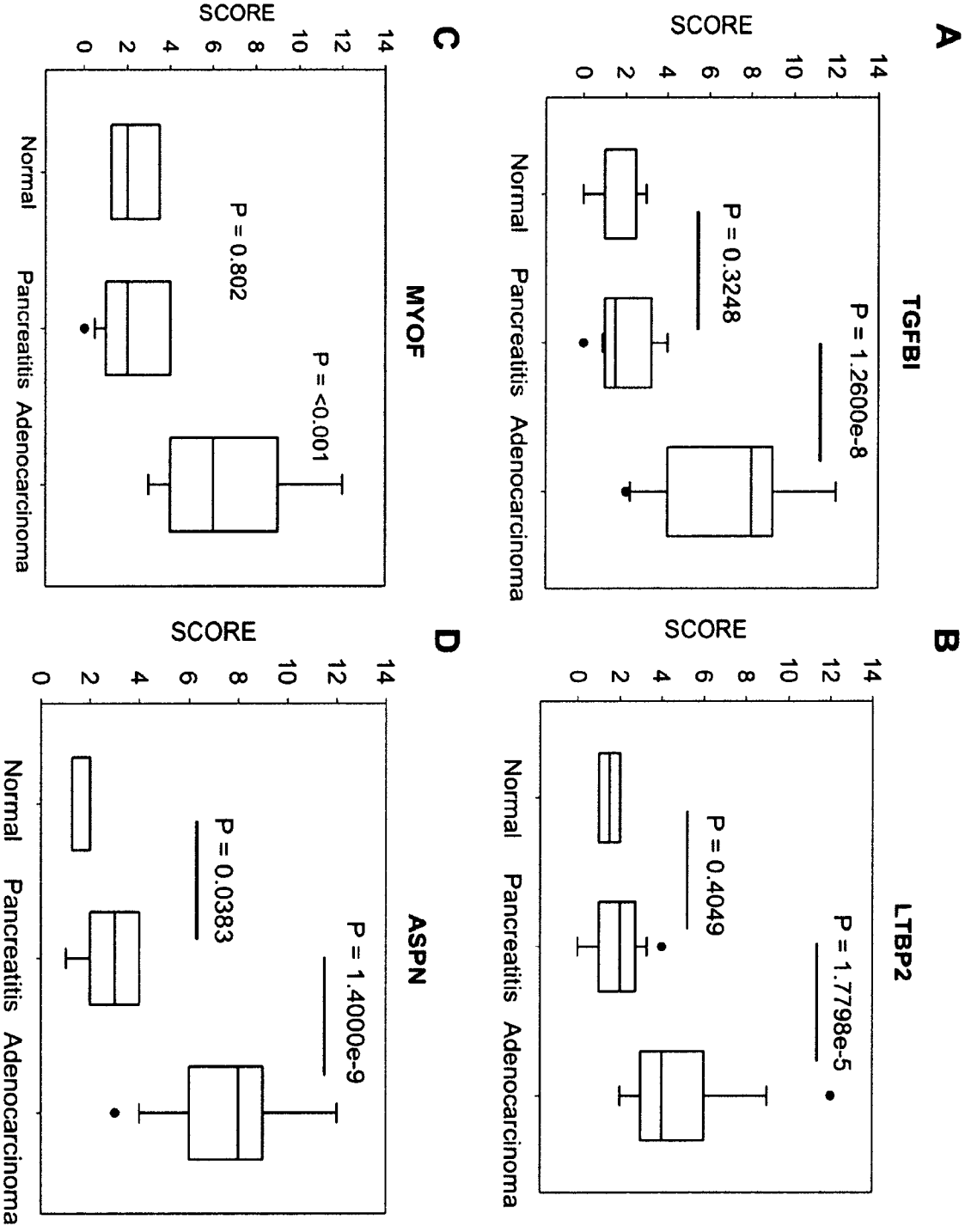
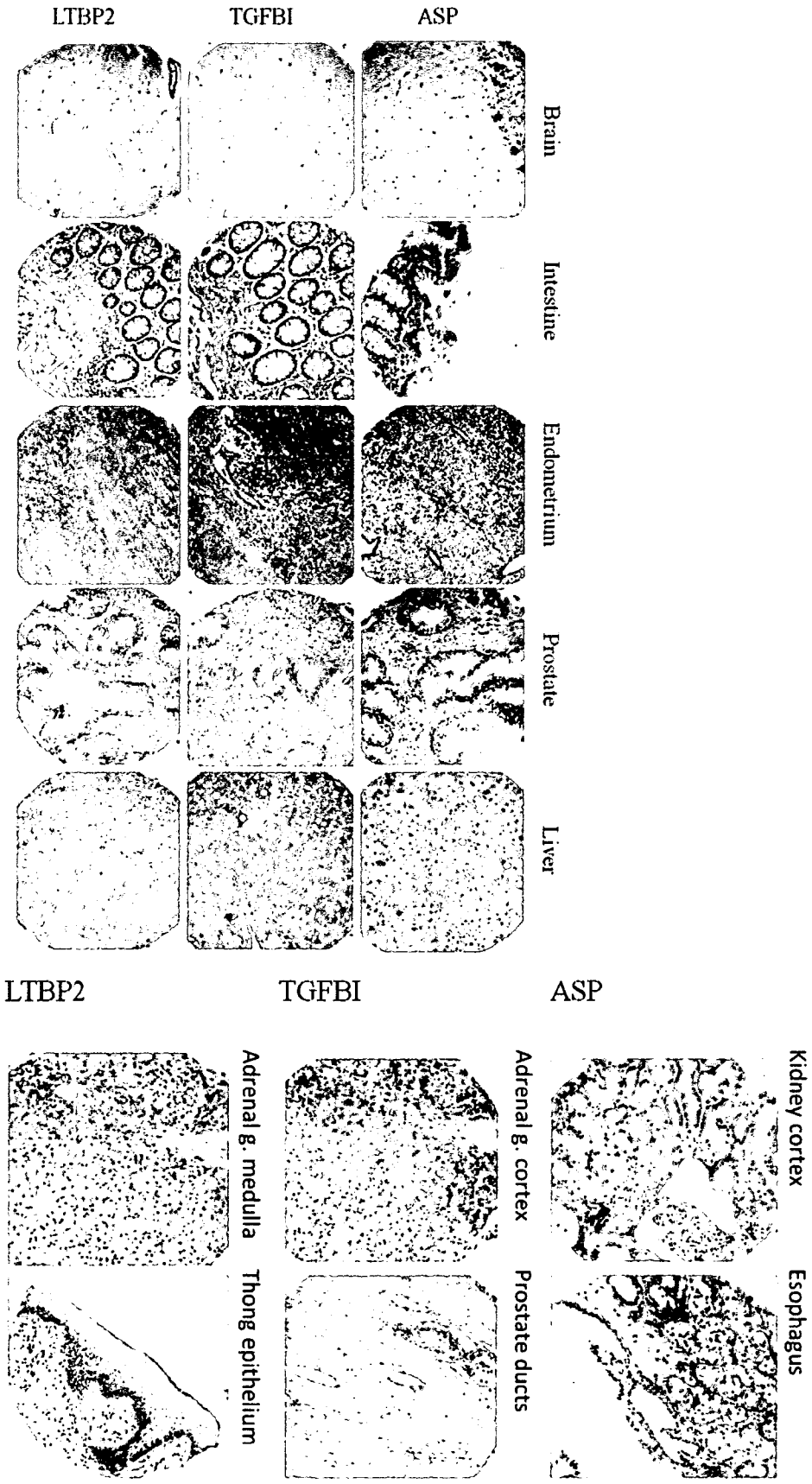

