



(11) **EP 2 977 760 A1**

(12) **EUROPEAN PATENT APPLICATION**  
published in accordance with Art. 153(4) EPC

(43) Date of publication:  
**27.01.2016 Bulletin 2016/04**

(51) Int Cl.:  
**G01N 33/574 (2006.01) G01N 33/53 (2006.01)**  
**G01N 33/48 (2006.01)**

(21) Application number: **14767707.4**

(86) International application number:  
**PCT/KR2014/002209**

(22) Date of filing: **17.03.2014**

(87) International publication number:  
**WO 2014/148780 (25.09.2014 Gazette 2014/39)**

(84) Designated Contracting States:  
**AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR**  
Designated Extension States:  
**BA ME**

(72) Inventors:  
• **KIM, Youngsoo**  
**Seoul 135-280 (KR)**  
• **YOON, Junghwan**  
**Seoul 110-744 (KR)**  
• **KIM, Hyunsoo**  
**Bucheon-si**  
**Gyeonggi-do 420-804 (KR)**  
• **JIN, Jonghwa**  
**Seoul 136-725 (KR)**

(30) Priority: **20.03.2013 KR 20130029626**

(74) Representative: **Ter Meer Steinmeister & Partner**  
**Patentanwälte mbB**  
**Nymphenburger Straße 4**  
**80335 München (DE)**

(71) Applicant: **Seoul National University R&DB**  
**Foundation**  
**Gwanak-gu**  
**Seoul 08826 (KR)**

(54) **BIOMARKER FOR DIAGNOSING LIVER CANCER**

(57) The present invention relates to a biomarker for diagnosis or prognosis of liver cancer which is at least one member selected from the group consisting of ACADVL, ANLN, BASP1, MTHFD1, CAPN1, C4A, FLNB and PABPC1, and a measurement kit, a composition and a method which are for diagnosis or prognosis of liver cancer which include the same. The biomarker according to the present disclosure enables, with high accuracy and sensitivity, an early diagnosis of or monitoring of hepatocellular carcinoma and effective prognosis or assessment of progression of disease. In addition, simple and effective diagnosis of or monitoring of liver cancer is possible by noninvasive tests using blood, urine or the like, and thus the present invention will be very effective in the early detection of liver cancer at home or general clinic through a development of, for example, point of care testing (POCT).

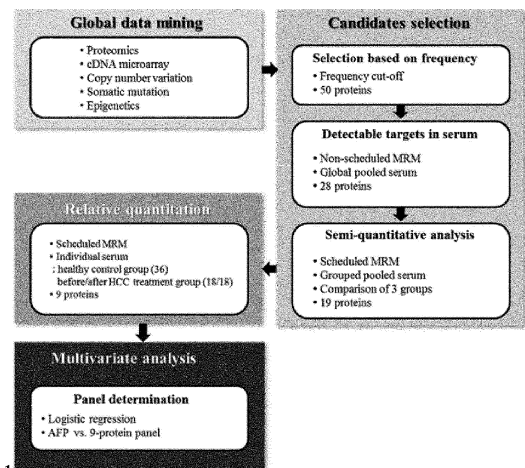


FIG. 1

EP 2 977 760 A1

**Description****BACKGROUND OF INVENTION**5 **Field of the Invention**

[0001] The present disclosure generally relates to biomarkers and its use to detect or diagnose liver cancer.

10 **Description of the Related Art**

[0002] Hepatocellular carcinoma (HCC) is a type of cancer which is most prevalent in adult liver cancer and is the 3<sup>rd</sup> leading cause of cancer related death worldwide (Stefaniuk P, et al., 2010, World J Gastroenterol 16: 418-424). Patients with HCC are usually asymptomatic until later stages. Thus the patients frequently miss the optimal timing for a proper treatment and thus usually have a very poor prognosis. Particularly the patients usually die within one year after diagnosis when it is not possible to remove the cancer by surgery. Therefore the patients can be greatly benefited by early diagnosis or detection.

[0003] At present, diagnosis of HCC is commonly performed using the serum marker alfafetoprotein (AFP) often in combination with ultrasonography. However the concentration of AFP is also found to be increased in non-HCC such as chronic or acute hepatitis and benign tumors. The specificity and sensitivity of AFP to detect HCC is considered to be low and AFP has been found to have about 66% overall sensitivity and 82% overall specificity. Thus AFP is not suitable for all types of liver cancer. AFP concentration of 200 (or 400) ng/mL is used as criteria in combination with other clinical findings in ultrasonography test for diagnosis of HCC.

[0004] Currently there are no biomarkers developed to accurately detect liver cancer in a person without symptoms and measuring the level of AFP is a main non-invasive method for early detection of liver cancer in people at high risk of developing the same. When AFP was first developed as a marker, the criteria considered having enough specificity and sensitivity to diagnose HCC was AFP level of 20 ng/mL. However the specificity remained at just 60%. The current level suggested by the American Association for the Study of Liver Diseases is 200ng/mL, in which case the specificity is increased. However the sensitivity is very low at just 22%. According to the previous studies, the overall specificity and sensitivity of AFP in diagnosing liver cancer is about 66% and 82%, respectively. Thus AFP has a limitation in applying for all types of liver cancer.

[0005] US Patent Application Publication No. 2010/030437 relates to liver cancer diagnosing methods and compositions and discloses methods of diagnosing liver cancer by detecting CpG islands present in BASF1, SPINT2, APC, CCND2, CFTR, RASSF1 and SRD5A2 genes.

[0006] KR Patent No.1161789 relates to novel biomarkers for diagnosing liver cancer and its use, and discloses the use of NK4 proteins for the diagnosis or prognosis of liver cancer.

[0007] Also other serum markers whose criteria for diagnosing are yet to be determined include such as Descarboxyprothrombin(DCP), Prothrombin Induced by Vitamin K Absence II (PIVKA-II), glycosylated AFP versus total AFP (L3 fraction), alpha fucosidase, glypican 3, HSP-70 and the like. However those are used as prognosis markers and not available as a marker for diagnosis due to their low specificity.

[0008] Therefore there are urgent needs to develop biomarkers with improved specificity and sensitivity to overcome the limitations of current testing.

**SUMMARY OF THE INVENTION**

[0009] The present disclosure is to provide biomarkers and its use, in alone or in combinations of markers, in diagnoses hepatocellular carcinoma non-invasively with improved specificity and sensitivity.

[0010] In one aspect, the present invention provides one or more biomarkers for diagnosing or prognosis of hepatocellular liver cancer that are selected from the group consisting of ACADVL, ANLN, BASP1, MTHFD1, CAPN1, C4A, FLNB and PABPC1.

[0011] The biomarkers according to the present disclosure make possible the accurate and sensitive detection of hepatocellular live cancer, considering the high sensitivity and the specificity and the convenience of the detection provided by the present markers, in alone or in combination with the other established markers and/or testing, for example AFP. In one embodiment, the biomarkers of the present disclosure are used as combination of markers and include ANLN and FLNB; BASP1 and PABPC1; BASP1, PABPC1, and CAPN1; BASP1, PABPC1, CAPN1 and ANLN; BASP1, PABPC1, CAPN1, ANLN and C4A; BASP, PABPC1, CAPN1, ANLN, C4A and MTHFD1; BASP1, PABPC1, CAPN1, ANLN, C4A, MTHFD1 and ACADVL; or BASP1, PABPC1, CAPN1, ANLN, C4A, MTHFD1, ACADVL and FLNB, but are not limited thereto.

[0012] The biomarkers according to the present disclosure can be used for diagnosing or prognosis of hepatocellular

live cancer by detecting the level (concentration) and/or the presence or absence of the biomarkers at nucleic acid or protein level.

[0013] Hence the biomarkers according to the present disclosure can be provided as a kit or a composition for diagnosis or prognosis of hepatocellular liver cancer comprising an agent to detect the present biomarkers.

[0014] The kit or composition for HCC detection can analyze the biomarkers in a sample using various detecting methods. In one embodiment, the kits may be provided as for example and ELISA kit, a dip stick rapid kit, a MRM kit, a microarray, a gene amplification or an immunoassay kit.

[0015] Various agents may be employed for detecting the present markers and include for example the agents to detect the level (concentration) and/or presence or absence of the biomarkers at nucleic acid protein level. In one embodiment, nucleic acid sequences of the biomarkers according to the present disclosure (genes encoding the present biomarkers), its complimentary sequences, and fragments thereof, or antibodies specific to the biomarker protein that is encoded by the corresponding nucleic acid sequence, antibody fragments thereof, aptamers, avidity multimers and peptidomimetics are included. In other embodiment, the agents for a western blot, an ELISA, a radioimmunoassay, an immunodiffusion assay, an immunoelectrophoresis, an immunostaining, an immunoprecipitation, a complementary fixation test, a FACS, a mass spectrometry and a protein microarray are included. In still other embodiment, a monoclonal antibody, a polyclonal antibody, a substrate, an aptamer, an avimer, a peptidomimetic, a receptor, a ligand and a cofactor are included. In still other embodiment, the agents for a polymerase chain reaction, a reverse transcription polymerase chain reaction, a competitive polymerase chain reaction, a Nuclease protection assay (RNase, S1 nuclease assay), an in situ hybridization assay, a DNA microarray and a northern blot assay are included. In still other embodiment, a primer set, or a probe, or a primer set and a probe that specifically recognizes each biomarker according to the present disclosure are included.

[0016] In other aspect, the present disclosure provides methods for detecting the present biomarkers in vitro to diagnose or prognosis HCC. The methods comprise detecting the level or the presence of at least one biomarker selected from the group consisting of ACADVL, ANLN, BASP1, MTHFD1, CAPN1, C4A, FLNB and PABPC1 at a nucleic acid and/or protein level in a sample from a subject in need thereof; and comparing the detection result with that of a corresponding marker from a control sample wherein the change in the concentration and/or the presence of the biomarkers at the nucleic acid and/or protein level in the subject from the control sample indicates the presence of HCC in the subject.

[0017] Among the biomarkers according to the present disclosure, the expressions of ACADVL, ANLN, MTHFD1, and FLNB are increased, and the expression of BASP1, CAPN1, C4A, and PABPC1 are decreased when the subject of interest is afflicted with HCC. The present detection method may be performed by employing various analysis methods such as a microarray, a gene amplification, an antigen-antibody reaction or a mass spectrometry. The combinations of makers used for the present method and other analysis methods to be employed are as described hereinbefore and after.

[0018] In still other aspect, the present disclosure provides methods for screening an agent for preventing or treating liver cancer, which comprise providing a cell comprising at least one nucleic acid sequence encoding the biomarkers selected from the group consisting of ACADVL, ANLN, BASP1, MTHFD1, CAPN1, C4A, FLNB and PABPC1; contacting the cell with a testing agent; determining the expression level of the gene at a protein and/or mRNA level in the cell treated with the testing agent; comparing the expression level of the cell to that of a control and selecting the testing agent as a potential therapeutic agent when the expression level of the nucleic acid sequence encoding at least one of the markers is decreased in the cell treated with the testing agents, compared to that of the control.

#### **ADVANTAGEOUS EFFECTS**

[0019] The biomarkers according to the present disclosure enable, with high accuracy and sensitivity, an early diagnosis or monitoring of hepatocellular carcinoma as well as an effective prognosis or assessment of the progression of the disease. The present methods employing the present markers provide a noninvasive test by using the sample such as blood and the like. The present biomarkers may be developed as a part of POCT device for a convenient and efficient early detection of HCC at home or in the clinics.

#### **BRIEF DESCRIPTION OF THE DRAWINGS**

[0020] These and/or other aspects and advantages of the invention will become apparent and more readily appreciated from the following description of the embodiments, taken in conjunction with the accompanying drawings of which:

Fig. 1 shows a schematic description of HCC biomarker discovery process employed in the present disclosure.

Figs. 2A to 2F show the result of data mining for liver cancer markers. Each figure shows the results of proteome analysis, gene expression analysis, gene copy-number variation, epigenetic analysis, and somatic mutation analysis results and the summary of the above analysis. References in Fig 2B to 2D are as follows: Lee et al., Nature Medicine

2006; 12:410-416; Roessler et al., Cancer Research 2010; 70:10202-10212; Satow et al., Clinical Cancer Research 2010; 16:2518-2528; Chen et al., Hepatology 2010; 52:1690-1701; Chochi et al., J. of Pathology 2009; 217:677-684; Hashimoto et al., Modern Pathology 2004; 17:617-622; Acevedo et al., Cancer Research 2008; 68:2641-2651; Hernandez-Vargas et al., PLoS ONE 2010; 5(3): e9749. doi:10.1371/journal.pone.0009749; ZHU, Cell Research 2005; 15:272-280.

Fig. 3 shows a response curve of IS peptide using ATP-dependent RNA helicase A(DHX9) peptide.

Fig. 4A to 4D show the result verified by Western blot of four markers, AFP, ANLN, C4A and FLNB respectively.

Fig 5A to 5C show ROC curves of AFP and 2-marker panel. LR analysis was performed by comparing AFP with the 2-marker panel. Using LR analysis models, false positive/negative rates were prepared in the classifier tables, in which error rates are also shown. LR analyses were performed for 3 types as indicated in each figures.

Fig 6A to 6J show ROC curves of the multi-marker panel combined using LR analysis models. The markers comprised in each combination are indicated in each figure.

### **DETAILED DESCRIPTION OF THE EMBODIMENTS**

**[0021]** It takes tremendous amount time and efforts to discover protein marker candidate specific to a disease. Thus, as an initial step, the present inventors have performed global data mining with regard to liver cancer in several disciplines including proteomics to select the most significant candidates.

**[0022]** The present invention is based on the findings of 50 most significant proteins in liver cancer which were selected through the data mining of 5 different types of datasets in proteomics, gene expression using cDNA microarray, copy number variation, somatic mutation, and epigenetics, and the findings of the proteins whose expression level have been significantly different between the healthy control group and liver cancer patient groups (untreated and treated).

**[0023]** Therefore in one aspect, the present disclosure relates to at least one biomarker for diagnosing or prognosis of hepatocellular carcinoma selected from the group consisting ACADVL, ANLN, BASP1, MTHFD1, CAPN1, C4A, FLNB and PABPC1.

**[0024]** The term diagnosis as used herein refers to determining the disease or disorder susceptibility of a subject, determining whether a subject has a specific disease or disorder, determining the prognosis (for example, identification of transitional cancer status, stages or progression of a cancer or determining the response to cancer treatments) of a subject who has specific disease or disorder, or therametrics(for example, monitoring the status of a subject to provide the information on the efficacy of treatment).

**[0025]** The term biomarker for diagnosing or diagnosis marker as used herein refers to an agent that may discriminate a liver cancer tissues or cells from normal cells or a treated liver cancer tissues or cells, and comprises an organic and biological molecule and the like, such as proteins or nucleic acid molecules, lipid, glycolipid, and glycoprotein that has increased or decreased in tissues or cells compared with normal control samples. In the present disclosure, the agent is one or more markers selected from the group consisting of long-chain specific acyl-CoA dehydrogenase (ACADVL), actin-binding protein, anillin (ANLN), brain acid-soluble protein 1 (BASP1), c-1-tetrahydrofolate synthase, cytoplasmic (MTHFD1), calpain-1 catalytic subunit (CAPN1), complementary C4-A (C4A), filamin-B (FLNB), and polyadenylate-binding protein 1 (PABPC1) or its corresponding genes or fragments, the expression level of which is increased or decreased at hepatocellular carcinoma tissues or cells. Particularly the expression levels of ACADVL, ANLN, MTHFD1, and FLNB is increased, and that of BASP1, CAPN1, C4A, and PABPC1 is decreased.

**[0026]** In the present disclosure, the present markers may be used in alone or two or more markers may be used in combination to further improve the specificity and/or sensitivity. For example, two, three, four, five, six, seven or more markers may be combined. Further the present markers may also be used together with the conventional markers, such as AFP, and/or with the other diagnosis test available, such as liver ultrasonography. The person skilled in the art would be able to select the combination of markers that show a desired sensitivity and specificity using the methods such as Logistic regression analysis and/or analysis of the biological samples from the subjects including a normal person and patient using the methods such as described in the examples of the present disclosure. In one embodiment of the present disclosure, the combination includes, but is not limited to, for example, ANLN and FLNB; BASP1 and PABPC1; BASP1, PABPC1, and CAPN1; BASP1, PABPC1, CAPN1 and ANLN; BASP1, PABPC1, CAPN1, ANLN and C4A; BASP, PABPC1, CAPN1, ANLN, C4A and MTHFD1; BASP1, PABPC1, CAPN1, ANLN, C4A, MTHFD1 and ACADVL; and BASP1, PABPC1, CAPN1, ANLN, C4A, MTHFD1, ACADVL and FLNB.