A
TGFBI
P = 1.2600e-8
P = 0.3248
SCORE
Normal
Pancreatitis
Adenocarcinoma
B
LTBP2
P = 1.7798e-5
P = 0.4049
C
MYOF
P = <0.001
P = 0.802
D
ASPN
P = 1.4000e-9
P = 0.0383

Figure 8

Figure 9


LTBP2
TGFBI
ASP
Brain
Intestine
Endometrium
Prostate
Liver
LTBP2
TGFBI
ASP
Adrenal g. medulla
Adrenal g. cortex
Kidney cortex
Thong epithelium
Prostate ducts
Esophagus

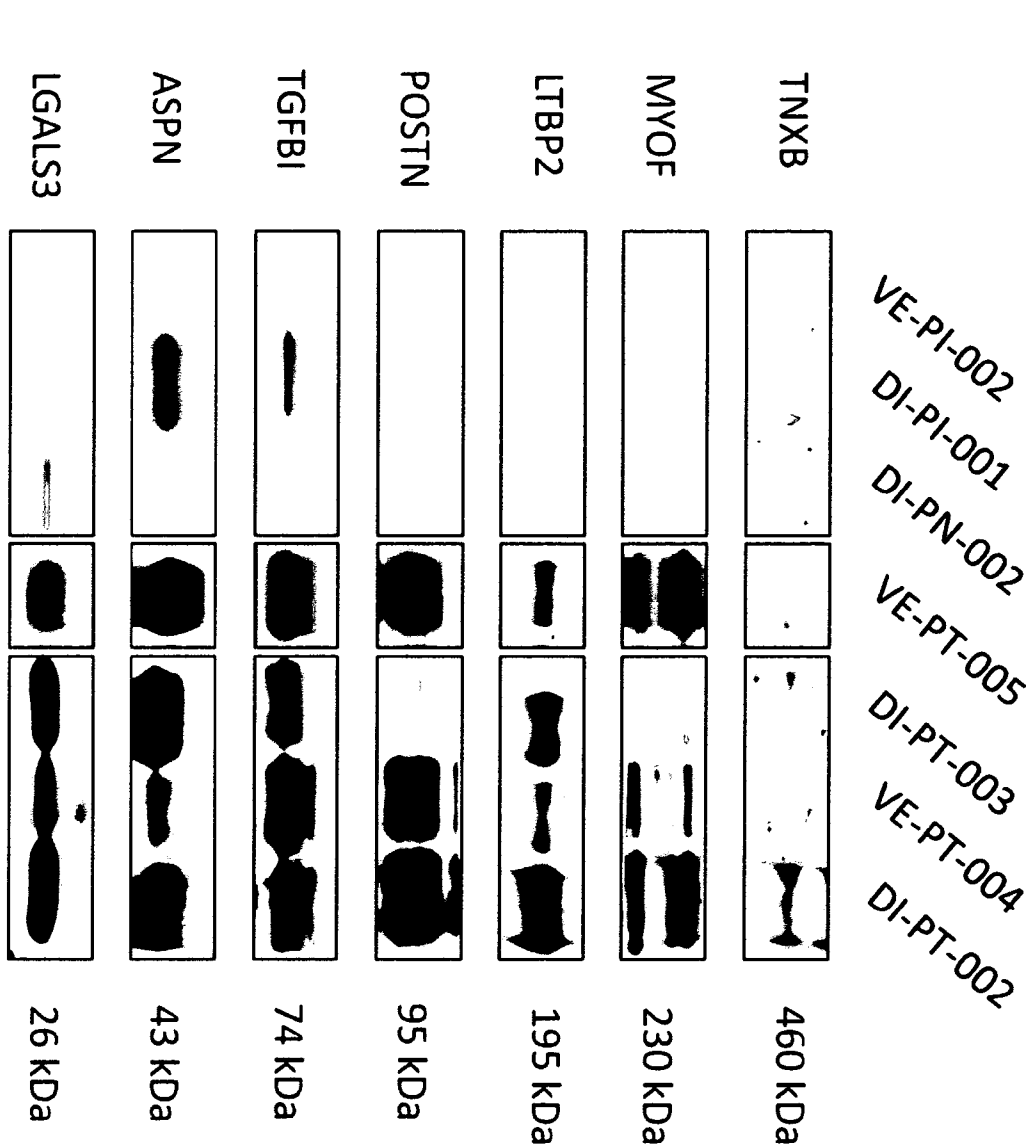

TNXB
MYOF
LTBP2
POSTN
TGFBI
ASPN
LGALS3
460 kDa
230 kDa
195 kDa
95 kDa
74 kDa
43 kDa
26 kDa
VE-PI-002
DI-PI-001
DI-PN-002
VE-PT-005
DI-PT-003
VE-PT-004
DI-PT-002

Figure 10

Figure 11

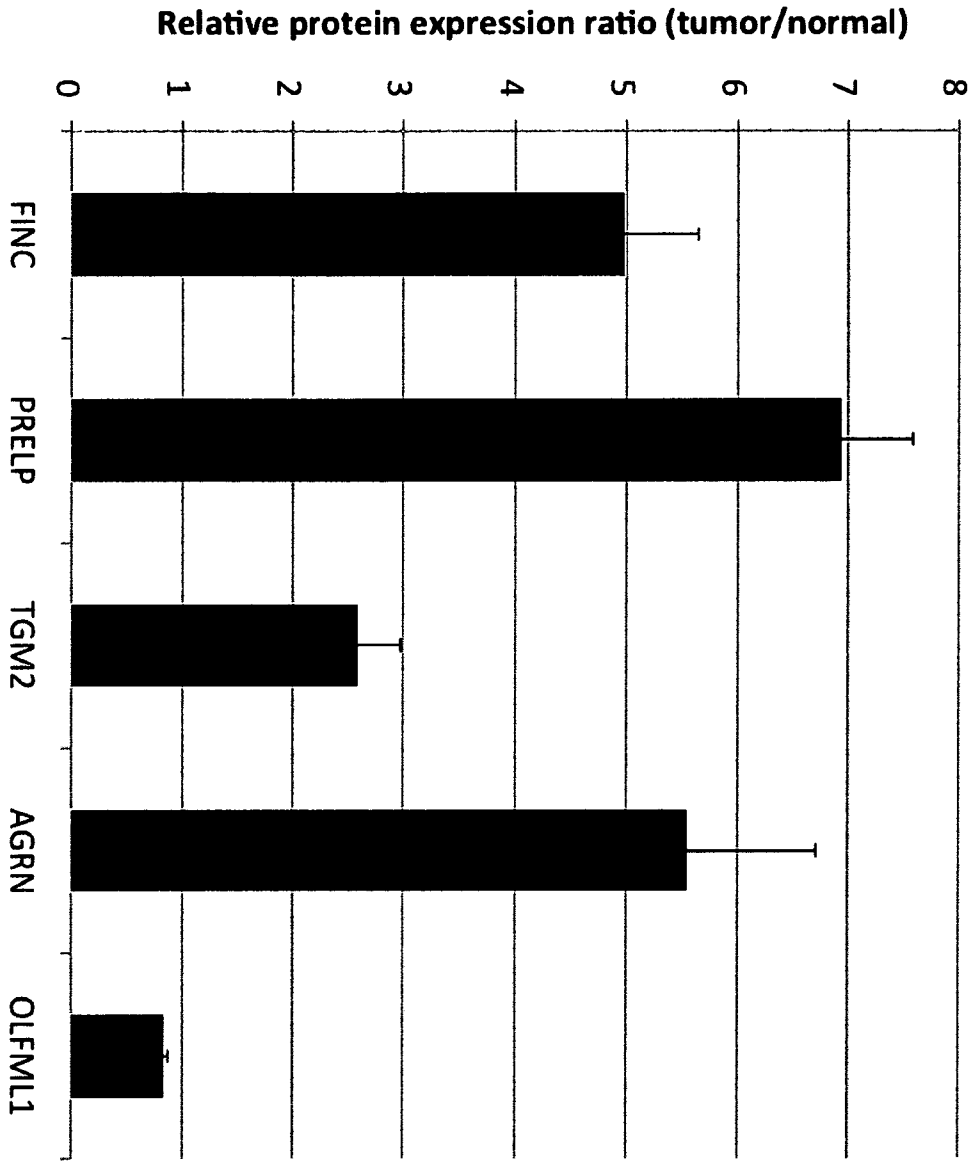
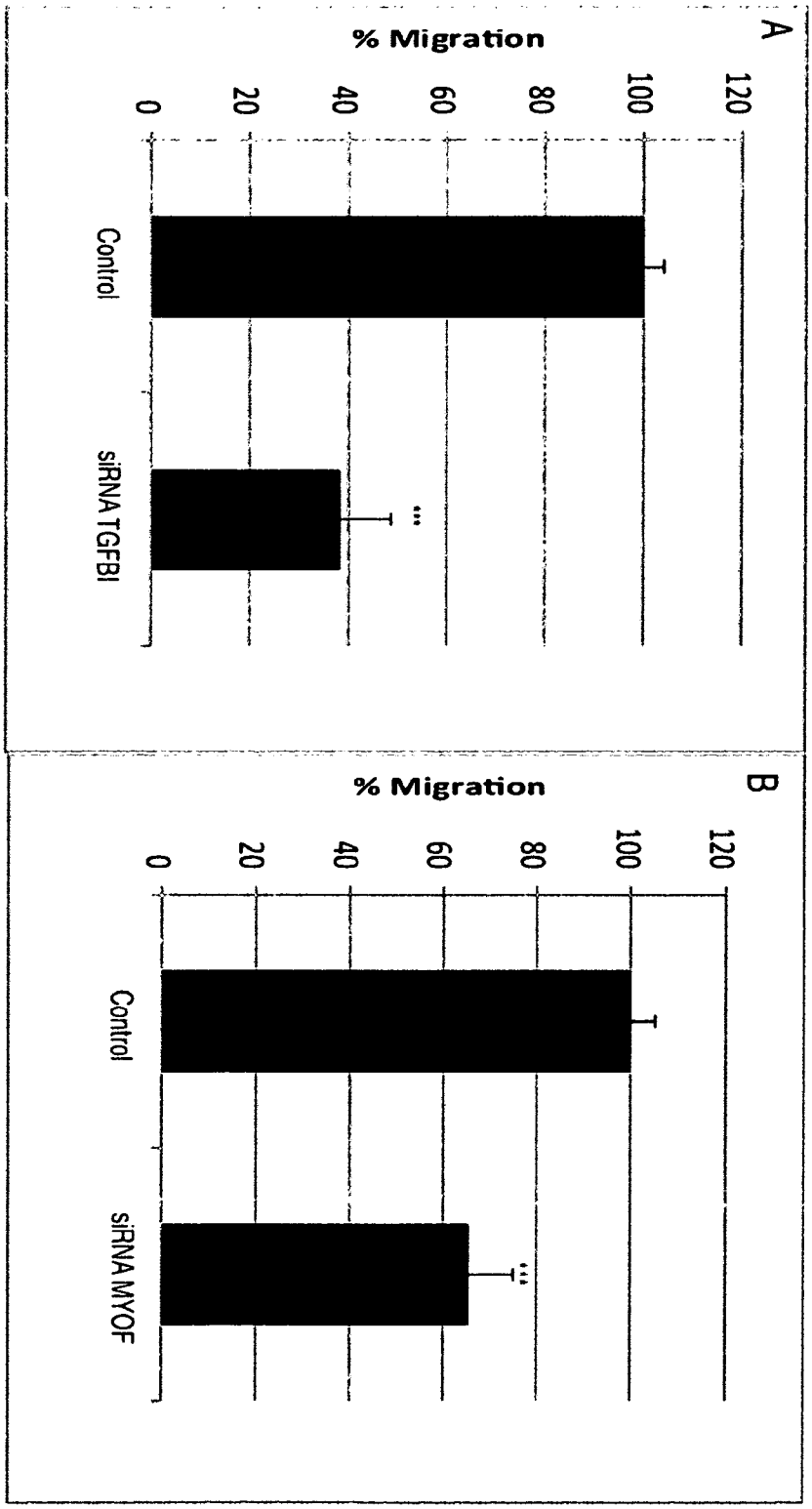

Relative protein expression ratio (tumor/normal)
0
1
2
3
4
5
6
7
8
FINC
PRELP
TGM2
AGRN
OLFML1

Figure 12


A
% Migration
0
20
40
60
80
100
120
Control
siRNA TGFBI
***
B
% Migration
0
20
40
60
80
100
120
Control
siRNA MYOF
***

Accession	Protein name	Score	Normal/Meta		Number of patients
			Ratio	P-Value	
Q01518	Adenylyl cyclase associated protein 1	2700.75	0.63	0.14	3
Q8IUX7	Adipocyte enhancer binding protein 1	2275.47	M	M	3
P04083	Annexin A1	3969.34	0.43	0.00	3
P07355	Annexin A2	7874.84	0.46	0.00	3
O95994	Anterior gradient protein 2 homolog	2438.31	M	M	2
Q8TD06	Anterior gradient protein 3 homolog	229.13	M	M	2
Q9BXN1	Asporin	174.17	M	M	2
P35613	Basigin	471.13	0.46	0.02	2
P61769	Beta 2 microglobulin	2303.87	0.41	0.03	2
P21810	Biglycan	8529.73	M	M	3
Q12864	Cadherin 17	268.25	M	M	2
P62158	Calmodulin	749.92	M	M	2
O43852	Calumenin	1612.65	0.38	0.02	2
P13688	Carcinoembryonic antigen related cell adhesion molecule 1	254.74	M	M	2
P06731	Carcinoembryonic antigen related cell adhesion molecule 5	2880.34	M	M	3
P40199	Carcinoembryonic antigen related cell adhesion molecule 6	1970.86	M	M	2
P16070	CD44 antigen	372.00	0.24	0.00	2
P48960	CD97 antigen	262.92	M	M	2
O00299	Chloride intracellular channel protein 1	3340.75	0.40	0.06	3
P09496	Clathrin light chain A	251.26	M	M	2
O15551	Claudin 3	895.78	M	M	2
P10909	Clusterin	3653.82	0.25	0.00	3