**[0027]** The hepatocellular carcinoma (HCC) of the present disclosure is a primary malignant tumor of liver tissue developed in a subject who has risk factors such as alcoholic abuse, viral hepatitis and metabolic liver disease. HCC which usually does not develop fibrous stroma leads to a bleeding and necrosis, and thus resulting in a vascular invasion

into hepatic portal vein system which may lead to a hepatorrhesis and abdominal plasma exudation in severe cases.

**[0028]** The biological sample of the present disclosure is a substance or a mixture of the substances that contain or is expected to contain/express one or more of the present biomarkers, and includes cells, tissues or bodily fluids from an organism, particularly human, for example, whole blood, urine, plasma, and serum, but is not limited thereto. Also the sample includes cells or tissues cultured in vitro as well as those derived directly from an organism. Various samples may be used for the detection of HCC markers according to the present disclosure. In one embodiment, urine, whole blood, plasma and/or blood serum can be used. In other embodiment, the liver tissues/cells or in vitro cell cultures from an organism of interest, where HCC has developed or HCC is plausible or likely to be developed, may be used, but the samples are not limited thereto. Also the fractions or derivatives of the blood, cells or tissues are included. When cells or tissues are used, lysates thereof may also be used.

**[0029]** The term detection or detecting as used herein refers to a determination in quantity, quality and/or changes in expression patterns and/or profiles of the expression. The detection includes a determination of the presence and/or absence as well as the levels of the present markers. The present markers may be detected using the methods known in the art, and the person skilled in the art would be easily able to select appropriate methods for the detection.

**[0030]** The markers according to the present disclosure can be detected to analyze the presence or absence of the markers and/or the expression level, the difference in the expression level or the changes in the expression level at mRNA and/or protein level through various quantitative and/or qualitative analysis methods known in the art.

**[0031]** The detection of the markers to diagnose HCC according to the present disclosure may be based on the functional and/or antigenic characteristics thereof. The markers can be detected using the agents that specifically interact with the markers at the nucleic acid level, particularly mRNA level and/or protein level, or the agents that activates the markers may also be used.

**[0032]** In this perspectives, the present disclosure relates to nucleic acid sequences or the sequences complementary thereto or the fragments thereof, or antibodies or aptamers that specifically recognize the protein encoded by the nucleic acid sequences.

**[0033]** The quantitative and qualitative analysis of the makers for HCC diagnosis according to the present disclosure may be done by various methods known in the art that can detect the nucleic acid molecules or proteins quantitatively or qualitatively. For example, the methods include a western blot, an ELISA, a radioimmunoassay, an immunodiffusion, an immunoelectrophoresis, an immunostaining, an immunoprecipitation, a complement fixation assay, an array system such as a nucleic acid array or a protein array such as antigen array, a nucleic acid transcription and amplification system, eTag system, a system based on tailed beads, a binding with a tailed antibody in solution/suspension and a detection by flow cytometry, or a mass spectrometry, and the like. These methods are known and documents such as chip-based capillary electrophoresis: Colyer et al. 1997. J Chromatogr A. 781(1-2):271-6; mass spectroscopy: Petricoin et al. 2002. Lancet 359: 572-77; eTag systems: Chan-Hui et al. 2004. Clinical Immunology 111:162-174; microparticle-enhanced nephelometric immunoassay: Montagne et al. 1992. Eur J Clin Chem Clin Biochem. 30:217-22 may be referred.

**[0034]** In one embodiment of the present disclosure, a mass spectrometry is used for detecting the present markers, wherein the proteins are extracted from the appropriate samples and analyzed using the method such as described in the Examples of the present disclosure, and in the literatures Kim, et al. 2010 J Proteome Res. 9: 689-99; Anderson, L et al. 2006. Mol Cell Proteomics 5: 573-88 may also be referred. In one embodiment Multiple Reaction Monitoring (MRM) technology utilizing Triple Quadrupole LC-MS/MS and QTRAP and the like may be used. MRM is a method for exactly quantifying multiple markers present in biological samples in minute amount. In MRM, by a first mass filter (Q1), parent or precursor ions are selected from the ion fragments generated in ionization source and transferred to a collision cell. And then the precursor ions arrived at the collision cell collide with internal collision gas, and are fragmented into products or daughter ions and transferred to a second mass filter (Q2), from which only the specific ions are delivered to a detector. In this way only the information of the desired target can be obtained with high selectivity and sensitivity. The literature Gillette et al., 2013, Nature Methods 10:28-34 and the like may be referred.

**[0035]** In other embodiment, the agents that specifically recognize or bind to the present protein markers or mRNA or DNA encoding the same or the arrays comprising such agents may be used.

**[0036]** In a still other embodiment, an immunoassay using sandwich system like ELISA (Enzyme Linked Immuno Sorbent Assay), or RIA (Radio Immuno Assay) and the like may be used for quantitative and/or qualitative detection of the present markers. In this system, the biological samples are reacted with a first antibody fixed to a solid substrate/support such as a glass, a plastic (for example, polystyrene), polysaccharides, a bead, a nylon or nitrocellulose membrane or a microplate well to form a complex and the complex is then allowed to react with an second antibody that is usually labeled with agents that can be detected directly or indirectly such as radioactive substances like  $^3\text{H}$  or  $^{125}\text{I}$ , fluorescent materials, chemiluminescent substances, hapten, biotin, or digoxigenin and the like. In some cases, the labeling materials are conjugated with an enzyme such as horseradish peroxidase, alkaline phosphatase, or maleate dehydrogenase that is able to produce colors or color changes or illuminate in the presence of appropriate substrates.

**[0037]** Other methods based on immune reaction may also be used. In other embodiment, an Immuno Electrophoresis such as an Ouchterlony plate, a Western blot, a Crossed IE, a Rocket IE, a Fused Rocket IE, or an Affinity IE, which

can detect the markers simply by antigen-antibody reaction may be used. The agents or materials that may be used in the methods described above are known the art. For example, the markers may be detected through an antigen-antibody reaction, or a reaction with a substrate, nucleic acid or peptide aptamers, receptors or ligands that specifically recognize the present markers, or cofactors or using mass spectrometry. The agents or materials that bind or interact specifically with the markers of the present disclosure can be utilized by means of chip or with nanoparticles. The immunoassay or immunostaining methods as described above are disclosed in the following literatures : Enzyme Immunoassay, E. T. Maggio, ed., CRC Press, Boca Raton, Florida, 1980; Gaastra, W., Enzyme-linked immunosorbent assay(ELISA), in Methods in Molecular Biology, Vol. 1, Walker, J.M. ed., Humana Press, NJ, 1984 etc. The intensities of the signals generated by the immunoassay mentioned above are then analyzed, namely compared with the signals from appropriate controls for the determination whether the sample has HCC or not.

**[0038]** Each marker of the present disclosure can be used for diagnosis or prognosis of HCC by detecting the marker(s) quantitatively and/or qualitatively at nucleic acid level, particularly at mRNA level. Various methods known in the relevant art may be used for this purpose. For example, for the quantitative or qualitative detection of the marker(s) or the expression pattern thereof at mRNA level, a polymerase chain reaction (PCR) combined with cDNA synthesis, a RT-PCR (reverse Transcription PCR), a competitive RT-PCR, a real-time RT-PCR, a Nuclease Protection Assay (NPA) such as a RNase or a S1 nuclease assay, an in situ hybridization assay, a DNA microarray or chip or Northern blot and the like may be used, the methods of which are known in the art and also can be performed using commercially available kits. The skilled person in the art would be able to select appropriate methods for practicing the present invention. For example, the transcriptome size that exists in cells can be determined by a northern blot, which has an advantage of being able to use diverse probes. A NPA is useful for multi-marker analysis. An in situ hybridization is useful for localizing the transcriptome such as mRNA in cells or tissues. A RT-PCR is particularly useful when small amount of samples are employed.

**[0039]** The reagents for detecting the presence markers in the amount or for determining the pattern of mRNA via RT-PCR comprise, for example, a primer set and/or probe specific to the marker(s) of the present disclosure. The terms "primer" or "probe" refers to a nucleic acid sequences that can bind to a template via complementarity and have free 3'-hydroxy group that can serves as a starting point for polymerization using the template by reverse transcriptase or DNA polymerases. The reagents used herein can also be labeled with appropriate materials for the signal detection such as color forming agent, illuminating or fluorescent materials. In one embodiment, a northern blot or a reverse-transcription PCR is used for the mRNA detection. The latter is a method to detect or determine the expression level of a specific gene expression in a sample, wherein the RNA, particularly mRNAs are isolated from the sample and cDNAs are synthesized from the isolated RNA, and then the cDNAs are amplified using a specific set of primers, or a combination of primers and probe. Such methods to detect the presence/absence or quantity of a specific gene or marker may be found in for example, Han, H. et al, 2002. Cancer Res. 62: 2890-6.

**[0040]** In other aspect, the present disclosure provides a kit, composition or system for diagnosing or prognosis of hepatocellular carcinoma comprising an agent to detect the expression level or presence of at least one biomarker selected from the group consisting of ACADVL, ANLN, BASP1, MTHFD1, CAPN1, C4A, FLNB and PABPC1. The detection agents and the methods of using the agent are described hereinbefore. The agents for detecting the present marker(s) can be provided in a container separated in sections, or can be provided separately. In this regards, the present disclosure provides a device that provides the agents for detecting the present markers in separate sections in a container.

**[0041]** In one embodiment, the detecting agents are antibodies specifically binding to one or more markers of the present disclosure, and can detect the protein in a sample such as serum quantitatively and/or qualitatively. In other embodiment, the antibodies may be provided as conjugated to a substrate or support, for example to a well of the multi-well plates or a surface of the glass slides or nitrocellulose filters.

**[0042]** In other aspect, the detecting agents which may be used for the present disclosure may be provided as a type of array such as microarray or chip, and the references for the array preparation technology may be found in for example, Schena et al., 1996, Proc Natl Acad Sci USA. 93(20):10614-9; Schena et al., 1995, Science 270(5235):467-70; and U.S. Pat. Nos. 5,599,695, 5,556,752 and 5,631,734. The detecting agents that can be provided as attached to an array include, but are not limited to, antibodies specifically recognizing the present marker, or antibody fragments, or aptamers, avimers(avidity multimer) or peptidomimetics.

**[0043]** The detecting agents may also be labeled for the direct or indirect signal detection via sandwich forms. In the case of direct labeling method, biological samples such as serum to be used for array may be labeled with a fluorescent material such as Cy3 and Cy5. In the case of indirect labeling, unlabeled samples are allowed to bind to the detecting agent which is attached to the array, and then the target or marker proteins are detected by the labeled antibodies which specifically recognize the marker or target. In the case of sandwich methods, the sensitivity and specificity of the detection is usually high and able to detect the marker at the pg/mL level. In addition, labeling agents including such as radioactive materials, agents that produce visible colors under proper condition, magnetic particles and high-density electron particle and the like may also be used.

[0044] For the detection of fluorescence signals, a scanning confocal microscopy may be used, which are available from for example, Affymetrix, Inc. or Agilent Technologies, Inc. and the like.

[0045] The kits or compositions of the present disclosure may also include one or more additional components as needed for the detection of the present marker(s), which include, for example, binding buffers, reagents for preparing biological samples, syringes for blood collecting or negative and/or positive controls.

[0046] The kits or compositions of the present disclosure in which various agents as described above may be contained may be provided for an ELISA, a dip stick rapid kit, a MRM, microarray, a gene amplification, or an immunoassay. The reagents which may be included for a particular kit may be appropriately selected from the known agents in view the present disclosure.

[0047] In one embodiment, an ELISA or a dip stick rapid kit is used, and in this case, the antibodies to detect the present markers may be provided as conjugated to a substrate/support, for example to wells of a multi-well plate or surface of a glass slide or a nitrocellulose paper. A dip stick is a form widely used in POCT (point of care technology) in which the biomarkers of the present disclosure may be detected by employing one or more antibodies that are conjugated to a substrate such as nitrocellulose paper which is then contacted with a sample such as serum for example by dipping one end of the stick to the serum and then the sample is moved through the substrate by a capillary action, and the markers are detected by a color that is developed when the markers of interest are bound by the antibody attached to the substrate.

[0048] In other embodiment, a MRM kit which is based on peptide analysis is provided. MRM analysis may be performed or used as described above. The MRM is a method using the peptides that selectively recognize a particular protein, and can detect the markers in biological samples in a more stable manner compared to the existing methods that employ antibodies that are sensitive to environmental conditions such as temperature and humidity and the like.

[0049] Also, the kits of the present disclosure may also include instructions for using the kit to detect the markers according to the present disclosure.

[0050] In still other aspect, the present disclosure provides compositions comprising at least one biomarker selected from the group consisting ACADVL, ANLN, BASP1, MTHFD1, CAPN1, C4A, FLNB and PABPC1, and/or the agents that are able to detect the markers.

[0051] In still other aspect, the present disclosure provides methods for diagnosing or prognosis of HCC, or detecting in vitro at least one maker of the present disclosure for diagnosis or prognosis of HCC, which comprises detecting the presence or amount of at least one biomarker selected from the group consisting of ACADVL, ANLN, BASP1, MTHFD1, CAPN1, C4A, FLNB and PABPC1 at a nucleic acid and/or protein level in a sample from a subject in need thereof.

[0052] In one embodiment, the present methods may additionally include a step of comparing the detection result of the subject sample to that of a corresponding marker in a control sample from healthy subject wherein the change in the amount and/or in the presence or absence of the biomarker at the nucleic acid and/or protein level in the subject from the control sample indicates that the subject has HCC. In one embodiment, the present biomarkers ACADVL, ANLN, MTHFD1 and FLNB are found to be increased in HCC sample and the biomarkers BASP1, CAPN1, C4A and PABPC1 are found to be decreased in HCC sample. Thus the subject can be diagnosed accordingly based on the detection result.

[0053] Samples such as biological samples, control or reference groups employed in the present methods, methods for determining the present biomarkers and the reagent used therefor, and methods for data analysis required for the interpretation or diagnosis are as described hereinbefore and in the following Examples.

[0054] The subject to be diagnosed by the present methods is a mammal, particularly human. The subjects include ones who are suspected of HCC or at high risk or others who need for HCC diagnosis or prognosis. In one embodiment, the present methods are applied for a subject who has symptoms indicative of liver related disease, such as an abdominal pain, hepatomegaly, ascites, jaundice, muscle wasting, hepatitis(for example HCV infection), or esophageal varix and the like. In other embodiment, the subject includes ones who appear to be normal and yet to develop HCC related symptoms.

[0055] In other embodiment, the subject may be a person whose concentration of serum AFP is low or in a normal range and thus not diagnosed as HCC by applying AFP criteria. In this case, the serum concentration of AFP may be about 0  $\mu\text{g/l}$  (not detected) to 20  $\mu\text{g/l}$ , e.g., 0  $\mu\text{g/l}$  (not detected) to 5  $\mu\text{g/l}$ , 5  $\mu\text{g/l}$  to 10  $\mu\text{g/l}$ , 10  $\mu\text{g/l}$  to 15  $\mu\text{g/l}$ , or 15  $\mu\text{g/l}$  to 20  $\mu\text{g/l}$ .

[0056] When the combinations of two or more markers are used for the present methods, a profile, namely a set of data including quantitative information in relation to the protein expressions in the sample of interest may be generated.

[0057] When the profiles are obtained using the present markers, HCC in the sample of a subject may be diagnosed or determined by comparing the result of the sample with that of the control or reference groups. The control or reference groups which may be used for the present methods include a negative control from a healthy or normal subject, or from a subject cured of HCC, and a positive control is from HCC patient diagnosed not by the present method but by the other methods, or from liver cirrhosis patient, or from hepatitis patient.

[0058] In one embodiment, as a control or reference groups, normal samples or samples from HCC-treated patient

are used and compared with the acquired profile from the subject who needs diagnosis.

**[0059]** For comparing the profiles obtained, methods known in the art may be used. For example, comparison of the digital images of expression profiles, comparison using expression database, or US Patent NOs 6,308,170 and 6,228,575 may be referred.