Accession	Protein name	Score	Normal/Meta		Number of patients
			Ratio	P-Value	
P02452	Collagen alpha 1 I chain	1713.56	0.25	0.09	3
Q99715	Collagen alpha 1 XII chain	272.24	M	M	2
Q05707	Collagen alpha 1 XIV chain	557.36	M	M	3
P08123	Collagen alpha 2 I chain	1390.79	0.48	0.00	3
P12110	Collagen alpha 2 VI chain	483.60	M	M	3
P07585	Decorin	326.30	M	M	2
Q9UGM3	Deleted in malignant brain tumors 1 protein	738.47	M	M	2
Q14126	Desmoglein 2	238.32	M	M	2
Q96C19	EF hand domain containing protein D2	729.65	M	M	2
Q9Y6C2	EMILIN 1	629.65	M	M	3
P16422	Epithelial cell adhesion molecule	1975.73	M	M	3
P02751	Fibronectin	2864.98	0.14	0.00	3
P09382	Galectin 1	1997.67	0.46	0.02	3
P17931	Galectin 3	2200.01	M	M	2
P28799	Granulins	222.46	M	M	2
P59768	Guanine nucleotide binding protein G I G S G O subunit gamma 2	1855.16	M	M	2
P04196	Histidine rich glycoprotein	1967.68	0.39	0.01	3
P04439	HLA class I histocompatibility antigen A 3 alpha chain	3257.02	M	M	2
P10314	HLA class I histocompatibility antigen A 32 alpha chain	653.82	M	M	2
Q09160	HLA class I histocompatibility antigen A 80 alpha chain	883.00	M	M	3

SUBSTITUTE SHEET (RULE 26)

Q07000	HLA class I histocompatibility antigen Cw 15 alpha chain	2512.69	0.23	0.05	3
					2
P01861	Ig gamma 4 chain C region	1367.27	M	M	

Accession	Protein name	Score	Normal/Meta		Number of patients
			Ratio	P-Value	
P11215	Integrin alpha M	311.67	M	M	2
P05556	Integrin beta 1	904.29	0.73	0.18	3
Q16363	Laminin subunit alpha 4	253.37	M	M	2
P07942	Laminin subunit beta 1	253.87	M	M	2
Q14766	Latent transforming growth factor beta binding protein 1	337.98	M	M	2
Q14767	Latent transforming growth factor beta binding protein 2	215.68	M	M	2
P51884	Lumican	10292.20	0.31	0.00	3
P33241	Lymphocyte specific protein 1	1470.87	0.27	0.00	3
P13473	Lysosome associated membrane glycoprotein 2	164.20	0.08	0.00	3
P20774	Mimecan	2782.97	0.43	0.00	3
Q9H3R2	Mucin 13	1433.17	M	M	2
P59665	Neutrophil defensin 1	4127.99	0.18	0.00	3
Q15063	Periostin	1699.05	M	M	3
P13796	Plastin 2	2334.51	0.81	0.21	3
P51888	Prolargin	11026.30	0.51	0.00	3
P02760	Protein AMBP	360.06	M	M	2
Q92597	Protein NDRG1	1008.31	0.14	0.00	2
P06703	Protein S100 A6	3181.63	0.43	0.00	3

Q92930	Ras related protein Rab 8B	1602.34	0.63	0.39	3
A6NIZ1	Ras related protein Rap 1b like protein	219.79	M	M	2
Q9Y3L5	Ras related protein Rap 2c	271.38	M	M	2
P08134	Rho related GTP binding protein RhoC	1391.68	0.63	0.28	2
					2
P13637	Sodium potassium transporting ATPase subunit alpha 3	307.45	0.42	0.16	

Accession	**Protein name**	**Score**	**Normal/Meta**		**Number of patients**
			Ratio	**P-Value**	
P24821	Tenascin	722.08	M	M	3
P07996	Thrombospondin 1	453.08	M	M	3
P35442	Thrombospondin 2	770.58	M	M	3
P04216	Thy 1 membrane glycoprotein	1796.84	M	M	3
Q15582	Transforming growth factor beta induced protein ig h3	894.58	M	M	3
P37802	Transgelin 2	6892.45	0.40	0.00	3
Q9UI15	Transgelin 3	534.39	M	M	3
P02766	Transthyretin	7561.09	0.22	0.01	2
P04004	Vitronectin	5682.05	0.16	0.00	3

Figure 13 (table over previous 4 pages)

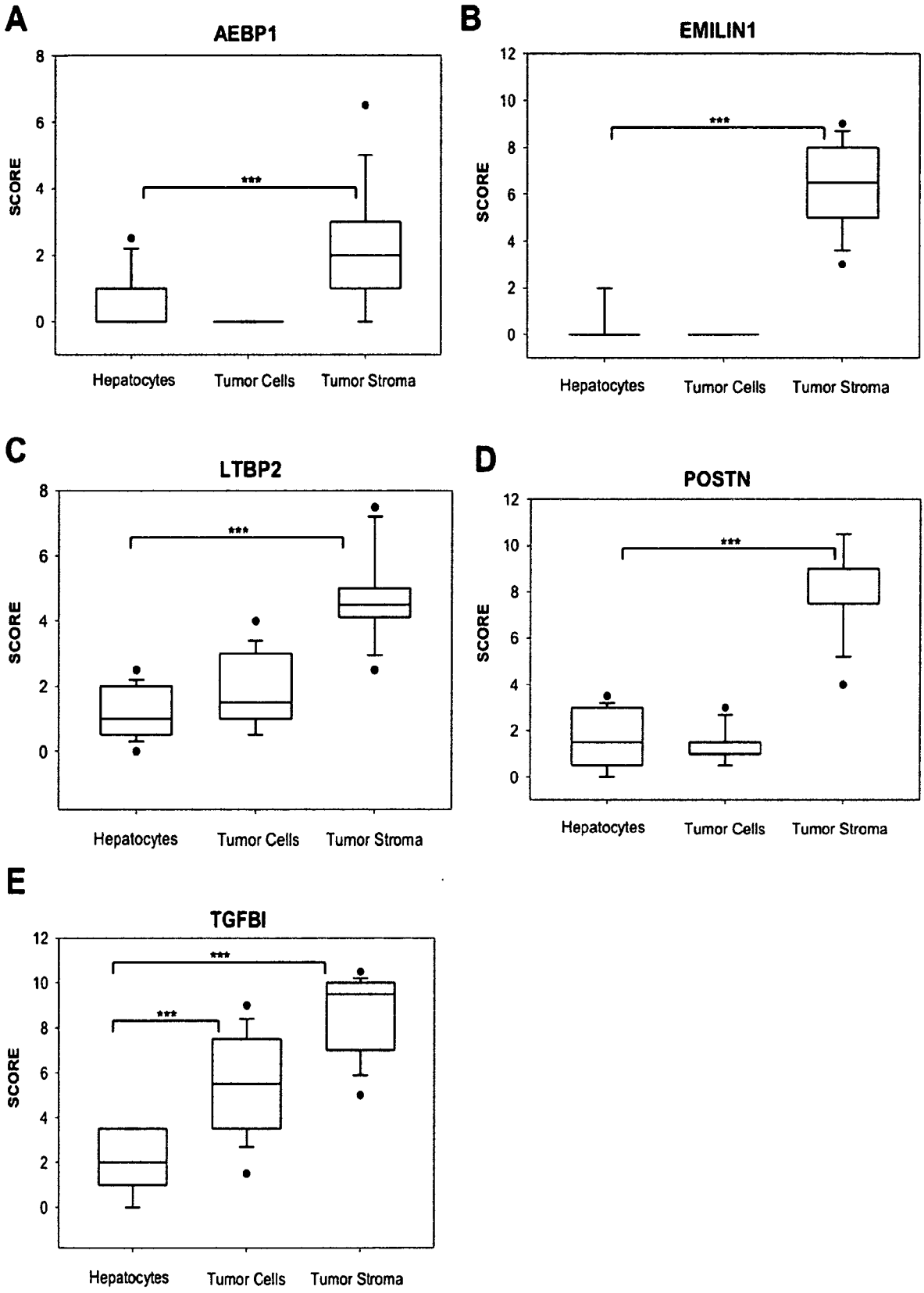

A
AEBP1
SCORE
0
2
4
6
8
***
Hepatocytes
Tumor Cells
Tumor Stroma
B
EMILIN1
SCORE
0
2
4
6
8
10
12
***
Hepatocytes
Tumor Cells
Tumor Stroma
C
LTBP2
SCORE
0
2
4
6
8
***
Hepatocytes
Tumor Cells
Tumor Stroma
D
POSTN
SCORE
0
2
4
6
8
10
12
***
Hepatocytes
Tumor Cells
Tumor Stroma
E
TGFBI
SCORE
0
2
4
6
8
10
12
***
***
Hepatocytes
Tumor Cells
Tumor Stroma

Figure 14

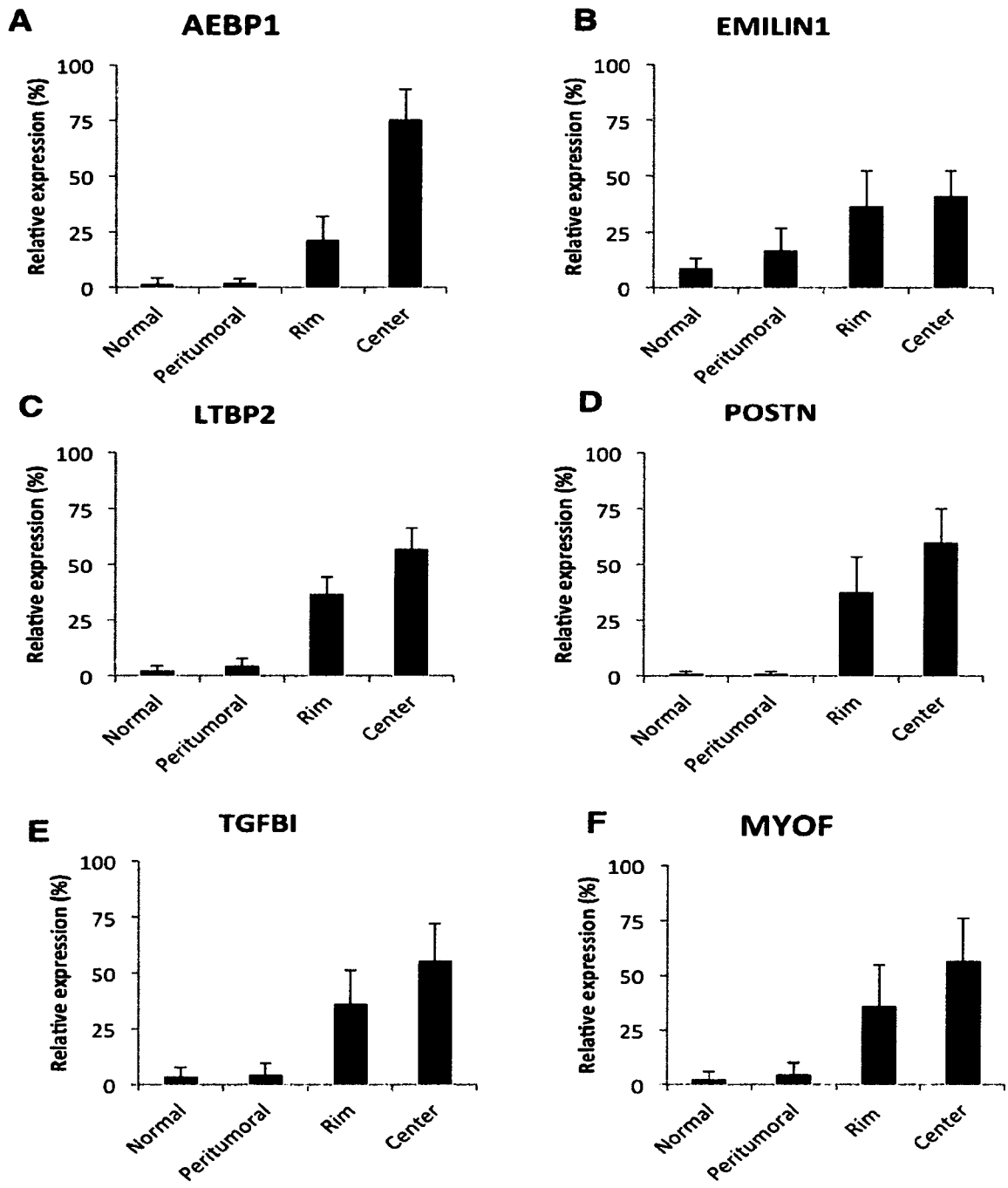

A
AEBP1
B
EMILIN1
C
LTBP2
D
POSTN
E
TGFBI
F
MYOF
Relative expression (%)
100
75
50
25
0
Normal
Peritumoral
Rim
Center