**[0060]** The acquired profiles using the present markers may be processed by data analysis methods known in the art. For example, the methods include but are not limited to nearest neighbor classifier, partial-least squares, SVM, AdaBoost and clustering-based classification. Or the methods described in the literatures such as Ben-Dor et al (2007, J. Comput. Biol. 7: 559-83), Nguyen et al (2002, Bioinformatics 18:39-50), Wang et al (2003, BMC Bioinformatics 4:60), Liu et al (2001, Genome Inform. Ser. Workshop Genome Inform. 12:14-23), Yeang et al (2001, Bioinformatics 17 Suppl 1:S316-22) and Xiong (2000, Biotechniques 29(6):1264-8, 1270) and the like may also be referred.

**[0061]** Further, various statistical analysis methods may be employed to calculate the significance of the data obtained using the present markers to diagnose or prognosis HCC. In one embodiment, a logic regression method is used for a statistical analysis and the literatures such as Ruczinski, 2003, Journal of Computational and Graphical Statistics 12:475-512 may be referred. A logic regression method is similar to CART method in which classifier is presented as a binary tree, but Boolean operators are used for each node that is more general and related to characteristics than "and" operator used in CART. There are other analysis methods such as nearest shrunken centroids (Tibshirani. 2002 PNAS. 99:6567-72), random forests (Breiman. 2001. Machine Learning 45:5-32 and MART (Hastie. 2001. The Elements of Statistical Learning, Springer).

**[0062]** In one embodiment, the confidence level of a significant difference between the test sample and the controls to diagnose the sample as HCC may be determined in a statistical analysis. The raw data used for a statistics are the values analyzed in duplicate, triplicate or more for each marker.

**[0063]** In one embodiment, it can be determined that the sample of interest is to be significantly different from the normal control when the values of at least one or at least two markers from ACADVL, ANLN, BASP1, MTHFD1, CAPN1, C4A, FLNB and PABPC1 in the sample of interest are found to be higher than the preset value to be significant. In other embodiment, the sample of interest is to be significantly different from the normal control when the markers ANLN and/or FLNB are significantly higher than the preset value.

**[0064]** This statistical analysis is useful to decide the clinical significance of the clinical and genetic data as well as the data produced from the biomarkers.

**[0065]** Also the present methods may be used to determine the severity of the HCC. For example, it may be diagnosed as a mild HCC, a moderate HCC or a severe HCC by comparing the data obtained with the profiles from the positive and/or negative control groups. Further, the marker profiles obtained from a certain HCC groups may be classified according to a certain criteria. In this way, it is possible to find the disease at the early stage so that expensive test such as magnetic resonance imaging (MRI) may not be needed.

**[0066]** In other embodiment, the present method may be performed multiple times with a certain interval over a period of time for example over one year to monitor the changes in the expression pattern of the present markers. The increase or decrease in the expression depending on the markers can be used to relate the sample with a particular status or severity of HCC. Also in one subject, the data may be used to determine the development, progression, and/or aggravation of HCC over time in comparison to the previous data from the same subject or to the data from the controls. The changes in the expression pattern or levels over time thus can be used to treat HCC or to prevent the progression of HCC into a more advanced stage. Also the present methods may be used together with conventional diagnosis methods such as AFP test, a ultrasonography, a computerized axial tomography(CT scan) or a magnetic resonance imaging (MRI).

**[0067]** In other aspect, the present disclosure relates to methods of screening an agent for treating or preventing HCC or symptoms associated therewith. The agents that are able to change or affect or influence the expression level or pattern of the present markers in HCC tissues or cells can be selected as candidates for HCC therapeutic agent. In one embodiment, the present screening method comprises providing a cell including at least one gene selected from the group consisting of ACADVL, ANLN, BASP1, MTHFD1, CAPN1, C4A, FLNB; contacting the cell with a testing agent; determining the expression level of the gene at the protein and/or mRNA level in the cell treated with the testing agent; comparing the expression level of the cell to that of a control and selecting the testing agent as a potential therapeutic agent when the cells treated with the testing agent shows a decreased level of expression of the markers compared to that of the controls.

**[0068]** In one embodiment, small molecules are used to identify agents useful for treating and/or preventing HCC or symptoms associated therewith. For example small molecules with molecular weight of about less than 1,000Da such as 400 Da, 600 Da, or 800Da are used. If desired, small molecules may form part of a library, the total number of small molecules included therein may vary from dozens to millions. Test substance of a library may include peptides, peptoides, circular or liner oligomeric compounds, template based compounds such as benzodiazepine, hydantoin, biaryls, carbocyclic and polycyclic compounds such as naphthalene, phenothiazine, acridine, steroids and the like, carbohydrate and amino acid derivatives, dihydropyridine, benzhydryl and heterocyclic compounds such as triazine, indole, thiazolidine and the like, but are not limited thereto.

[0069] The present disclosure is further explained in more detail with reference to the following examples. These examples, however, should not be interpreted as limiting the scope of the present invention in any manner.

## EXAMPLES

[0070] In the present examples, the expression levels of the 50 proteins originated from the serum samples of 3 subject groups (healthy control group, untreated HCC group and treated HCC group) were verified via multiple reaction monitoring (MRM) using mass spectrometry (Triple Quadrupole LC-MS/MS), and the markers for HCC that have shown significant differences between the groups were discovered.

### Example 1. Materials and Methods

#### 1-1. Clinical sample information

[0071] The institutional review board of Seoul National University Hospital (approval No. H-1103-056-355) approved the protocol of the present invention, and the written informed consent was obtained from each patient or their legally authorized representative. The clinical characteristics of the patients are as in Table 1

[0072] In the present invention, 36 samples from a healthy control group (male/female = 18/18), 18 untreated HCC samples from a patient group diagnosed with HCC (male/female = 13/5) and 18 treated HCC samples from HCC patients who underwent HCC therapy (male/female = 13/5) were used. Untreated and treated HCC samples were collected from the same patients. The HCC serum samples were collected before the start of treatment for liver cancer, and then the treated samples were collected and used for the analysis only when liver cancer wasn't recurred and metastasized and completely cured based on the data followed for more than a year after the treatment.

[0073] The rates of liver cancer is 70% more prevalent in male than in female, so the samples were collected to include 70% more male samples than the female samples. To avoid the complexity of the data analysis arising from the nature of the samples used, the samples from the patients who had been treated only with transarterial chemoembolization (TACE) and percutaneous ethanol injection (PEIT) were used for the analysis. Besides, Virus (HBV, HCV) and alcohol are known as the two main cause of liver cancer, but the recent study showed that in Asia and Africa, 70% of liver cancer is arise from HBV infection and 20% from HCV and 10 % from alcohols. Based on this report, the samples of liver cancer only form HBV infection which is the most prevalent cause in Asia were used for the study. The clinical information of the samples used is as follows.

[Table 1]

Clinical characteristics of the subject

	MRM analysis		Western blot analysis	
	Untreated HCC group and corresponding treated HCC Group	Healthy control group	Untreated HCC group and corresponding treated HCC Group	Healthy control group
<b>Total patient number</b>	18 in each group	36	13 in each group	13
<b>Gender (Male/Female)</b>	S / 13	18 / 18	10 / 3	8 / 5
<b>Age (Mean, Range)</b>	00.6 (47-81)	58.7 (50-69)	62.4 (48-79)	56.2 (52-67)
<b>Etiology of river disease</b>	HBV. 18 (100%)		HBV, 13 (100%)	
<b>Locoregional modality</b>				
TACE	7		4	
PEIT	11		9	
<b>AFP value (Mean,Range)</b>	1079.4 (141-6900)		245.2 (16-730)	
<20 ng/ml	2		1	
20-200 ng/ml	4		6	
200-1000 ng/ml	7		6	
>1000 ng/ml	S		0	
<b>PIVKA value (Mean,Range)*</b>	916 (5-10720)		117.6 (28-612)	

(continued)

**Clinical characteristics of the subject**

	MRM analysis		Western blot analysis	
	Untreated HCC group and corresponding treated HCC Group	Healthy control group	Untreated HCC group and corresponding treated HCC Group	Healthy control group
<20 ng/ml	4		0	
20-100 ng/ml	6		7	
100-1000 ng/ml	3		6	
>1000 ng/ml	3		0	

\*PIVKA values were provided for 16 (M11/F5) among a total of 18 untreated HCC group

Abbreviationsn:

AFP : Alpha-Fetoprotein

PIVKA: Protein induced by vitamin K absence or antagonist

TACE : Transcatheter arterial chemoembolization

PEIT :Percutaneous ethanol injection therapy

**[0074]** The blood samples were collected and centrifuged immediately at 3000 rpm for 10 min at 4°C to fractionate the serum. The resulting supernatants were then aliquoted(50 $\mu\ell$ ) and stored at -80°C until analysis.

**1-2. Tryptic Digestion of serum samples**

**[0075]** Serum proteins were treated for MRM analysis as follows. Serum proteins were quantified by bicinchoninic acid (BCA) assay. The 200 $\mu\text{g}$  aliquots of the serum samples were denatured with 6 M urea, 50 mM Tris, pH 8.0, and 30 mM dithiothreitol (DTT) at 37°C for 60 min and alkylated with 50 mM iodoacetamide (IAA) at room temperature in the dark for 30 min. The urea was diluted 15-fold with 50 mM Tris, pH 8.0 prior to overnight digestion at 37°C with trypsin (Promega, sequencing-grade modified) using a 1:50 (w/w) enzyme-to-serum concentration ratio.

**[0076]** Tryptic digestion was stopped with formic acid at a final concentration of 1% and desalted on Sep-pak tC18 cartridges (Waters Corp., USA). The Sep-pak tC18 cartridges were equilibrated sequentially with 3 ml methanol and 5 ml water that contained 0.1% trifluoroacetic acid (TFA) prior to loading of the tryptic digestion. The cartridges were washed with 3 ml 0.1% trifluoroacetic acid (TFA) and eluted with 1 ml of 60% ACN, 0.1 % TFA. The eluted samples were frozen and lyophilized on a speed vacuum. Prior to MRM analysis, the samples were reconstituted in 0.1% formic acid to 2  $\mu\text{g}/\mu\text{l}$ .

**1-3. Standard peptide**

**[0077]** Peak areas for the MRM runs were normalized using a standard peptide. A synthetic stable isotope-labeled peptide from ATP-dependent RNA helicase A (DHX9) (sequence ELDALDANDELTP $\mathbf{LGR}$ ) with a C-terminal  $^{15}\text{N}$ -,  $^{13}\text{C}$ -labeled arginine residue ( $\geq 99$  atom % isotopic enrichment) (21<sup>st</sup> Century Biochemicals Inc., MA) was added to the HCC serum peptide mixture before MRM analysis at 50 fmol. After MRM analysis, the peak area of the internal standard peptide in each sample was used to normalize of the peak area of each target peptide in each MRM run.

**1-4. MRM**

**[0078]** For each target protein, we selected peptides and fragment ions for MRM using Skyline (<http://proteome.gs.washington.edu/software/skyline>), an open-source software application for developing MRM methods and analyzing MRM data (Stergachis AB, et al., 2011, Nat Methods 8: 1041-1043)

**[0079]** In brief, the full-length protein sequences were imported into Skyline in FASTA format and designed into peptides, each with a list of product ions for monitoring by MRM. In selecting transitions through Skyline, the peptide filter condition was as follows: maximum length of peptide of 20, including at least 8 amino acids. Peptides with repeat arginines (Arg, R) or lysines (Lys, K) were discarded. If methionine (Met, M) was included in the peptide, it was discarded to avoid the risk of modification. If proline (Pro) lay next to arginine (Arg, R) or lysine (Lys, K), the peptide was discarded. If a peptide contained histidine (His, H), it was discarded to avoid the risk of charge alteration.

**[0080]** Peptides that satisfied these conditions were used as Q1 transitions. Next, we selected a maximum of 5 Q3 transitions from the fragmentation ions that were derived from the Q1 transitions in descending order. The .csv output

files that contained the transition information were exported for MRM into Analyst (AB SCIEX, Foster City, MA), and MRM was run using the nonscheduled MRM method. The MRM data were generated in.wiff and wiff.scan files, which were converted to mzXML format using mzWiff and input into Skyline for MRM data processing, such as selecting the peak intensities of MRM transitions.

5

### 1-5. Quantification by multiple reaction monitoring

**[0081]** MRM was performed on a nano LC system, which was connected to a hybrid triple quadrupole/ion trap mass spectrometer (4000 QTRAP, AB SCIEX, Foster City, CA) that was equipped with a nanoelectrospray interface. The 4000 QTRAP was operated in positive ion MRM mode, in which Q1 and Q3 were set to transmit different precursor/product ion pairs.

10

**[0082]** The LC buffer system used was as follows: mobile phase A, 2% acetonitrile/0.1% formic acid and mobile phase B, 98% acetonitrile/0.1% formic acid. The peptides were separated and eluted at a flow rate of 300 ml/min on a linear gradient of mobile phase B from 2% to 40% B in 43 min. The gradient was ramped up to 70% B for 5 min and 2% B for 10 min to equilibrate the column for the next run. The total LC run time was 60 min. The analytical column was 75  $\mu\text{m}$  x 15 cm, packed with Magic C18AQ resin (5  $\mu\text{m}$ , 100  $\text{\AA}$ , Michrom Bioresources).

15

**[0083]** Typical instrument settings were as follows: ion spray (IS) voltage of 2.3 kV, an interface heater temperature of 200°C, a GS1 (nebulizer gas) setting of 12, and curtain gas set to 15. MS parameters for declustering potential (DP) and collision energy (CE) were determined by linear regression of previously optimized values in Skyline. MRM experiments were performed with a scan time of 50 ms and scan width of 0.002 m/z, using a unit resolution of 0.7 Da (FWHM) for Q1 and Q3. In the MRM runs, scan time was maintained at 50 ms for each transition, and the pause between transition scans was set to 3 ms. MRM transitions were selected to monitor the fragment ion with the highest intensity. Preference for MRM transitions was given to fragment ions with m/z values that were higher than that of the parent (Kirkpatrick DS, et al., 2005, Methods 35: 265-273).

20

**[0084]** The optimal transitions for each targeted peptide were determined empirically using the heavy isotope-labeled synthetic reference peptides. To normalize the amount of each protein, 50 fmol of heavy isotope-labeled forms of the peptides in 0.1% formic acid was injected with all serum samples.

25

### 1-6. Statistical analysis for verification of biomarker candidates

30

**[0085]** Raw data files from the MRM analysis were processed using Skyline. For each targeted peptide, the results of the peak area integration were inspected manually to correct wrong integrations. The default peak integration and Savitzky-Golay smoothing algorithm were applied. Peptides with at least 3-fold signal-to-noise ratios were considered detectable.

35

**[0086]** Two approaches were used to assess HCC candidate proteins. First, we distinguished disease group (untreated HCC group) from nondisease group (healthy control group, treated HCC group). Comparisons between the untreated HCC group (n=18) versus healthy control group (n=36) and untreated HCC group (n=18) versus treated HCC group (n=18) were made using analysis of variance (ANOVA). Based on ANOVA, we selected target peptides that had a significance level below 0.05 in mean intensity level between groups.

40

**[0087]** Second, to evaluate the efficacy of serum biomarkers in distinguishing disease group from nondisease group, we analyzed receiver operator characteristic (ROC) curves and scatter plots. ROC curves and scatter plots were used to determine the optimal cutoff point that yielded the highest total accuracy with respect to discriminating sample groups. The AUC is equal to the probability that a classifier will rank a randomly chosen positive instance higher than a randomly chosen negative one. We performed all statistical analyses and generated all scatter plots and ROC curves with MedCalc (MedCalc, Mariakerke, Belgium, version 12.2.1).

45

### 1-7. Western blot analysis

**[0088]** To confirm the quantification accuracy of the MRM results, following 9 proteins were analyzed using Western blot: long-chain specific acyl-CoA dehydrogenase, mitochondria (ACADVL), actin-binding protein, anillin (ANLN), brain acid-soluble protein 1 (BASP1), c-1-tetrahydrofolate synthase, cytoplasmic (MTHFD1), calpain-1 catalytic subunit (CAPN1), complementary C4-A (C4A), alpha-fetoprotein (AFP), filamin-B (FLNB), and polyadenylate-binding protein 1 (PABPC1).

50

**[0089]** The clinical samples in the Western blot experiment is consisted of 13 individual samples from the healthy control group, untreated HCC group, and treated HCC group (Table 1). Serum sample concentrations were determined by BCA protein assay. Equal amounts of protein (30  $\mu\text{g}$ ) were mixed with SDS loading buffer (62 mM Tris-HCl, pH 6.8, 10% glycerol, 2% SDS, 2%  $\beta$ -mercaptoethanol, and bromophenol blue), boiled for 10 min, and separated by 15-well-comb SDS-PAGE on a 12% acrylamide gel. After separation, serum samples were transferred to polyvinylidene difluoride

55

(PVDF) membranes (Bio-Rad, Cat. #162-0177), which were blocked with 5% BSA (w/v) in TBS-T (25 mM Tris, pH 7.5, 150 mM NaCl, and 0.05% (w/v) Tween-20) for 2 hr at room temperature.

[0090] Membranes were incubated overnight at 4°C with individual primary antibodies (diluted 1:100 to 1:1000). Membranes were washed 5 times with TBS-T and incubated for 2 hr with the appropriate secondary peroxidase-conjugated antibody (1:5000, Santa Cruz Biotechnology, USA). The membranes were then washed 5 times with TBS-T, and target protein bands were visualized using the Chemiluminescent Substrate Kit (Gen-DEPOT, W3651-012). Western blot band intensities were quantified using Multi Gauge. (Fujifilm, ScienceLab 2005, version 3.0). Pooled serum mixture was used as a loading control. All blots were normalized to the band intensity of the pooled serum. Band intensities were analyzed by T-test to identify meaningful differences between sample groups.

### 1-8. Statistical analysis to construct multimarker panel

[0091] In this study, we compared non-disease group (healthy control group, treated HCC group) with the disease group (untreated HCC group) using multimarker panel proteins. Logistic regression (LR) analysis was performed to generate the multimarker panel, consisting of several individual markers that differentiated cancerous from non-cancerous subjects. The discriminatory power was examined in 3 situations: healthy control group versus healthy control group, untreated HCC group versus treated HCC group, and healthy control group plus treated HCC group versus untreated HCC group.

[0092] LR and ROC curves were constructed using MedCalc, and the analysis was performed as follows: Case 1 was healthy control group versus untreated HCC group, Case 2 was untreated HCC group versus treated HCC group, and Case 3 of healthy control group plus treated HCC group versus untreated HCC group. The "Enter" method was used, and the classification table cutoff value was set to 0.5.

[0093] ROC curves were used to evaluate the efficacy of the multimarker panel-in this case, a 2-marker panel. AUC values for individual markers and the 2-marker panel were calculated to examine the distinguishing power of various combinations of HCC candidate markers. Multicollinearity of the panel was checked using IBM SPSS Statistics (version 20, commuter license).

### Example 2. Candidate biomarker selection using global data mining

[0094] To obtain a list of biomarker candidate proteins requires tremendous amount of time and efforts. Thus established liver cancer-specific markers were used to avoid redundant experiments to save time and thus to find new markers more rapidly. Thus, the present inventors have performed global data mining with regard to liver cancer in several disciplines as well as proteomics to select the most significant candidates.

[0095] With regard to HCC, data mining in proteomics, cDNA microarray, copy number variation, epigenetics, and somatic mutation data were processed. The term "frequency" was used for each target protein. Frequency was defined as the total number of occurrences in 5 biological fields. As a result of data mining in the 5 areas, 5,222 liver cancer-related proteins were selected and filtered by prioritizing candidate proteins. The top 50 high-frequency proteins were chosen and examined with regard to whether they were secreted into plasma, and final candidates were selected by confirmation with the Plasma Proteome Database (<http://www.plasmaproteomedatabase.org>). Then, sequence files of the 50 selected candidates were prepared in FASTA format and harvested using the Uniprot website (<http://www.uniprot.org>). The 50 FASTA files were input into Skyline to generate theoretical transitions for the MRM analysis. The overall scheme for this study is outlined in Fig. 1.

#### 2-1. Data mining of proteomic research

[0096] Proteomic studies were screened to determine whether they described quantitative analyses that were based on labeling methods, such as ICAT, SILAC, and iTRAQ, using liver cancer cells, tissues, and human blood, excluding label-free methods. Ten research papers, all published after 2004, were selected with impact factors above 4.0. Analyses that were based on unlabeled methods were excluded. All proteins in the 10 journals put together and analyzed in Ben diagram. The result is shown in Fig.2a.

#### 2-2. Data mining of gene expression

[0097] cDNA microarray research papers that examined gene expression using liver cancer and control tissues were examined. Nine such reports were selected (all published after 2003), with impact factors above 6.7. using these reports, the data mining of gene expression was processed and the result is shown in Fig.2b.

### 2-3. Data mining of copy number variation research

[0098] Earlier publications on copy number variation (CNV) by amplification or mutation of liver cancer were investigated, yielding 3 papers after 2004 with impact factors above 4.4. OncoDB (<http://oncodb.hcc.ibms.sinica.edu.tw/index.htm>) was also used to report copy number variation. Using the 3 papers and OncoDB, the data mining copy number variation research was processed and the result is shown in Fig.2c.

### 2-4. Data mining of epigenetic research

[0099] We selected 3 studies from after 2005 with impact factors above 4.3, in which DNA methylation of CpG islands in liver cancer is disclosed. Using the 3 studies, the data mining of epigenetic research was processed and the result is shown in Fig.2d.

### 2-5. Data mining of somatic mutation research

[0100] Using the OncoDB (<http://oncodb.hcc.ibms.sinica.edu.tw/index.htm>), Japan Liver Cancer (NCC, Riken), and International Cancer Genome Consortium (<http://www.icgc.org/icgc/cgp/66/420/824>) databases, the data mining of somatic mutation research was processed and the result is shown in Fig.2e.

[0101] Summary of above data mining results is shown in Fig. 2f.

#### Example 3. Examination of the linearity of the standard peptide

[0102] The peak area of the internal standard peptide in each sample was used to normalize that of each target peptide in each MRM run. Thus, to assess the quantitative linearity of MRM, a response curve of the standard peptide (sequence of ELDALDANDELTPPLGR, ATP-dependent RNA helicase A (DHX9)) in the presence of matrix peptides was generated. The standard peptide was diluted serially to 1, 5, 10, 50, 100, 250, and 500 fmol in 200  $\mu$ g of serum peptide mixture, which is similar to the MRM conditions for target peptides. Further, to verify the endogenous signal of the target peptide, a blank sample that lacked the internal standard peptide was run. Each MRM analysis was performed three times repeatedly.

[0103] The result is shown in Fig 3. The CV% of the experiment was below 20%, and the correlation coefficient was 0.9984. The response curve of the standard peptide with the serum peptide mixture as matrix indicated that the quantitative linearity of serum MRM at the given concentrations was sufficiently valid to obtain relative quantities of target peptides.

#### Example 4. Detection of target candidates in pooled serum

[0104] Before conducting individual MRM analyses using the 72 serum samples (36 healthy control group, 18 untreated HCC group, and 18 treated HCC group), a preliminary MRM analysis was performed on the target peptides/transitions using pooled serum of all patients to obtain transition information, such as the detectability in serum and the suitability of the transition. The FASTA files of the 50 proteins yielded 498 peptides and 2174 transitions on applying the hierarchy data in Skyline software. After obtaining the resulting MRM data, the final transition was selected using the following criteria: at least 2 peptides were selected per protein, and at least 3 transitions per peptide were chosen as detectable transitions that had a signal-to-noise (S/N) ratio above 3.

[0105] Ultimately, 28 of 50 candidates met the criteria in the first screening step

#### Example 5. Selection of the transitions with technical reproducibility for MRM analysis

[0106] The reproducibility of MRM analyses is critical in making quantitative measurements. In this example, the analyses were performed as follows. The pooled serum mixture from the 3 groups (the pooled serum mixture of 36 healthy control group, 18 the pooled serum mixture of untreated HCC group, and the pooled serum mixture of 18 treated HCC group) was analyzed to select the target protein that is applied to each sample from 333 transitions for 111 peptides, corresponding to the 28 candidate proteins from the first screening step.

[0107] Five repetitive scheduled MRM runs for the pooled serum mixture were performed using the retention time from the detectability experiment, using 500~1000 transitions with window size of 120 seconds. The MRM analysis was conducted once for each sample in the order of healthy control sample, untreated sample (non-treatment), treated sample, and was repeated 5 times. The data from the MRM analysis was imported into Skyline software, exported into the peak area of each peptide transition. The peak area using the internal standard peptides was normalized prior to compare between the averaged relative quantities of each transition.

[0108] Peptides, which had low variance in intensity of normalized peptides (CV < 30%) and showed a confident

difference in quantity (p-value < 0.05) between healthy control group versus untreated HCC group, and untreated HCC group versus treated HCC group, were chosen as target peptides to be applied in individual analysis. We processed MedCalc (version 12.2) statistical programs (t-TEST, ANOVA). As a result, 30 peptides, comprising 90 transitions that corresponded to 19 proteins, were selected as target protein to be applied in individual analysis

#### Example 6. Selection of target marker through MRM analysis using individual serum samples

**[0109]** Individual MRM analysis was performed using the 90 transitions for 30 peptides selected from the second screening in Example 5, corresponding to 19 proteins. The individual sample, of which expression pattern between groups is similar to that of the pooled sample, was selected as a final target protein. The analysis was conducted once for each sample in the order of healthy control sample no. 1, no. 2, untreated HCC sample no. 1, treated HCC sample no. 1 and then healthy control sample no. 3, no. 4, untreated HCC sample no. 2, treated HCC sample no. 2. The data from the MRM analysis were imported into the Skyline software, and then analyzed using MedCalc (version 12.2) statistical program.

**[0110]** Through this, following 9 proteins, that had similar expression patterns in the analysis of the pooled and individual samples (healthy control group, untreated HCC group, treated HCC group), were selected: long-chain specific acyl-CoA dehydrogenase, mitochondria (ACADVL), actin-binding protein, anillin (ANLN), brain acid-soluble protein 1 (BASP1), c-1-tetrahydrofolate synthase, cytoplasmic (MTHFD1), calpain-1 catalytic subunit (CAPN1), complementary C4-A (C4A), alpha-fetoprotein (AFP), filamin-B (FLNB), and polyadenylate-binding protein 1 (PABPC1).

**[0111]** The 9 final target proteins were verified by ROC curve. ROC curve were expressed on a two-dimensional plane to evaluate the relationship of sensitivity and specificity. The bigger AUC (area under the ROC curve) value is the better diagnosis method. The AUC values for corresponding target proteins are as follow table 2, and ROC curve and the interaction plot diagram were generated using MedCalc (version 12.2) statistical program.

**Table 2.**  
**List of proteins showing significant difference among different groups and their AUC value**

Number	Protein Name	Gene Symbol	Peptide Sequence	Area Under ROC Curve	
				Healthy control group vs. Untreated HCC group	Untreated HCC group vs. Treated HCC group
1	Very long-chain specific acyl-CoA dehydrogenase, mitochondrial	ACADVL	ASNTAEVFFDGV R	0.701	0.691
2	Actin-binding protein anillin	ANLN	TQSLPVTEK	0.920	0.744
3	Brain acid soluble protein 1	BASP1	AEGAATEEEEGTPK	0.951	0.719
4	C-1-tetrahydrofolate synthase, cytoplasmic	MTHFD1	GVPTGFILPIR	0.829	0.685
5	Calpain-1 catalytic subunit	CAPN1	YLGQDYEQLR	0.946	0.586
6	Complement C4-A	C4A	VGDTLNLNLR	0.867	0.710
7	Alpha-fetoprotein	AFP	GYQELLEK	0.756	0.735
8	Filamin-B	FLNB	APLNVQFNSPLPGDAVK	0.660	0.707
9	Polyadenylate binding protein 1	PABPC1	EFSPFGTITSAK	0.949	0.608

**[0112]** The AUC values of ACADVL, ANLN, MTHFD1, AFP, and FLNB increased in the healthy control group versus the untreated HCC group and declined in untreated HCC group versus treated HCC group. Those of BASP1, CAPN1, C4A, and PABPC1 declined in the healthy control group versus untreated HCC group and in the untreated HCC group versus treated HCC group.

**[0113]** In the statistical analysis, above 9 proteins were differentially expressed between healthy control group, untreated HCC group, and treated HCC group, with P-values < 0.05. The AUC values of the 9 target proteins ranged from

0.608 to 0.951 (Table 2). The AUC values of 6 proteins exceeded 0.8, those of ANLN, BASP1, CAPN1 and PABPC1 were 0.920, 0.951, 0.946 and 0.949, respectively, in the healthy control group versus untreated HCC group, reflecting the excellent specificity and sensitivity of the markers of the present disclosure.

#### 5 **Example 7. Correlation of MRM analysis with Western blot**

[0114] The 9 candidates from the verification step were verified. Prior to this experiment, a preliminary study was performed to confirm the abundance of these proteins in serum using PaxDb (<http://pax-db.org>) to determine the dilution of primary antibody (data not shown). For the Western blot, 13 of 36 healthy control group from the individual MRM analysis were selected randomly, and 13 untreated HCC group and 13 treated HCC group were analyzed. For the untreated HCC group and treated HCC group, the 13 serum samples differed from what was used in the individual MRM analysis.

[0115] To limit the variability between SDS-PAGE gels, serum from the 13 healthy control group, 13 untreated HCC group, and 13 treated HCC group were pooled and loaded onto the first lane of all SDS-PAGE gels (12%), and its intensity on each gel was used to normalize the intensities of individual samples. As a result, the bands that were generated by ACADVL and PABPC1 were inadequate to calculate intensities; thus, 7 of the 9 candidate proteins were verified by Western blot. Subsequently, the correlation between the protein quantification by MRM and Western blot was determined.

[0116] AFP, ANLN, C4A, and FLNB had the same patterns of expression by MRM analysis (Fig. 4). In the non-disease group (healthy control group, treated HCC group), AFP, ANLN, and FLNB expression decreased compared with disease group (untreated HCC group). AFP is an established serum biomarker for HCC and was verified to have consistent expression by MRM and Western blot. C4A declined in the healthy control group versus untreated HCC group and in untreated HCC group versus treated HCC group. In the comparison between MRM- and antibody-based verification, the relative quantities of 4 proteins (AFP, ANLN, FLNB, and C4A) were similar. In contrast, by Western blot, BASP1, MTHFD1, and CAPN1 had disparate expression patterns compared with MRM analysis.

#### **Example 8. Combination and analysis of the multimarker panel**

[0117] It has been found that a multiprotein panel provide an improved discriminatory power over single markers.

[0118] Four proteins that were validated by MRM analysis and Western blot were used to generate the multimarker panel. AFP and C4A showed colinearity, with a variance inflation factor (VIF) above 10; thus, these proteins were excluded in the final multi-marker model. The ultimate model comprised ANLN and FLNB, for which AFP was the control model as a single marker. Three types of comparisons were made: healthy control group versus untreated HCC group, untreated HCC group versus treated HCC group, and healthy control group plus treated HCC group versus untreated HCC group (Figs. 5A, 5B, and 5C).

[0119] When AFP was used as the lone classifier, 8 of 18 untreated HCC group and 29 of 36 healthy control group were classified correctly, demonstrating an AUC of 0.756 and 31 % error rate.

[0120] In contrast, when the 2-marker panel of the present disclosure was used, 17 of 18 untreated HCC group and all 36 healthy control group were classified correctly, with an AUC of 0.981 and 2% error rate, indicating that the 2-marker panel had improved discriminatory power compared with the AFP model (Fig. 5A).

[0121] As shown in Fig 5B, the AFP-only model classified 12 of 18 untreated HCC group and 9 of 18 treated HCC group correctly (AUC = 0.735, error rate = 42%), whereas the 2-marker panel of the present disclosure separated 15 of 18 untreated HCC group and 11 of 18 treated HCC group successfully (AUC=0.895, error rate=28%). Thus, in identifying patients in the untreated HCC group versus treated HCC group, the discriminatory power of the 2-marker panel was outstanding compared with AFP alone.

[0122] The AFP-only model classified 4 of 18 disease group (untreated HCC group) and 50 of 54 nondisease group (healthy control group, treated HCC group), with an AUC of 0.749 and error rate of 25% (Fig. 5C). Whereas the 2-marker panel identified 15 of 18 disease group and 50 of 54 nondisease group (AUC=0.953, error rate=10%), outstanding compared with AFP alone.

[0123] These results demonstrate that in all comparisons, the 2-marker panel of the present disclosure has greater discriminatory power compared to the traditional single marker AFP.