Figure 15

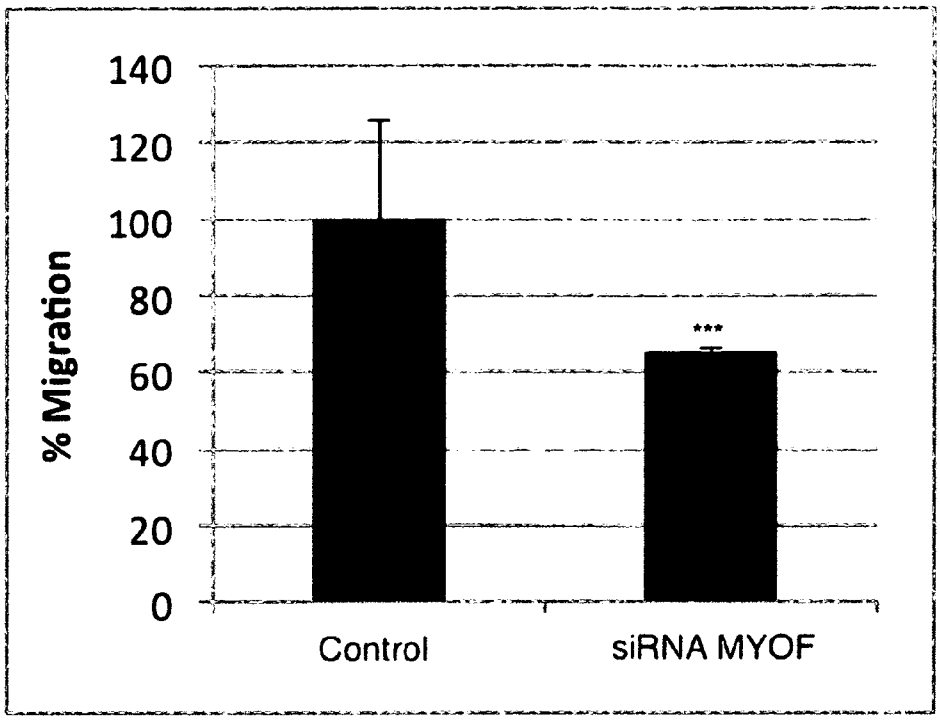


Figure 16

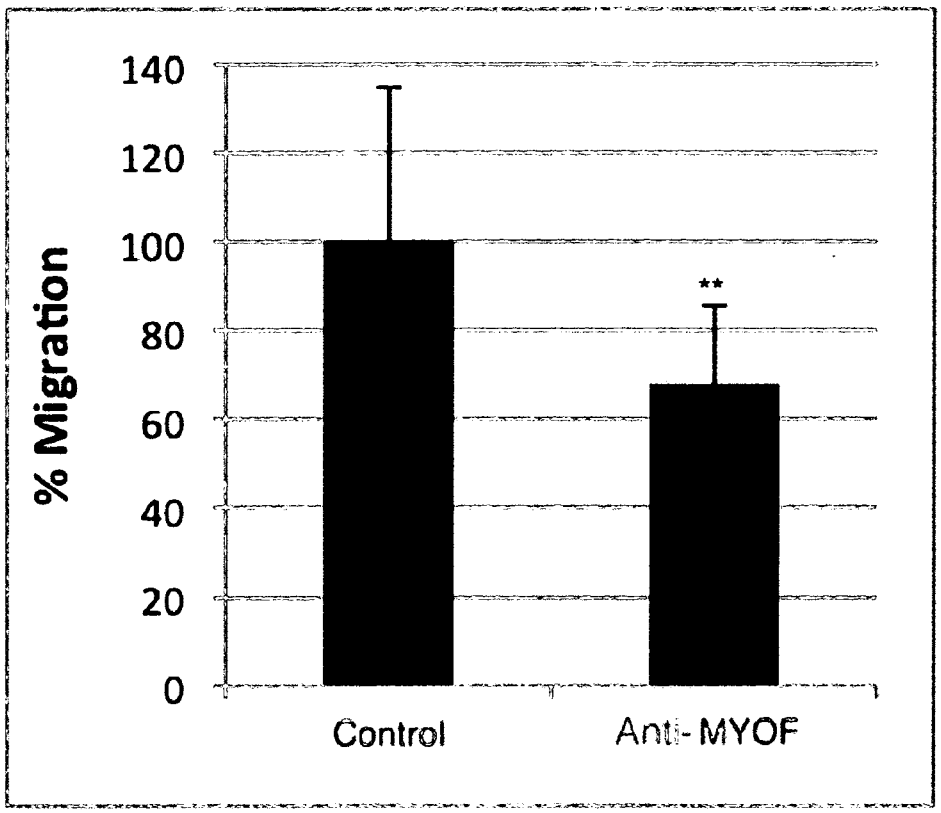


Figure 17

Accession	Protein name	Breast Tumor			Breast Tumor Bone Meta			Lymphoma		Pancreas Tumor			Glioblastoma		Liver Meta
		N	T	Ratio T/N	T	M	Ratio T/M	N	T	N	T	Ratio T/N	N	T	Ratio T/N
Q9BXN1	Asporin	8	8	1.20	8	1	1.06	0	1	0	2				tumor
O15230	Laminin alpha 5	1	3	1.00	0	1									
Q16363	Laminin alpha 4	4	4	0.44	1	3	1.00	3	4	0	2				tumor
P11047	Laminin gamma 1	5	4	0.74	2	5	0.05	3	5	2	1	0.67			tumor
Q14766	Latent transforming growth factor beta binding protein 1							0	1						tumor
Q14767	Latent transforming growth factor beta binding protein 2				2	0		0	2	0	2				
P20774	Mimecan	6	5	0.72	6	1		2	2	2	3	2.19	+		0.73
Q9UGT4	Sushi domain containing protein 2				5	0				0	1				
Q9NZM1	Myoferlin	0	3		1	0				0	1				
Q15582	Transforming growth factor-beta-induced protein ig-h3	6	8	1.03	7	2	0.12	2	5	0	2				6.82
O00468	Agrin	1	3	1.00	0	1		1	0	0	1			+	
P05556	Integrin beta 1	5	4	0.81	6	1	1.67	3	5	2	3	0.92		+	1.39
P16144	Integrin beta 4														tumor
P18564	Integrin beta 6							3	4					+	
P23229	Integrin alpha 6	1	0		1	0									tumor
P06756	Integrin alpha V	3	3	1.18	4	2	0.15	0	1					+	tumor
Q5ZPR3	CD276	0	2		1	0									
P05362	Intercellular adhesion molecule 1	3	2	0.95	2	0		3	4						tumor
Q6UWY5	Olfactomedin like protein 1	0	2							0	2				
Q9NRN5	Olfactomedin like protein 3	3	3	0.78	2	2	0.23			0	1				
P23142	Fibulin 1	3	3	1.24	5	1	2.75	0	2	0	1				
Q05707	Collagen A1 (XIV)	9	9	1.04	7	2	0.06	3	5	2	2	3.17			1.17

Accession	Protein name	Breast Tumor			Breast Tumor Bone Meta			Lymphoma		Pancreas Tumor			Glioblastoma		Liver Meta
		N	T	Ratio T/N	T	M	Ratio T/M	N	T	N	T	Ratio T/N	N	T	Ratio T/N
P12109	Collagen A1 (VI)	9	9	1.06	8	6	3.44	2	5	2	3	1.62		+	0.82
P12111	Collagen A3 (VI)	9	9	0.85	8	8	1.93	3	5	3	3	2.33		+	0.61
P08123	Collagen A2 (I)	5	5	1.00	5	4	1.50			1	0			+	3.32
P07355	Annexin A2	6	6	1.17	6	6	1.49	3	4	3	3	0.78		+	4.14
P08758	Annexin A5	5	5	1.11	6	5	2.80	3	5	3	3	0.53	+	+	0.54
P08133	Annexin A6	6	6		6	2	0.65	3	5	3	3	5.00	+	+	0.28
P20073	Annexin A7	0	1					0	1	0	1				
P07585	Decorin	9	9	0.71	8	4	3.60	2	4	2	2	2.17			1.42
P51888	Prolargin	9	7	0.69	8	5	3.88	3	4	2	2	10.36			1.73
P51884	Lumican	9	9	1.06	8	6	5.06	3	5	3	3	0.96	+	+	0.71
P80303	Nucleobindin 2	0	1		0	3		1	0	2	0				tumor
Q8IUX7	Adipocyte enhancer binding protein 1	4	5	1.32	8	3	8.56			0	2				tumor
Q9Y6C2	EMILIN 1	4	13	2.89											tumor
Q99715	Collagen alpha 1 XII chain	4	13	22.87											tumor
Q9NX62	Inositol monophosphatase 3	0	5												
Q9HD45	Transmembrane 9 superfamily member 3	2	3	2.8											
Q9UBG0	C type mannose receptor 2	2	4	1.68											
P21810	biglycan														2.08
P16422	Epithelial cell adhesion molecule														tumor
P04216	Thy 1 membrane glycoprotein														tumor

SUBSTITUTE SHEET (RULE 26)