[0124] Also as shown in Fig 6a to j, the various other liver cancer marker panels combining several individual markers were generated by performing Logistic regression (LR) analysis,. The markers BASP1 and ANLN having the highest AUC value and AFP, and 9 markers (ACADVL, ANLN, BASP1, MTHFD1, CAPN1, C4A, AFP, FLNB, and PABPC1) were combined and these combinations generated a multiprotein panel. We analyzed these combinations to discriminate liver cancer.

[0125] By Logistic regression statistical analysis, 9-protein panel, which comprised 9 proteins, was constructed. This panel had improved discriminatory power compared with the AFP panel in identifying patients in the untreated HCC

group versus treated HCC group, treated HCC group versus normal healthy control group, and untreated HCC group versus normal healthy control group successfully (AUC = 1.00).

[0126] Further 2~8 protein marker panels, which comprised one or more marker proteins selected from 8 marker proteins except AFP, were constructed according to the criteria of ROC curve of BASP1 of which AUC was the highest among markers. And we constructed the ROC curve and compared the AUC value of untreated HCC group with that of normal healthy control group.

[0127] As a result, the AUC value of 2-protein panel was 0.949, that of 3-protein panel was 0.955, and that of 4,5,6,7,8-protein panel were all 1.000. So, two or more combination panel as well as one-panel had outstanding specificity and sensitivity.

[0128] The various singular/plural permutations may be expressly set forth herein for sake of clarity. Although a few embodiments of the present disclosure have been shown and described, it would be appreciated by those skilled in the art that changes may be made in this embodiment without departing from the principles and spirit of the invention, the scope of which is defined in the claims and their equivalents.

[0129] Unless defined or interpreted otherwise, all technical and scientific terms and any acronyms used herein have the same meanings as commonly understood by one of ordinary skill in the art in the field of the invention.

### Claims

1. A kit for diagnosing or prognosis of hepatocellular liver cancer comprising an agent to detect the level of at least one biomarkers selected from the group consisting ACADVL, ANLN, BASP1, MTHFD1, CAPN1, C4A, FLNB and PABPC1.
2. The kit of claim 1, wherein the at least one biomarkers are  
 ANLN and FLNB;  
 BASP1 and PABPC1;  
 BASP1, PABPC1, and CAPN1;  
 BASP1, PABPC1, CAPN1 and ANLN;  
 BASP1, PABPC1, CAPN1, ANLN and C4A;  
 BASP, PABPC1, CAPN1, ANLN, C4A and MTHFD1;  
 BASP1, PABPC1, CAPN1, ANLN, C4A, MTHFD1 and ACADVL; or  
 BASP1, PABPC1, CAPN1, ANLN, C4A, MTHFD1, ACADVL and FLNB.
3. The kit of claim 1, wherein the agent to detect the level of at least one biomarkers is an antibody or fragments thereof, an aptamer, an avidity multimer or peptidomimetics that specifically recognizing the biomarkers.
4. The kit of claim 1, wherein the agents is to detect the concentration and/or presence or absence of the biomarkers at a nucleic acid and/or protein level.
5. The kit of claim 4, wherein the agent to detect the concentration and/or presence or absence of the biomarkers at the protein level is an agent which is employed for a western blot, an ELISA, an radioimmunoassay, an immunodiffusion assay, and immunoelectrophoresis, an immunostaining, an immunoprecipitation, a complement fixation assay, a FACS, an mass spectrometry, or a protein microarray.
6. The kit of claim 4, wherein the agent to detect the concentration and/or the presence or absence of the biomarkers at the protein level is a monoclonal antibody, a polyclonal antibody, a substrate, an aptamer, an avimer, an peptidomimetic, a receptor, a ligand or a cofactor.
7. The kit of claim 4, wherein the agent to detect the concentration and/or the existence of the biomarkers at the nucleic acid level is an agent which is employed for a polymerase chain reaction, a reverse transcription polymerase chain reaction, a competitive polymerase chain reaction, a RNase S1 nuclease assay, an in situ hybridization, a DNA microarray or a northern blot assay.
8. The kit of claim 4, wherein the agent to detect the concentration and/or the presence or absence of the biomarkers at the nucleic acid level is a primer pair, or a probe, or a primer set and a probe that specifically recognize each of the biomarker.
9. The kit of claim 1, wherein the kit is formulated as a ELISA kit, a dip stick rapid kit, an MRM analysis kit, a microarray,

kit, a nucleic acid amplification kit, or an immunoassay kit.

10. The kit of claim 1, further comprising a biomarker AFP.

5 11. A method for diagnosing or prognosing HCC comprising:

detecting the concentration and/or the presence of at least one biomarker selected from the group consisting of ACADVL, ANLN, BASP1, MTHFD1, CAPN1, C4A, FLNB and PABPC1 at a nucleic acid and/or protein level in a sample from a subject in need thereof; and

10 comparing the detection result to that of a corresponding marker from a control sample wherein the change in the concentration and/or presence of the biomarker at a nucleic acid and/or protein level in the subject from the control sample indicates the presence of HCC in the subject.

12. The method of claim 11, wherein the change is the increased concentration of the biomarkers ACADVL, ANLN, MTHFD1 and FLNB and the decreased concentration of biomarkers BASP1, CAPN1, C4A and PABPC1 in the subject compared to the control.

13. The method of claim 11, wherein the detection is performed by at least one of a microarray, a gene amplification, an antigen-antibody reaction of a mass spectrometry.

20 14. The method of claim 11, wherein the detection of the concentration or the presence of the biomarkers at the nucleic acid level is performed by at least one of a polymerase chain reaction, a reverse transcription polymerase chain reaction, a competitive polymerase chain reaction, a RNase S1 nuclease assay, an in situ hybridization, a DNA microarray or a norther blot assay.

25 15. The method of claim 11, wherein the detection of the concentration or the presence of the biomarkers at the protein level is performed by at least one of a western blot, an ELISA, a radioimmunoassay, an immunodiffusion assay, and immunoelectrophoresis, an immunostaining, an immunoprecipitation, a complement fixation assay, a FACS, an mass spectrometry, or a protein microarray.

30 16. The method of claim 11, wherein the at least one biomarkers is a combination of:

- ANLN and FLNB;
- BASP1 and PABPC1;
- 35 BASP1, PABPC1, and CAPN1;
- BASP1, PABPC1, CAPN1 and ANLN;
- BASP1, PABPC1, CAPN1, ANLN and C4A;
- BASP, PABPC1, CAPN1, ANLN, C4A and MTHFD1;
- BASP1, PABPC1, CAPN1, ANLN, C4A, MTHFD1 and ACADVL; or
- 40 BASP1, PABPC1, CAPN1, ANLN, C4A, MTHFD1, ACADVL and FLNB.

17. The method of claim 11, further comprising detecting the concentration of a biomarker AFP.

45 18. The method for screening an agent for treating hepatocellular carcinoma comprising:

providing a cell including at least one gene selected from the group consisting of ACADVL, ANLN, BASP1, MTHFD1, CAPN1, C4A, FLNB and

50 contacting the cell with a testing agent;  
determining the expression level of the gene at a protein and/or mRNA level in the cell treated with the testing agent;

comparing the expression level of the gene to that of a control and selecting the testing agent as a candidate wherein the cells treated with the testing agents shows a decreased level of expression of the genes compared to that of the control.

55

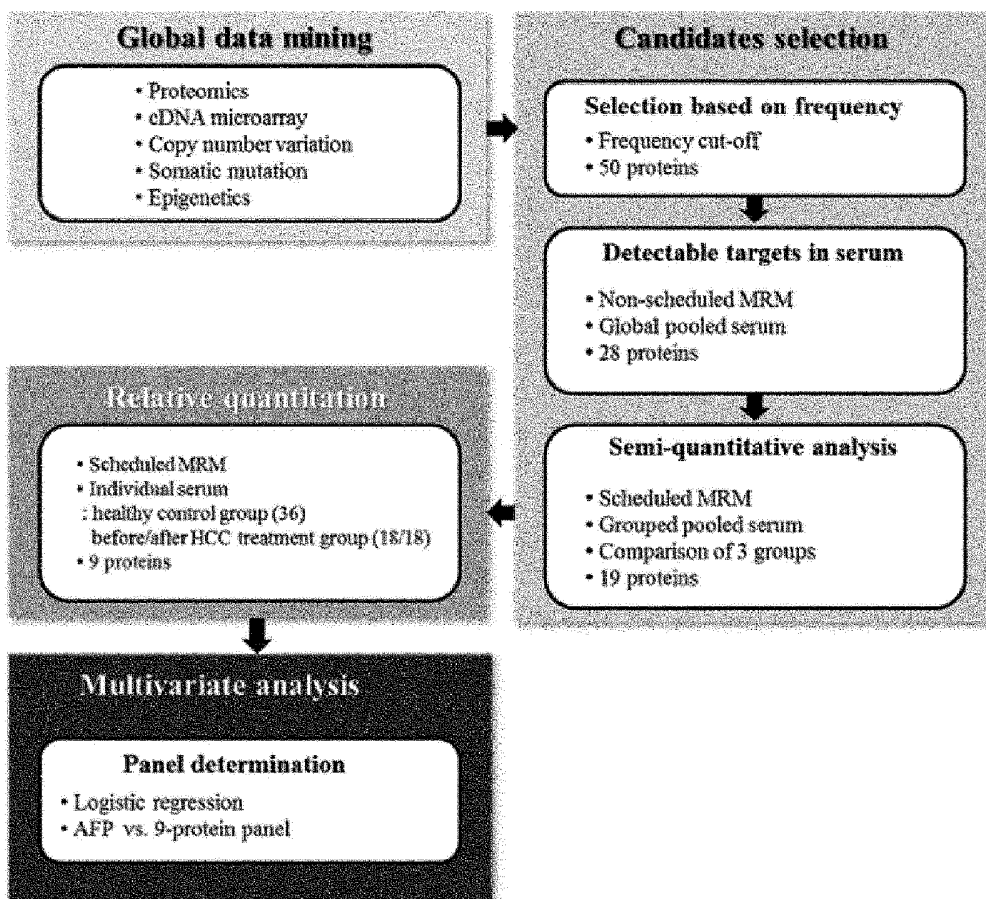


FIG. 1

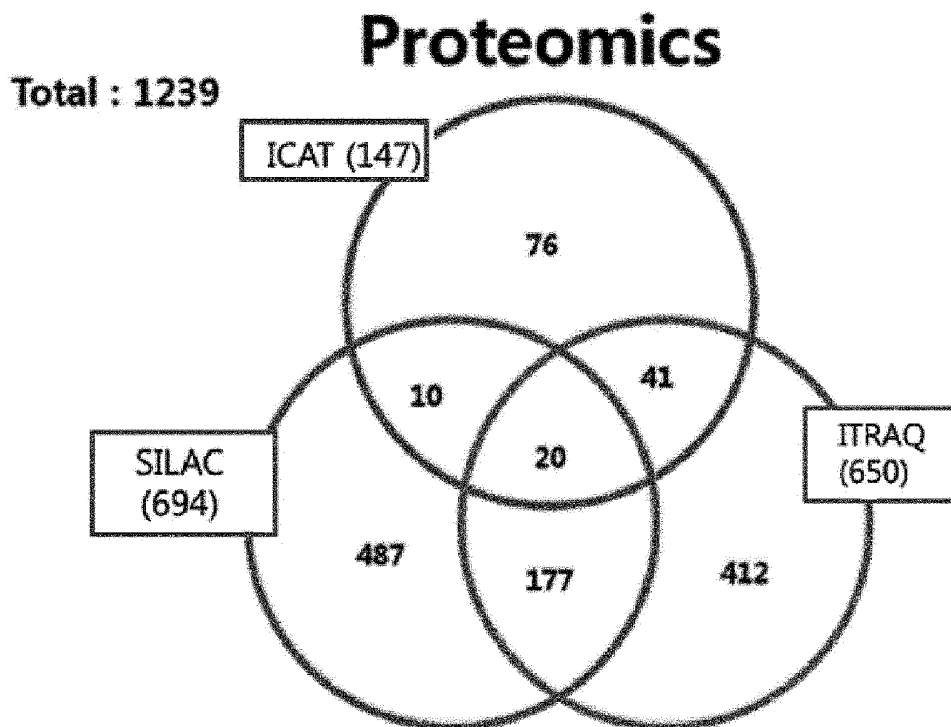


FIG. 2A

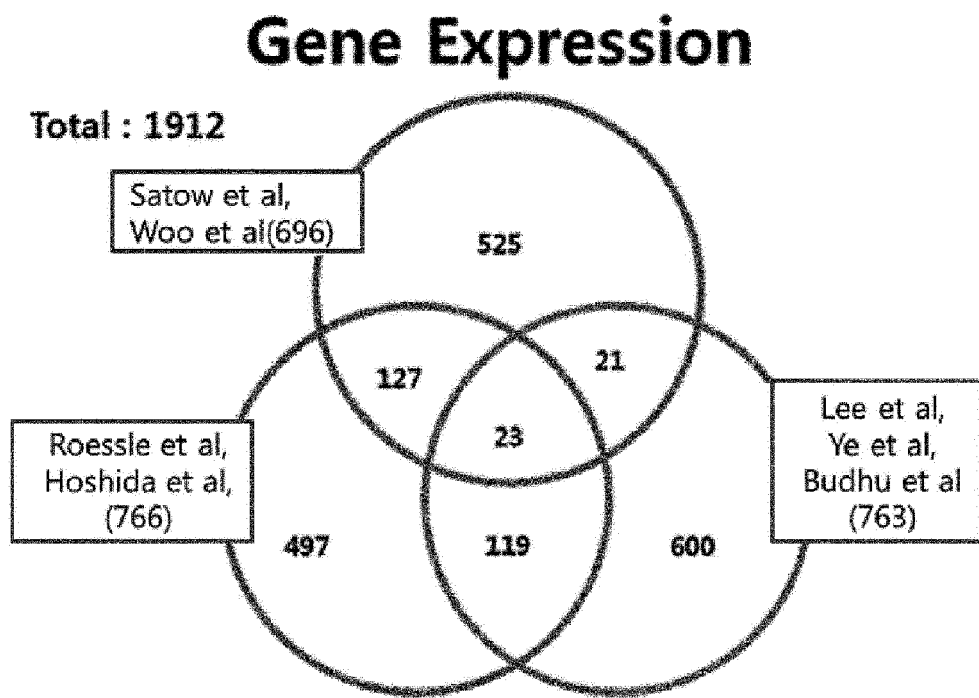


FIG. 2B



Total : 113, no overlap

Data Source	Gene count
Japanese liver cancer (NCC) data import from ICGC	59
Japanese liver cancer (Riken) data import from ICGC	30
OncoDB	24

FIG. 2E

### Summary

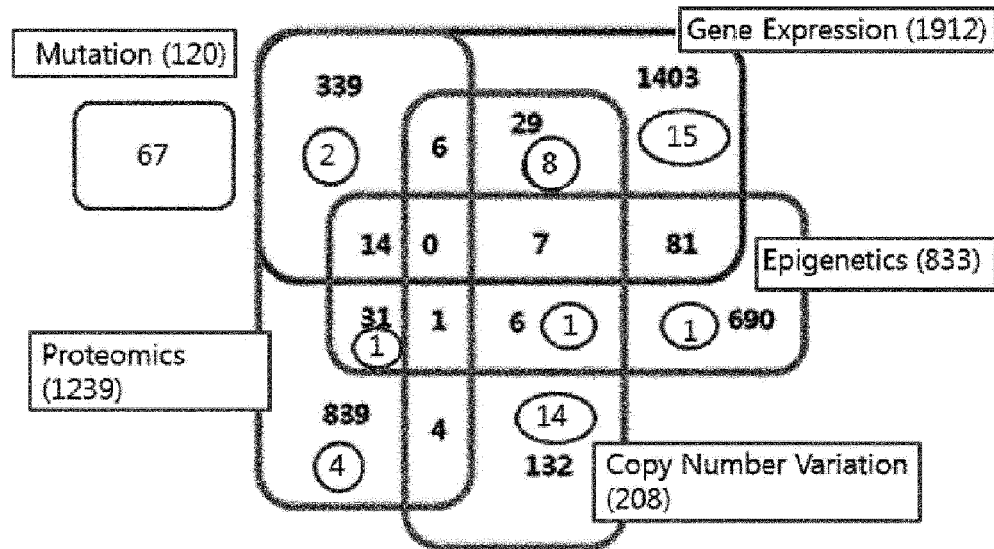
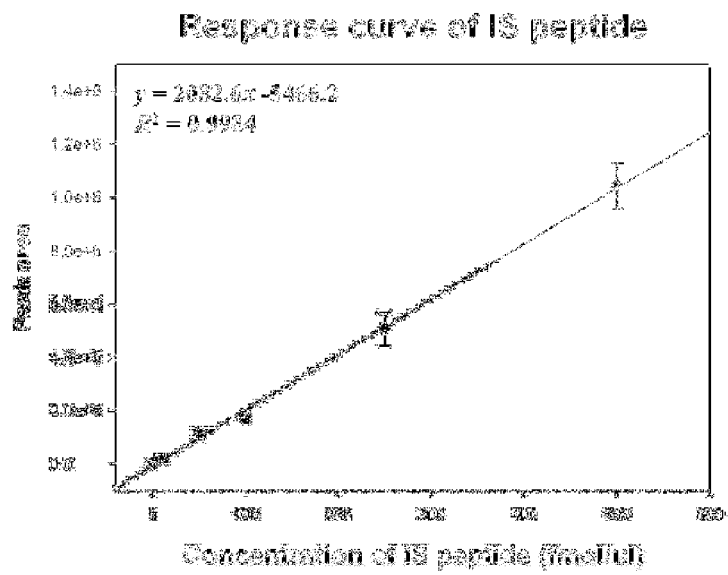