P48960	CD97												tumor
Q9H3K2	Growth hormone inducible transmembrane protein									0	1		tumor
P24043	Laminin alpha-2									0	2		
P61225	Ras-related protein Rap-2b												Tumor
P38606	V type proton ATPase catalytic subunit A									0	2		
		Breast Tumor			Breast Tumor Bone Meta			Lymphoma		Pancreas Tumor			Glioblastoma		Liver Meta
Accession	Protein name	N	T	Ratio T/N	T	M	Ratio T/M	N	T	N	T	Ratio T/N	N	T	Ratio T/N
P21980	Protein-glutamine gamma-glutamyl transferase 2									2	2	8.5			
Q0P6H9	transmembrane protein 62									0	1				

Accession	Protein name	Representative peptide sequence as identified by the MS analysis
Q9BXN1	Asporin	ISTVELEDFKR ; GLTSLYGLILNNNK ; SLYSAISLFNNPVK
		LNASIADLQSQLR ;
Q15230	Laminin alpha 5	TLSELMSQTGHLGLANASAPSGEQLL R
		RPASNVSASIQR ;
Q16363	Laminin alpha 4	TLFPVVLEQLDDYNAK ;
		LSDLQEALDQALNHVR ;
		ELVDEEADEAYELLSQAESWQR
P11047	Laminin gamma 1	NTIEETGNLAEQAR ;
		EAQQALGSAAADATEAK
Q14766	Latent transforming growth factor beta binding protein 1	TSTDLDVDVDQPKEEK ;
		LCQIPVHGASVPK
Q14767	Latent transforming growth factor beta binding protein 2	GAGGQSMSEAPTGDHAPAPTR
		LSLLEELSLAENQLLK ;
P20774	Mimecan	LDFTGNLIEDIEDGTFSK
		GEYVLLEAALTDLR ;
Q9UGT4	Sushi domain containing protein 2	HDPTFEPLFPSETTLNPSLAQEAAK
		ILHQHLGAPEER
Q9NZM1	Myoferlin	VLTPPMGTVMDVLK ;
Q15582	Transforming growth factor-beta-induced protein ig-h3	FSMLVAAIQSAGLTETLNR

O00468	Agrin	NELMLNSSLMR
P05556	Integrin beta 1	LSENNIQTIFAVTEEFQPVYK ; SLGTDLMNEMR
P16144	Integrin beta 4	VLSTSSTLTR ; NVISLTEDVDEFR ; DYIPVEGELLFQPGEAWK ; LTVPGLSENVPYK
P18564	Integrin beta 6	LGFGSFVEK ; PLSVGR
P23229	Integrin alpha 6	LWNSTFLEEYSK
P06756	Integrin alpha V	IYIGDDNPLTLIVK ; ANTTQPGIVEGGQVLK ; ISSLQTTEKNDTVAGQGER
Q5ZPR3	CD276	VVLGANGTYSCLVR ; TALFPDLLAQGNASLR
P05362	Intercellular adhesion molecule 1	ASVSVTAEDEGTQR ; DGTFPLPIGESVTVTR ; EPAVGEPAEVTTTVLVR
Accession	**Protein name**	**Representative peptide sequence as identified by the MS analysis**
Q6UWY5	Olfactomedin like protein 1	VYLLIGSR ; TLAEMLLQEAEEEK
Q9NRN5	Olfactomedin like protein 3	IYVLDGTQNDTAFVFPR
P23142	Fibulin 1	TGYYFDGISR ; SQETGDLDVGGLQETDK ; IIEVEEEQEDPYLNDR
Q05707	Collagen A1 (XIV)	MMEMFGLVEK ; TLFLGVTNLQAK ; IFASQASGFSDALTGMVK
P12109	Collagen A1 (VI)	TAEYDVAYGESHLFR ; NLVWNAGALHYSDEVEIIQGLTR ;

P12111	Collagen A3 (VI)	SLQWMAGGTFTGEALQYTR LLTPITLTSEQIQK ; ALILVGLER ; VVIHFTDGADGDLADLHR ; MKPLDGSALYTGSALDFVR GETGPSGPVGPAGAVGPR ; GPSGPQGIR ; GEAGAAGPAGPAGPR
P08123	Collagen A2 (I)	SALSGHLETVILGLLK ; RAEDGSVIDYELIDQDAR ;
P07355	Annexin A2	SLYYYIQQDTK ETSGNLEQLLLAVVK ;
P08758	Annexin A5	GLGTDEESILTLLTSR ; SIPAYLAETLYYAMK GSIHDFPGFDPNQDAEALYTAMK ;
P08133	Annexin A6	GLGTDEDTIIDIITHR ; SLEDALSSDTSGHFR ; TNEQMHQLVAAYK
P20073	Annexin A7	DENQSINHQMAQEDAQR ; SEIDLVQIK
P07585	Decorin	ASYSGVSLFSNPVQYWEIQPSTFR ; NLHALILVNNK DLPPDTTLLDLQNNK
P51888	Prolargin	LPGLVFLYMEK ; NLMQLNLAHNILR ; LEHLYLNNNSIEK ; LENLLLLDLQHNR
P51884	Lumican	LKEDAVSAAFK ; SLEDLQLTHNK ; SLEYLDLSFNQIAR
P80303	Nucleobindin 2	LHDVNSDGFLDEQELEALFTK ; LVTLEEFLK ;

		EVWEETDGLDPNDFDPK ; VQNIHPVESAK
Q8IUX7	Adipocyte enhancer binding protein 1	TPTQEQLLAAAMAAAR ; GEDEDEVSEAQETPDHAIFR
Accession	**Protein name**	**Representative peptide sequence as identified by the MS analysis**
Q9Y6C2	EMILIN 1	LEQLGGLLA**NVS**GELGGR ; LGAL**NSS**LQLLEDR ; LDTVAGGLQGLR
Q99715	Collagen alpha 1 XII chain	EAG**NIT**TDGYEILGK ; EAAGDDEVTVVEPASSTSVVLSSLKP ETLYLV**NVT**AEYEDGFSIPLAGEETTE EVK ; LNW**NPS**PSPVTGYK
Q9NX62	Inositol monophosphatase 3	QVALQTFG**NQT**TIIPAGGAGYK ; YVTTMVCVAVNGKPMLGVIHKPFSE YTAWAMVDGGSNVKAR ; ALGGHMTTLSGEEISYTGSDGIEGGLL ASIR
Q9HD45	Transmembrane 9 superfamily member 3	IVDV**NLT**SEGK ; **NLS**GQPNFPCRVNAVPRPIPEK ; MYGLFQTSFYFGYMAVFSTALGIMC GAIGYMGTSAFVRK
Q9UBG0	C type mannose receptor 2	VTPAC**NTS**LPAQR ; KKP**NAT**AEPTPPDR ; GGCVALATGSAMGLWEVK**NCT**SFR
P21810	biglycan	(K)EISPDTTLLDLQNNDISELR(K) ; (K)IQAIELEDLLR(Y) ;

P16422	Epithelial cell adhesion molecule	(K)DLPETLNELHLDHNK(I) (K)TQNDVDIADVAYYFEK(D) ; (R)TDKDTEITCSER(V) ; (K)EMGEMHRELNA(-)
P04216	Thy 1 membrane glycoprotein	(K)HVLFGTVGVPEHTYR(S) ; (K)VLYLSAFTSK(D) ; (R)TNFTSKYNMK(V)
P48960	CD97	(K)QAELEEIYESSIR(G) ; (K)LVDELMEAPGDVEALAPPVR(H) ; (R)RLSAVNSIFLSHNNTKELNSPILFAF
Q9H3K2	Growth hormone inducible transmembrane protein	SHLESSDGEAGR(D) (R)VATMLATGGNR(K) ; (K)DRIHSTYMYLAGSIGLTALSAIAIS RTPVLMNFMMR(G) ; (R)SIPYDQSPGPK(H)
P24043	Laminin alpha-2	TAVADNLLFYLGSAK
P61225	Ras-related protein Rap-2b	ASVDELFAEIVR
P38606	V type proton ATPase catalytic subunit A	EILQEEEDLAEIVQLVGK ; VGSHITGGDIYGIVSENSLIK ;
P21980	Protein-glutamine gamma-glutamyl transferase 2	VVTNYNSAHDQNSNLLIEYFR ; ALLVEPVINSYLLAER
Accession	**Protein name**	**Representative peptide sequence as identified by the MS analysis**
Q0P6H9	transmembrane protein 62	TKWLDIK

Figure 18 (table over previous 8 pages)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/EP2011/071335

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

3, 24, 28(completely); 1, 2, 14-23, 25-27(partially)

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2011/071335

A. CLASSIFICATION OF SUBJECT MATTER
INV. G01N33/574
ADD.

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G01N

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

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

EPO-Internal, BIOSIS, EMBASE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CHEN ET AL: "Pancreatic Cancer Proteome: The Proteins That Underlie Invasion, Metastasis, and Immunologic Escape", GASTROENTEROLOGY, ELSEVIER, PHILADELPHIA, PA, vol. 129, no. 4, 1 October 2005 (2005-10-01), pages 1187-1197, XP005155705, ISSN: 0016-5085, DOI: DOI:10.1053/J.GASTRO.2005.08.001 the whole document ----- -/--	1

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
9 March 2012	30/05/2012

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Rosin, Oliver

3

Form PCT/ISA/210 (second sheet) (April 2005)

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	Castronovo: "Proteomic analysis of human pancreas cancers for the identification of targetable biomarkers", FASEB J , April 2009 (2009-04), XP002621086, Retrieved from the Internet: URL:http://www.fasebj.org/cgi/content/meeting_abstract/24/1_MeetingAbstracts/354.5 [retrieved on 2011-01-28] abstract	1
A	HARSHA H C ET AL: "A compendium of potential biomarkers of pancreatic cancer.", PLOS MEDICINE 7 APR 2009 LNKD-PUBMED:19360088, vol. 6, no. 4, 7 April 2009 (2009-04-07), page E1000046, XP002621087, ISSN: 1549-1676 suppl information	1
X	LIU XIAOJUN ET AL: "Membrane proteomic analysis of pancreatic cancer cells", JOURNAL OF BIOMEDICAL SCIENCE, KLUWER ACADEMIC PUBLISHERS, DO, vol. 17, no. 1, 13 September 2010 (2010-09-13), page 74, XP021071257, ISSN: 1423-0127, DOI: 10.1186/1423-0127-17-74 table 1	1-3
X	WO 2006/024283 A2 (UNIV DRESDEN TECH [DE]; KLOEPPEL GUENTER [DE]; LUETTGES JUTTA [DE]; KA) 9 March 2006 (2006-03-09) page 108 page 101 page 123 - page 124	1-3, 14-28
X	WO 2004/074510 A1 (GARVAN INST MED RES [AU]; BIANKIN ANDREW [AU]; SEGARA DAVENDRA [AU]; H) 2 September 2004 (2004-09-02) abstract page 138 page 160 page 23 page 120 - page 124 page 98 - page 99	1-3, 14-28
X	WO 03/030725 A2 (UNIV JOHNS HOPKINS [US]; HRUBAN RALPH H [US]; ARGANI PEDRAM [US]; IACO) 17 April 2003 (2003-04-17) page 15	1-3, 14-28

-/--

3

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2003/073144 A1 (BENSON DARIN R [US] ET AL) 17 April 2003 (2003-04-17) abstract; table 4	1-3, 14-28
X	LACOBUZIO-DONAHUE C A ET AL: "DISCOVERY OF NOVEL TUMOR MARKERS OF PANCREATIC CANCER USING GLOBAL GENE EXPRESSION TECHNOLOGY", AMERICAN JOURNAL OF PATHOLOGY; [10640], AMERICAN SOCIETY FOR INVESTIGATIVE PATHOLOGY, US, vol. 160, no. 4, 1 April 2002 (2002-04-01) , pages 1239-1249, XP008003747, ISSN: 0002-9440 table 1	1-3, 14-28
X	ANSON W. LOWE ET AL: "Gene Expression Patterns in Pancreatic Tumors, Cells and Tissues", PLOS ONE, vol. 2, no. 3, 28 March 2007 (2007-03-28), page E323, XP055021296, DOI: 10.1371/journal.pone.0000323 suppl. information, fig 2	1-3, 14-28
X	IACOBUZIO-DONAHUE C A ET AL: "Highly expressed genes in pancreatic ductal adenocarcinomas: a comprehensive characterization and comparison of the transcription profiles obtained from three major technologies", CANCER RESEARCH, AMERICAN ASSOCIATION FOR CANCER RESEARCH, US, vol. 63, no. 24, 15 December 2003 (2003-12-15), pages 8614-8622, XP002403086, ISSN: 0008-5472 table 2	1-3, 14-28
X	IACOBUZIO-DONAHUE ET AL: "Exploration of global gene expression patterns in pancreatic adenocarcinoma using cDNA microarrays", AMERICAN JOURNAL OF PATHOLOGY; [10640], AMERICAN SOCIETY FOR INVESTIGATIVE PATHOLOGY, US, vol. 162, no. 4, 1 April 2003 (2003-04-01) , pages 1151-1162, XP009027111, ISSN: 0002-9440 table 1	1-3, 14-28

3

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/EP2011/071335

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2006024283 A2	09-03-2006	DE 102004042822 A1	16-03-2006
		DE 112005002742 A5	09-08-2007
		WO 2006024283 A2	09-03-2006
WO 2004074510 A1	02-09-2004	CA 2516290 A1	02-09-2004
		EP 1599603 A1	30-11-2005
		US 2006269921 A1	30-11-2006
		US 2009233286 A1	17-09-2009
		WO 2004074510 A1	02-09-2004
WO 03030725 A2	17-04-2003	AU 2002342053 A1	22-04-2003
		US 2003180747 A1	25-09-2003
		WO 03030725 A2	17-04-2003
US 2003073144 A1	17-04-2003	NONE	

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. claims: 3, 24, 28(completely); 1, 2, 14-23, 25-27(partially)

   Uses, methods and entities for pancreatic cancer comprising protein myoferlin.

   ---

2. claims: 4(completely); 1, 2, 14-23, 25-27(partially)

   Uses, methods and entities for pancreatic cancer comprising protein latent-transforming growth factor beta binding protein 2.

3-23. claims: 1, 2, 5-22, 25, 26(all partially)

   Uses, methods and entities for pancreatic cancer comprising protein transforming growth factor beta induced protein ig-h3 (invention 3), ... up to ... protein transmembrane protein 62 (invention 23), in as far as mentioned in the claims.