FIG. 2F

FIG. 3



Conc. point	Peak area								Retention time
	0	1	5	10	50	100	250	500	
Average	1135.7	4349.0	2189.7	10352.0	51223.0	100361.7	500046.7	1004049.0	20.0
STDEV	200.4	235.2	323.0	1371.0	17091.0	12121.0	52690.0	85126.0	0.0
CV(%)	7.5	5.4	15.0	13.0	15.1	12.2	10.5	8.2	0.0

FIG. 4A

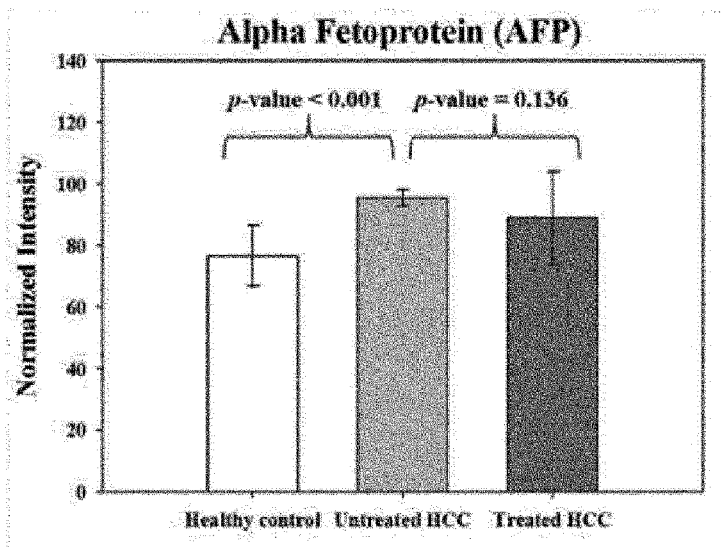
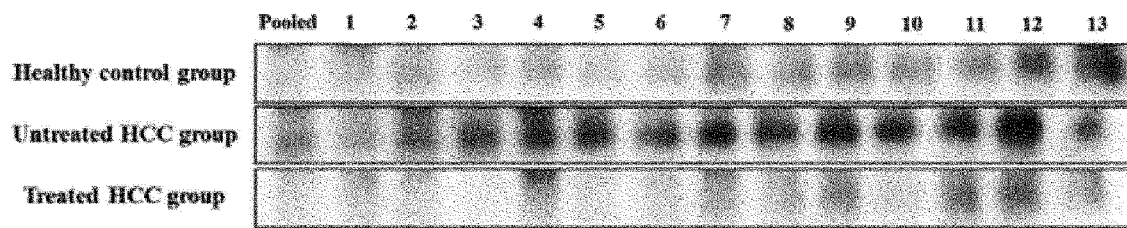


FIG. 4B

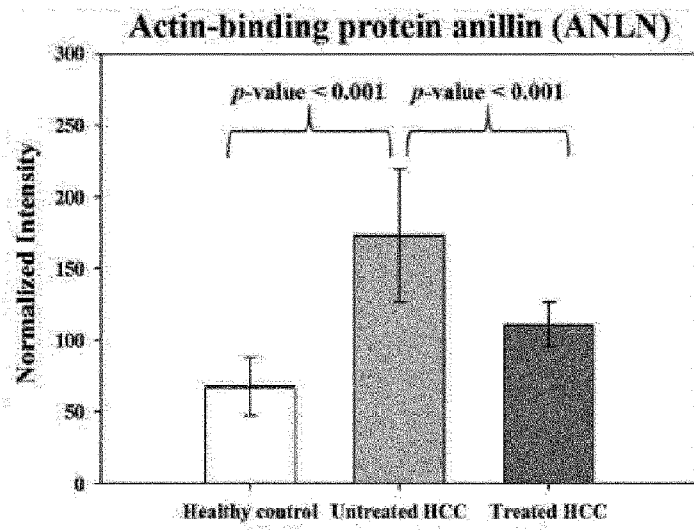
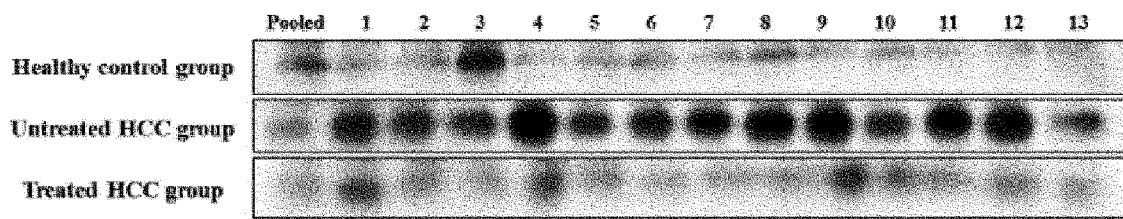


FIG. 4C

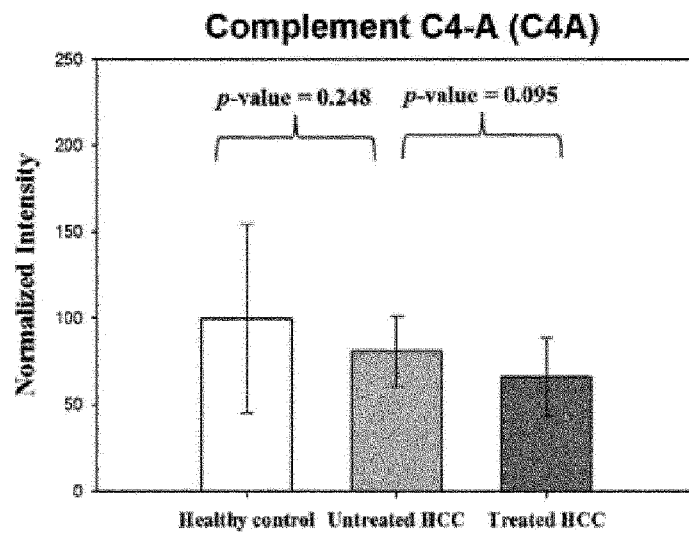
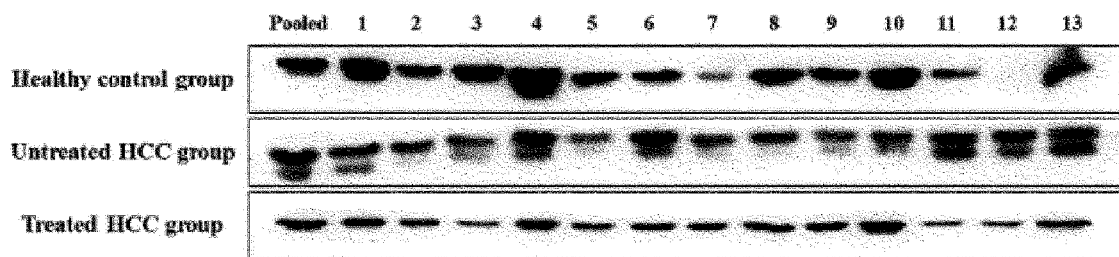
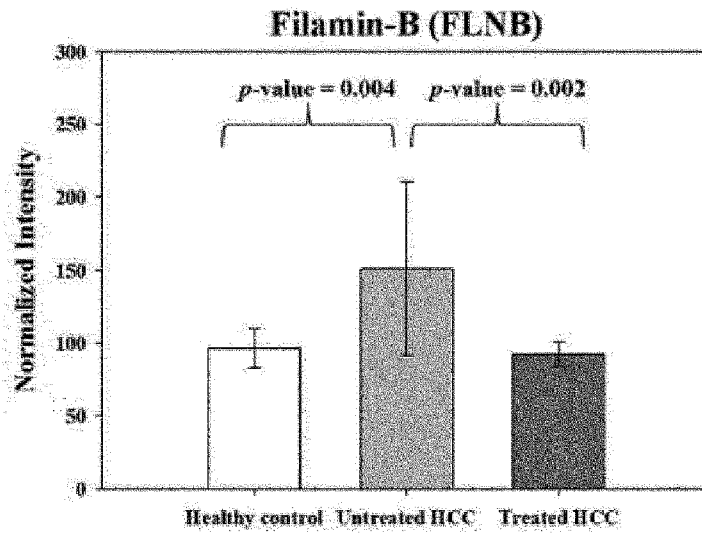
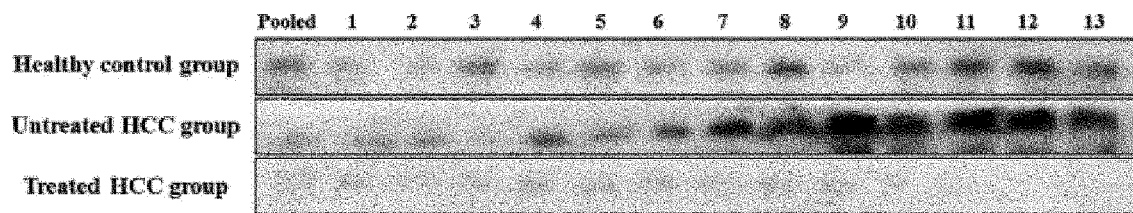
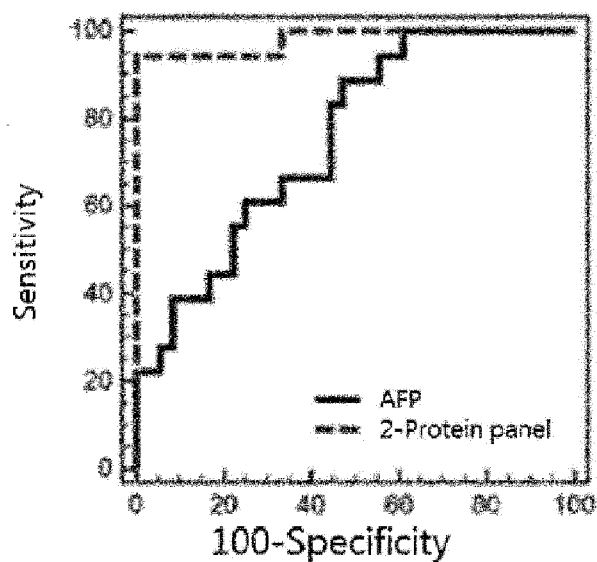


FIG. 4D



**Healthy control group vs. Untreated HCC group**



Classifier : AFP(AUC=0.756)

	Healthy control group(36)	Untreated HCC group(18)
Healthy control group(36)	29	7
Untreated HCC group(18)	10	8

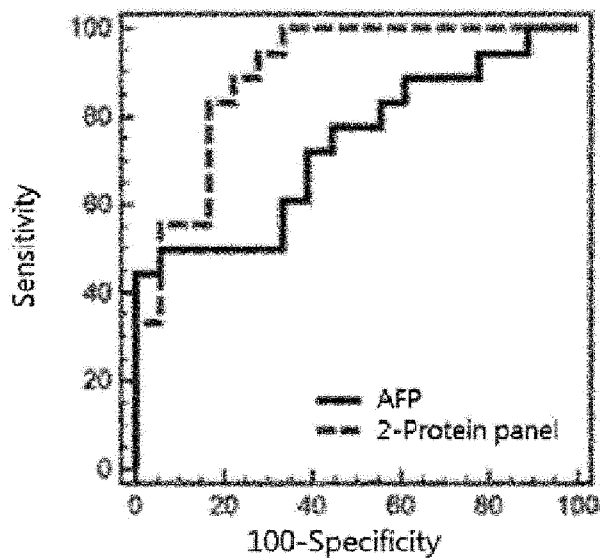
\* Error rate : 0.31

Classifier : 2-Protein panel (AUC=0.981)

	Healthy control group(36)	Untreated HCC group(18)
Healthy control group(36)	36	0
Untreated HCC group(18)	1	17

FIG. 5A \* Error rate : 0.02

**Untreated HCC group vs. Treated HCC group**



Classifier : AFP(AUC=0.735)

	Untreated HCC group(18)	Treated HCC group(18)
<b>Untreated HCC group(18)</b>	12	6
<b>Treated HCC group(18)</b>	9	9

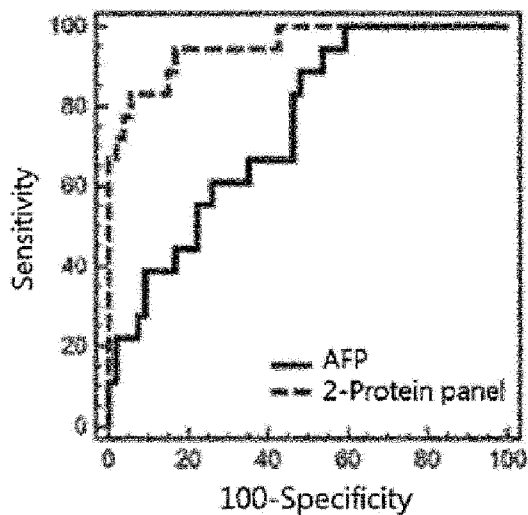
\* Error rate : 0.42

Classifier : 2-protein panel (AUC=0.895)

	Untreated HCC group(18)	Treated HCC group(18)
<b>Untreated HCC group(18)</b>	15	3
<b>Treated HCC group(18)</b>	7	11

FIG. 5B \* Error rate : 0.28

**Healthy control group + treated HCC group  
vs. Untreated HCC group**



Classifier : AFP (AUC=0.749)

	Healthy control group + Treated HCC group(54)	Untreated HCC group(18)
Healthy control group +Treated HCC group(54)	50	4
Untreated HCC group(18)	14	4

✘ Error rate : 0.25

Classifier : 2-protein panel (AUC=0.953)

	Healthy control group + Treated HCC group(54)	Untreated HCC group(18)
Healthy control group +Treated HCC group(54)	50	4
Untreated HCC group(18)	3	15

✘ Error rate : 0.10

FIG. 5C

FIG. 6A

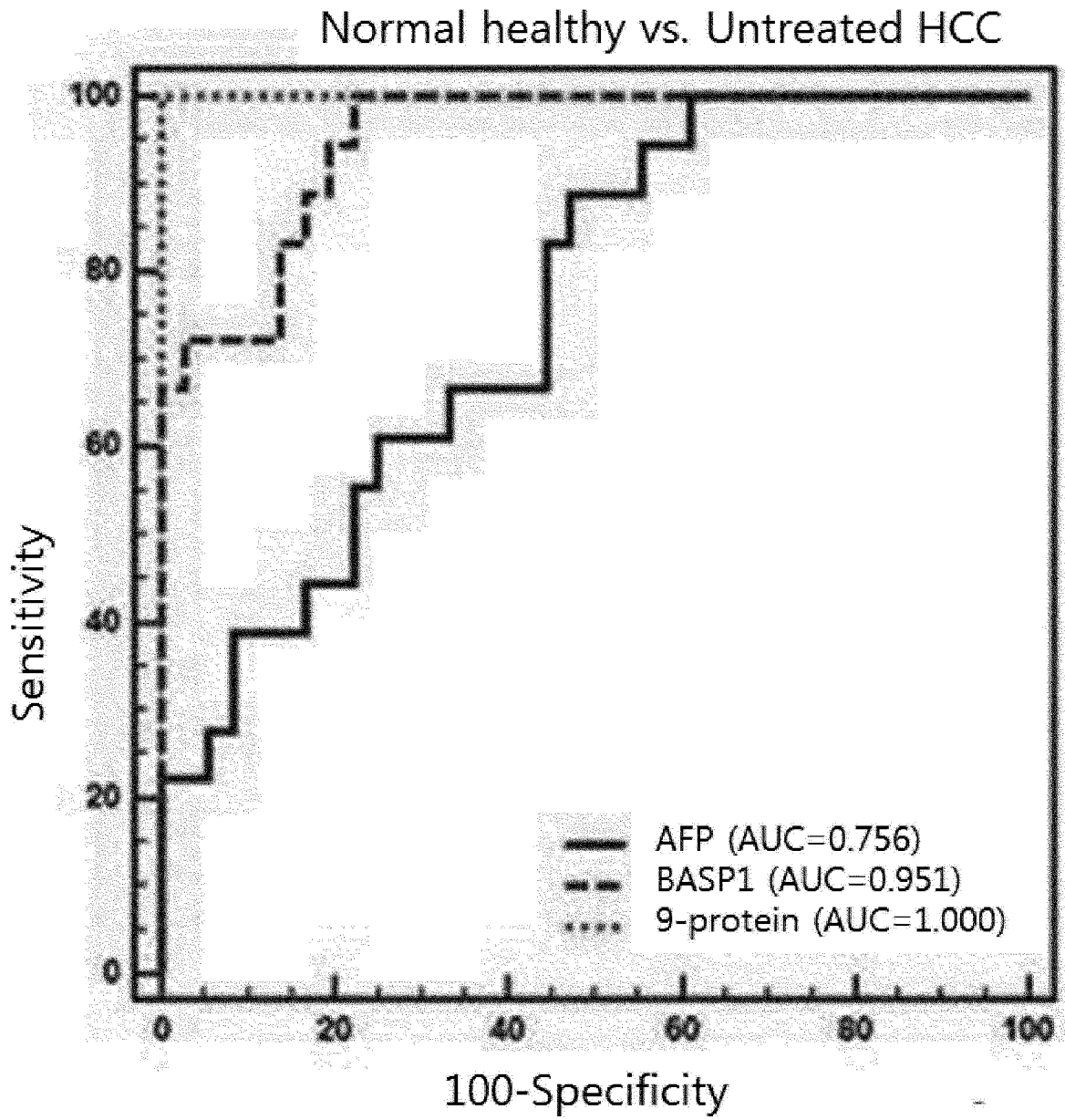


FIG. 6B

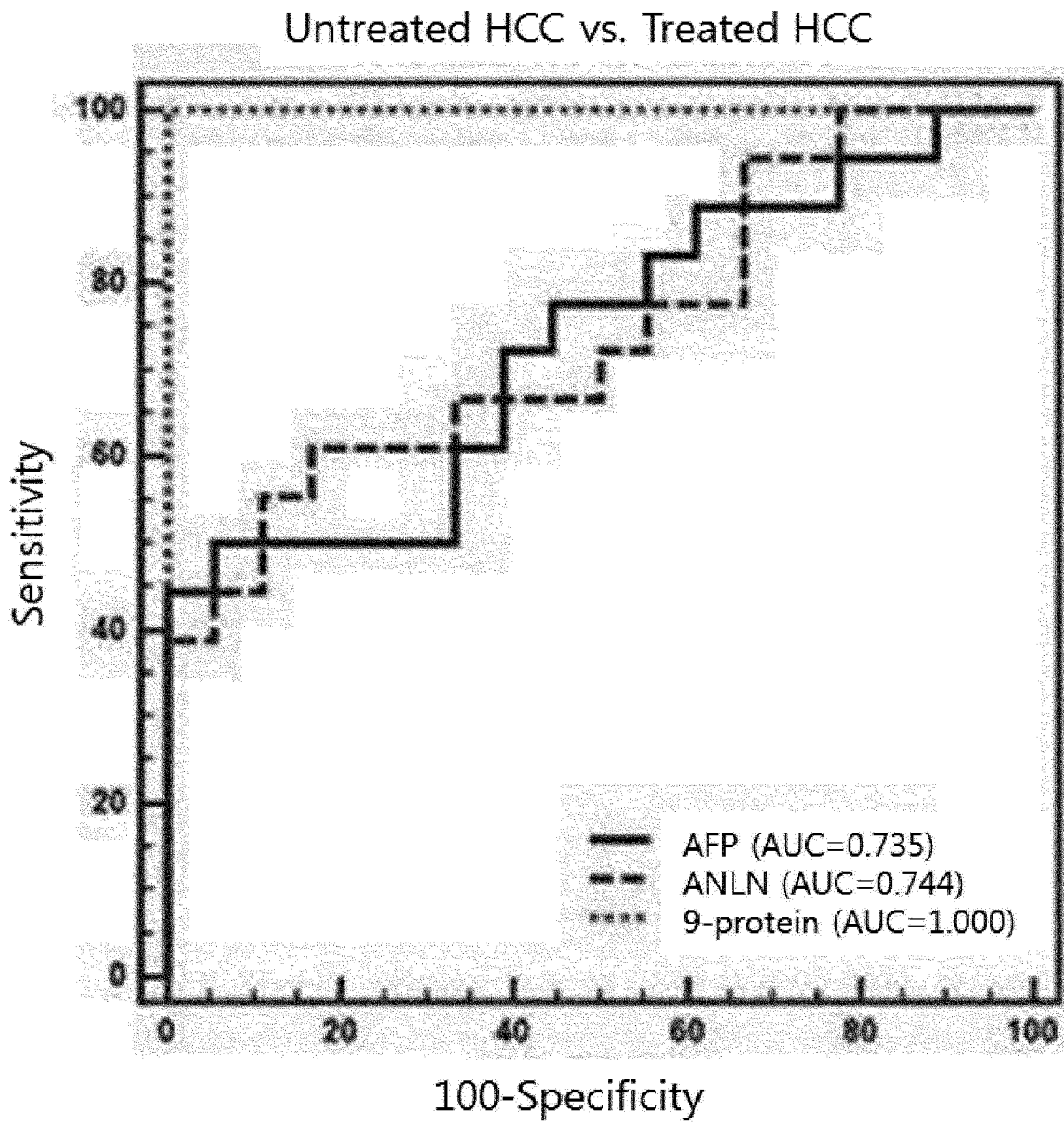


FIG. 6C

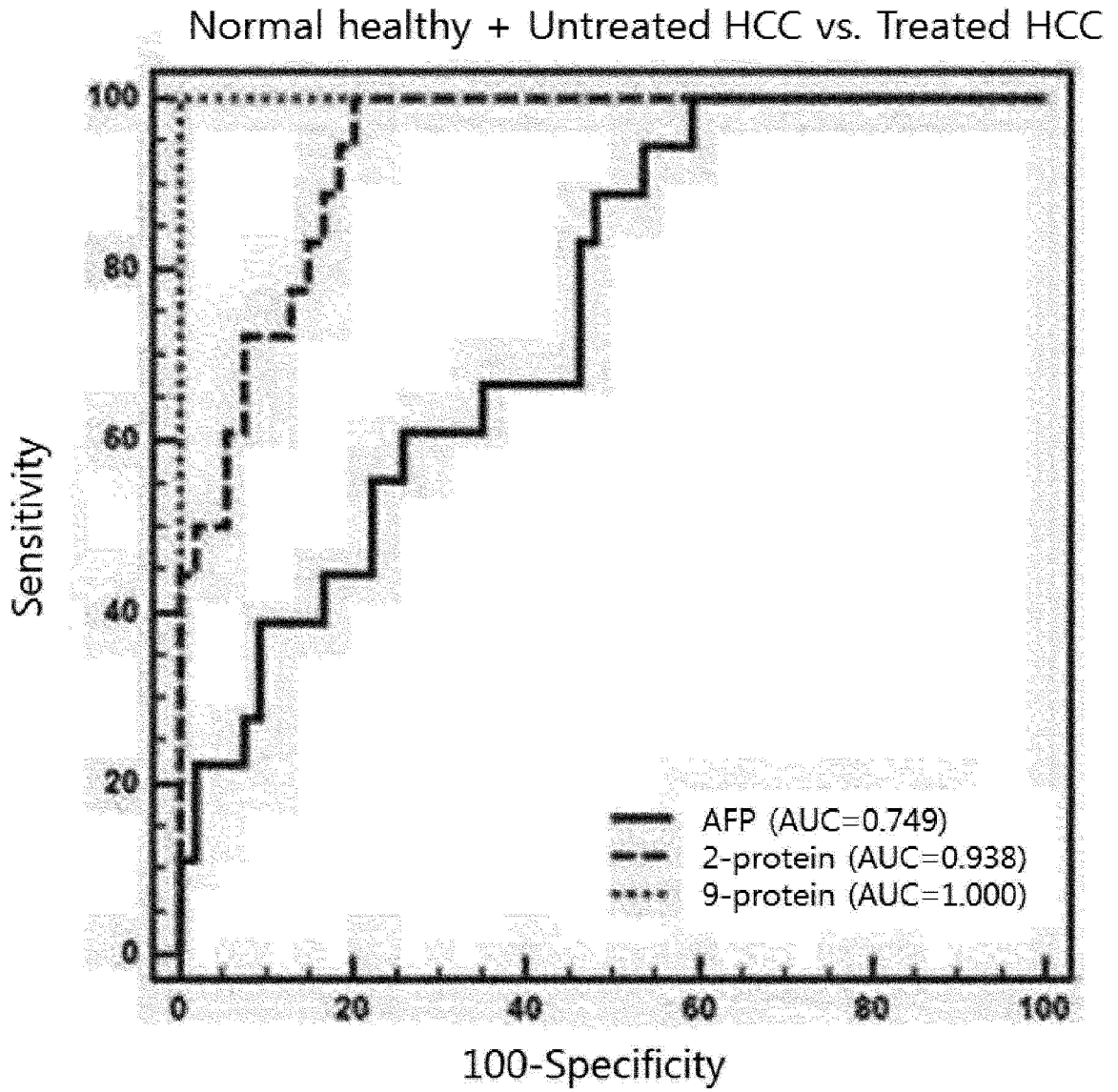


FIG. 6D

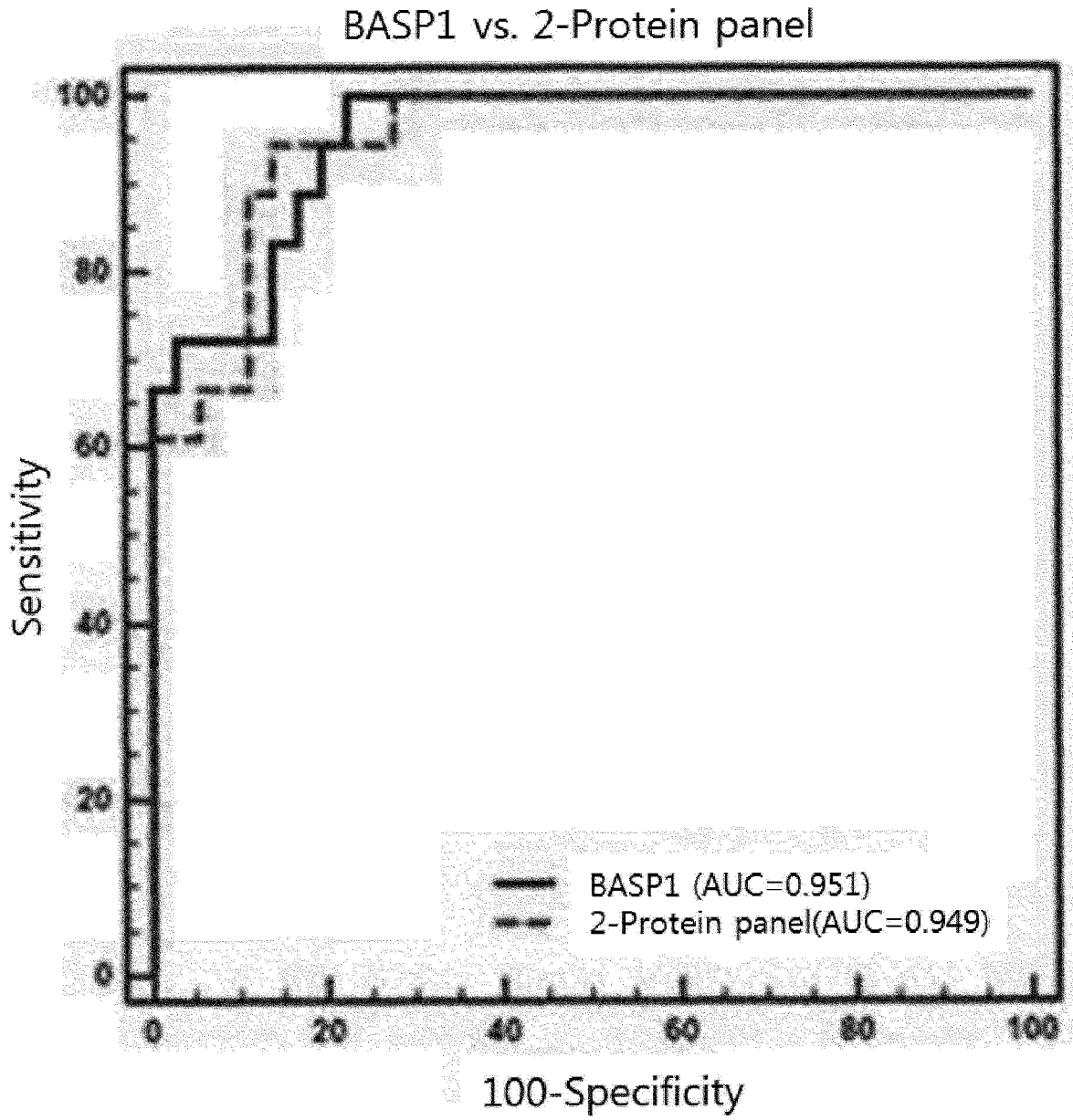


FIG. 6E

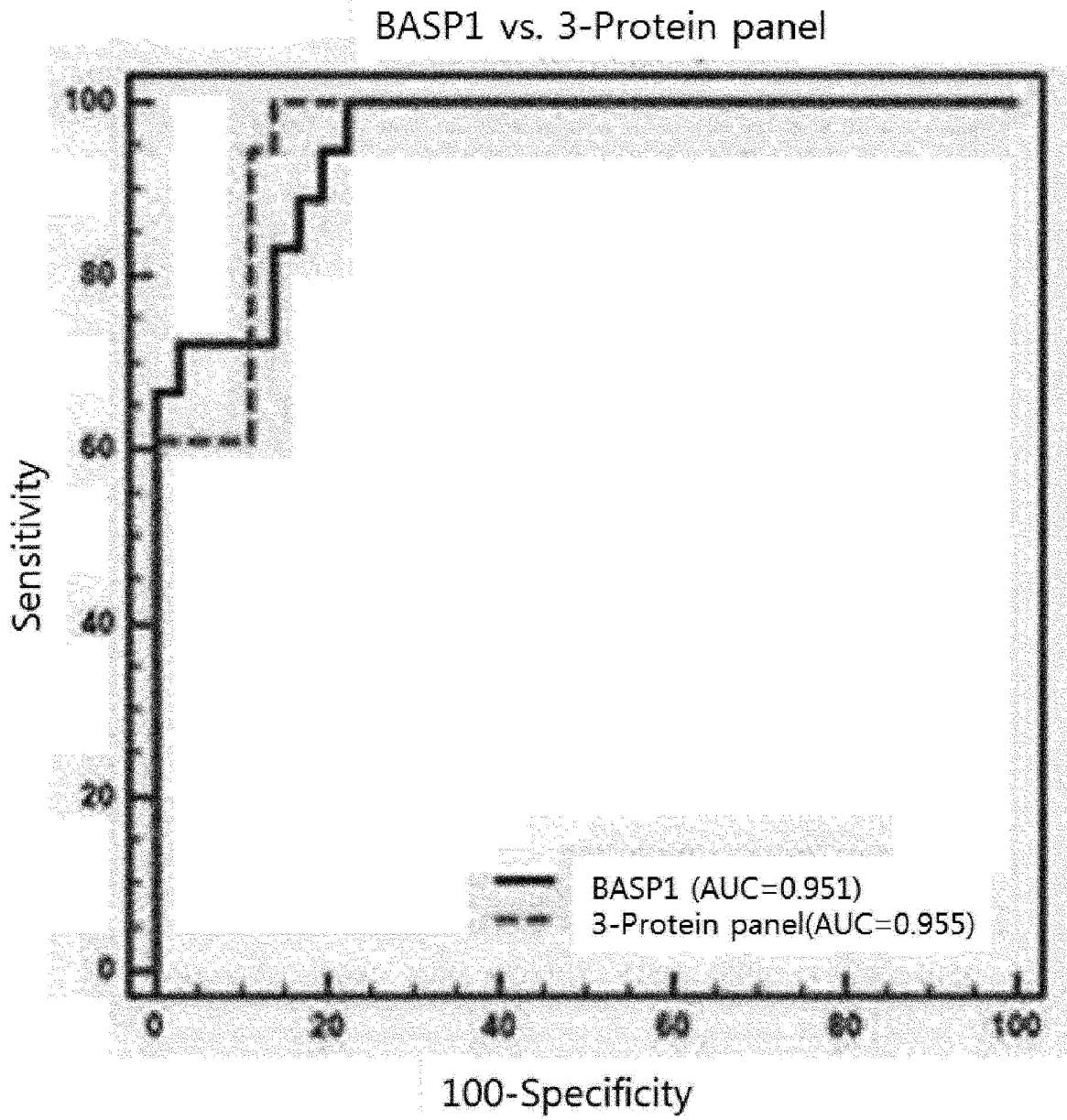


FIG. 6F

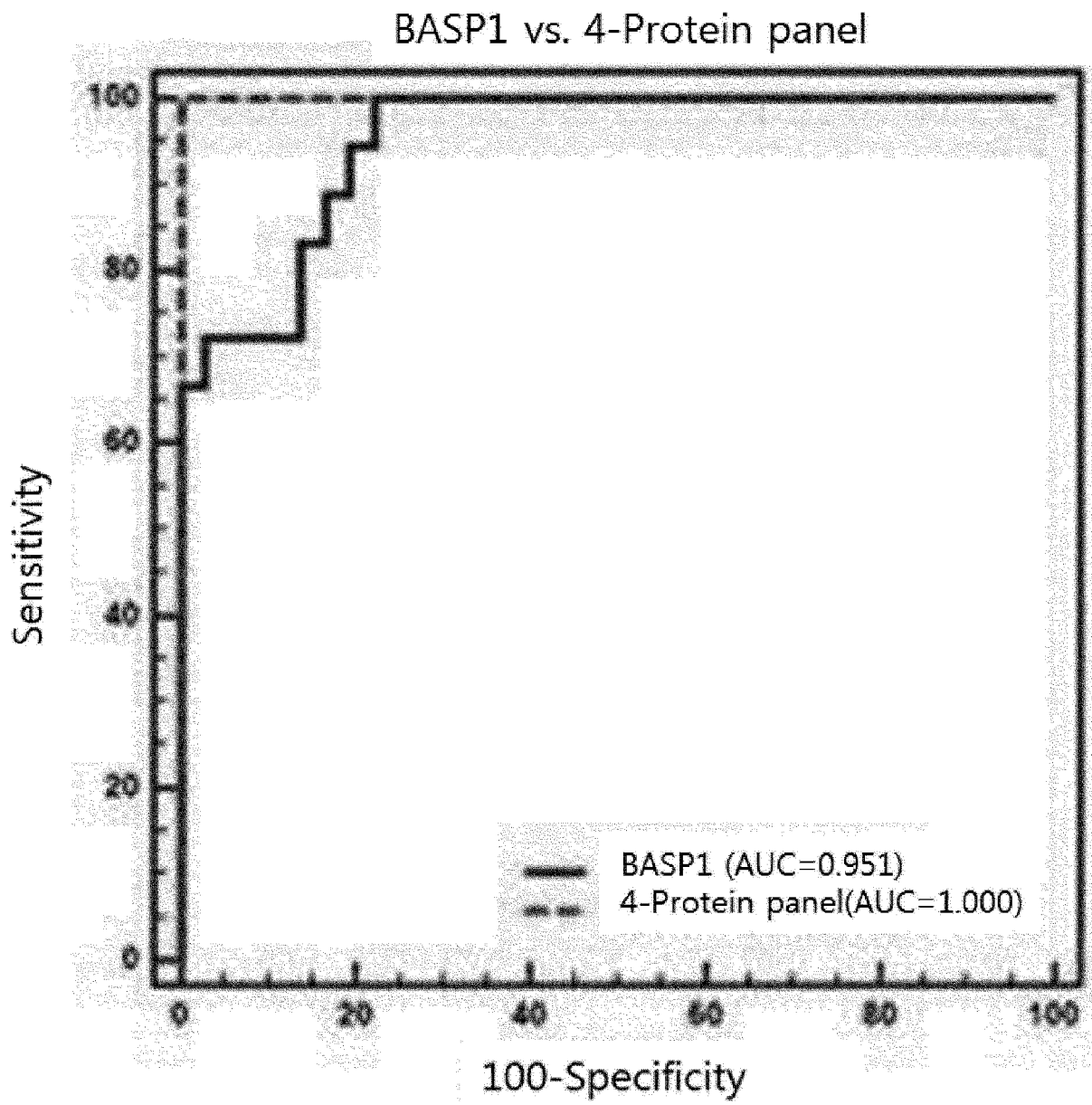


FIG. 6G

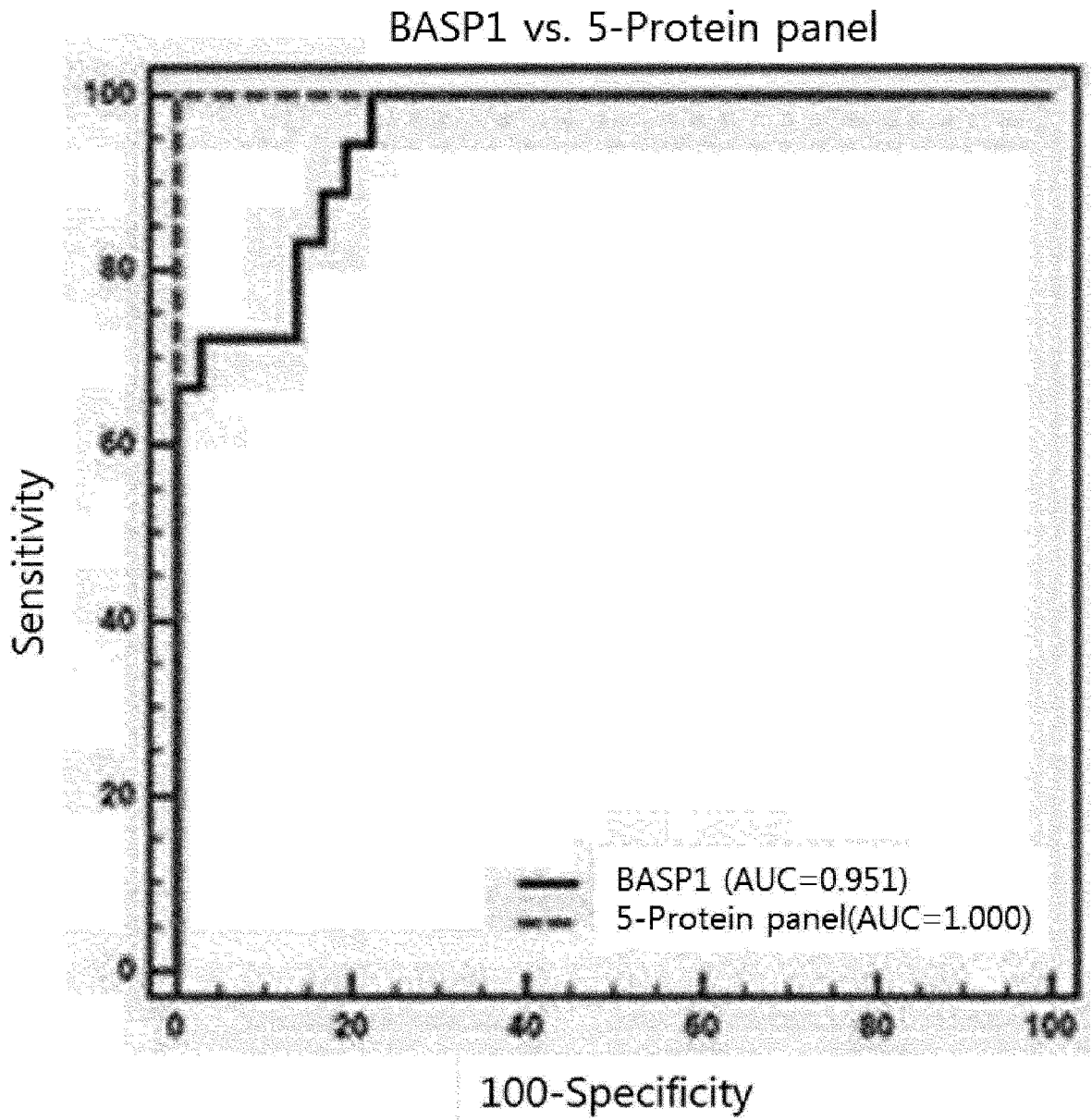


FIG. 6H

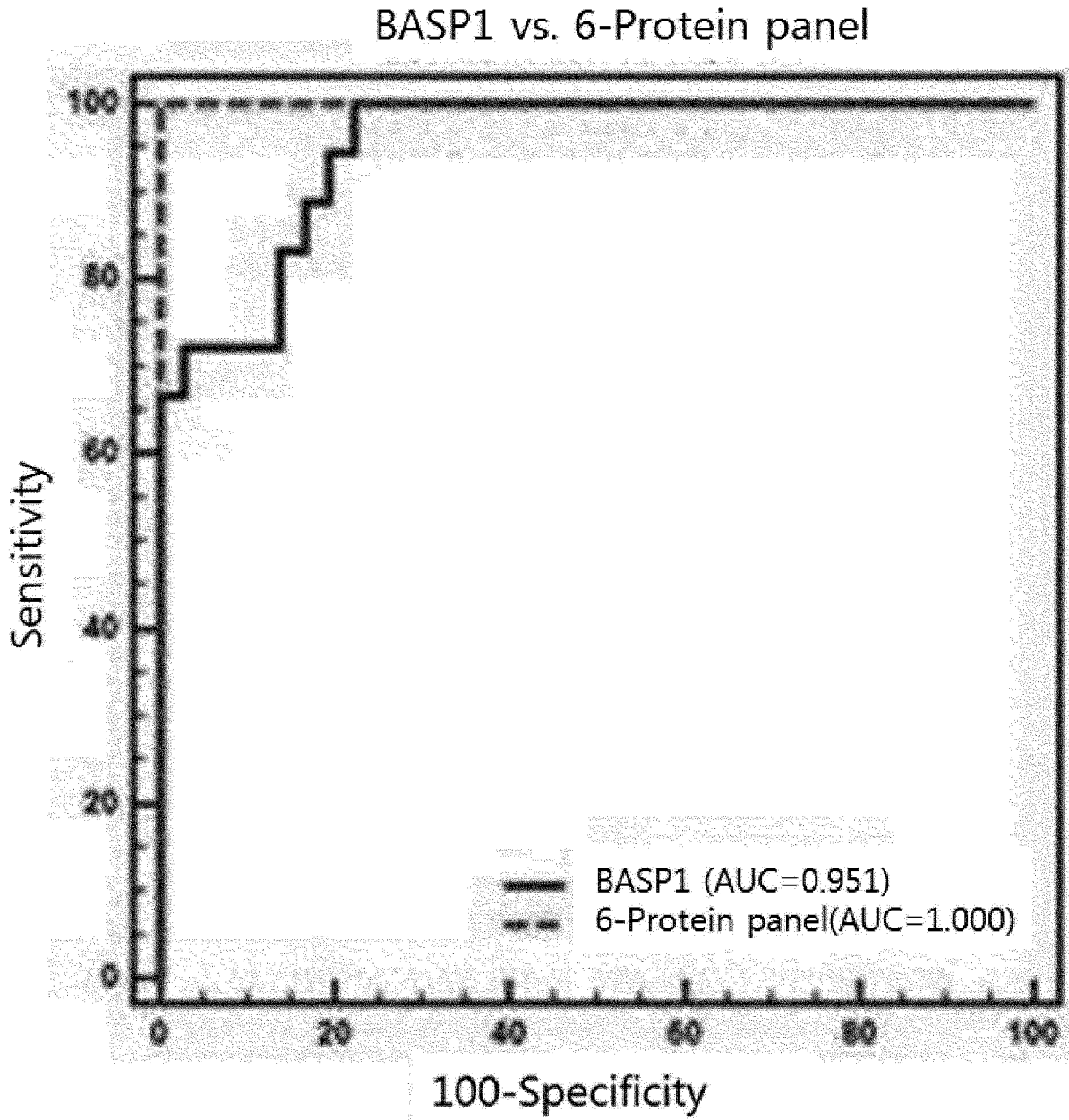


FIG. 6I

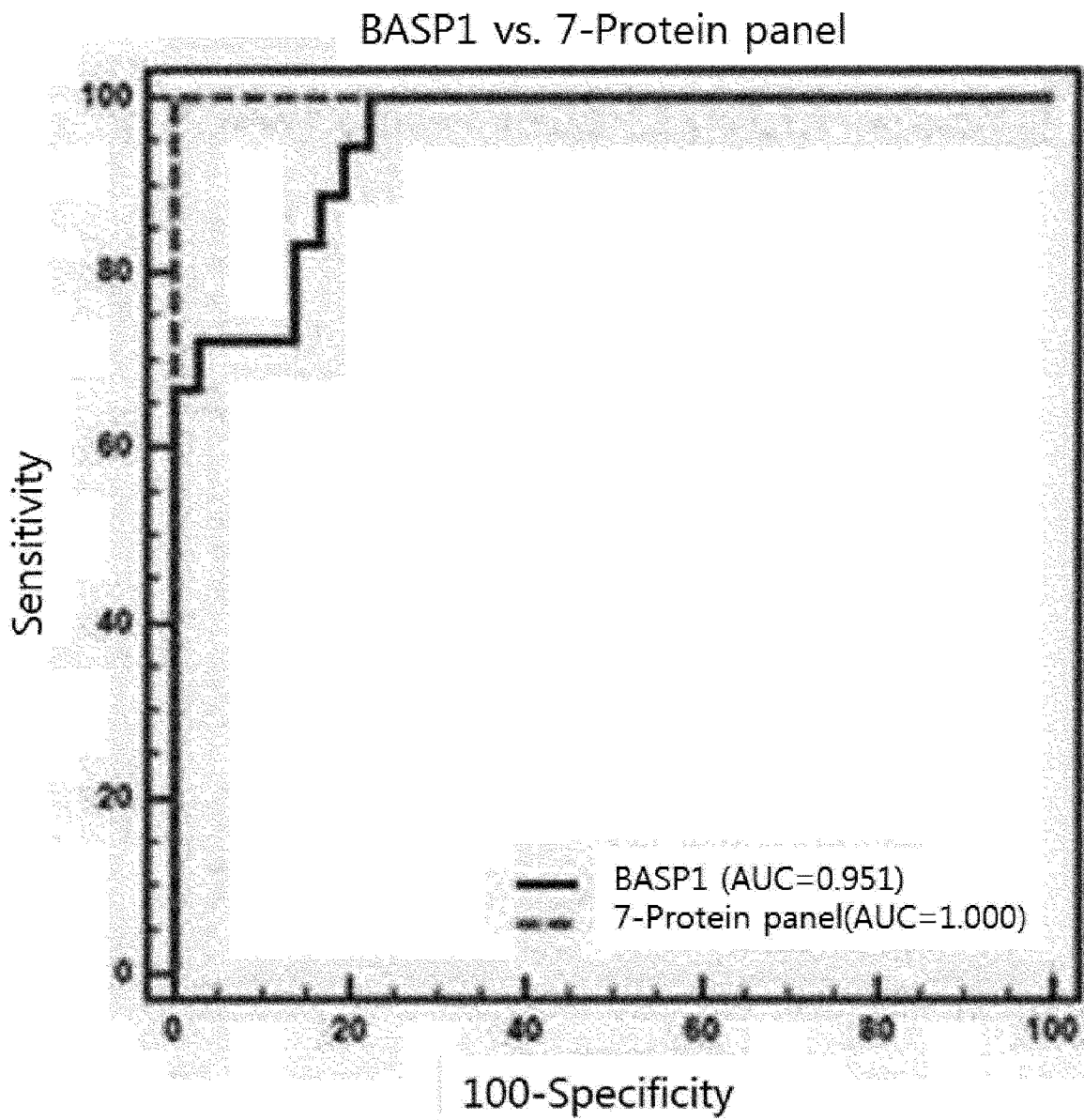
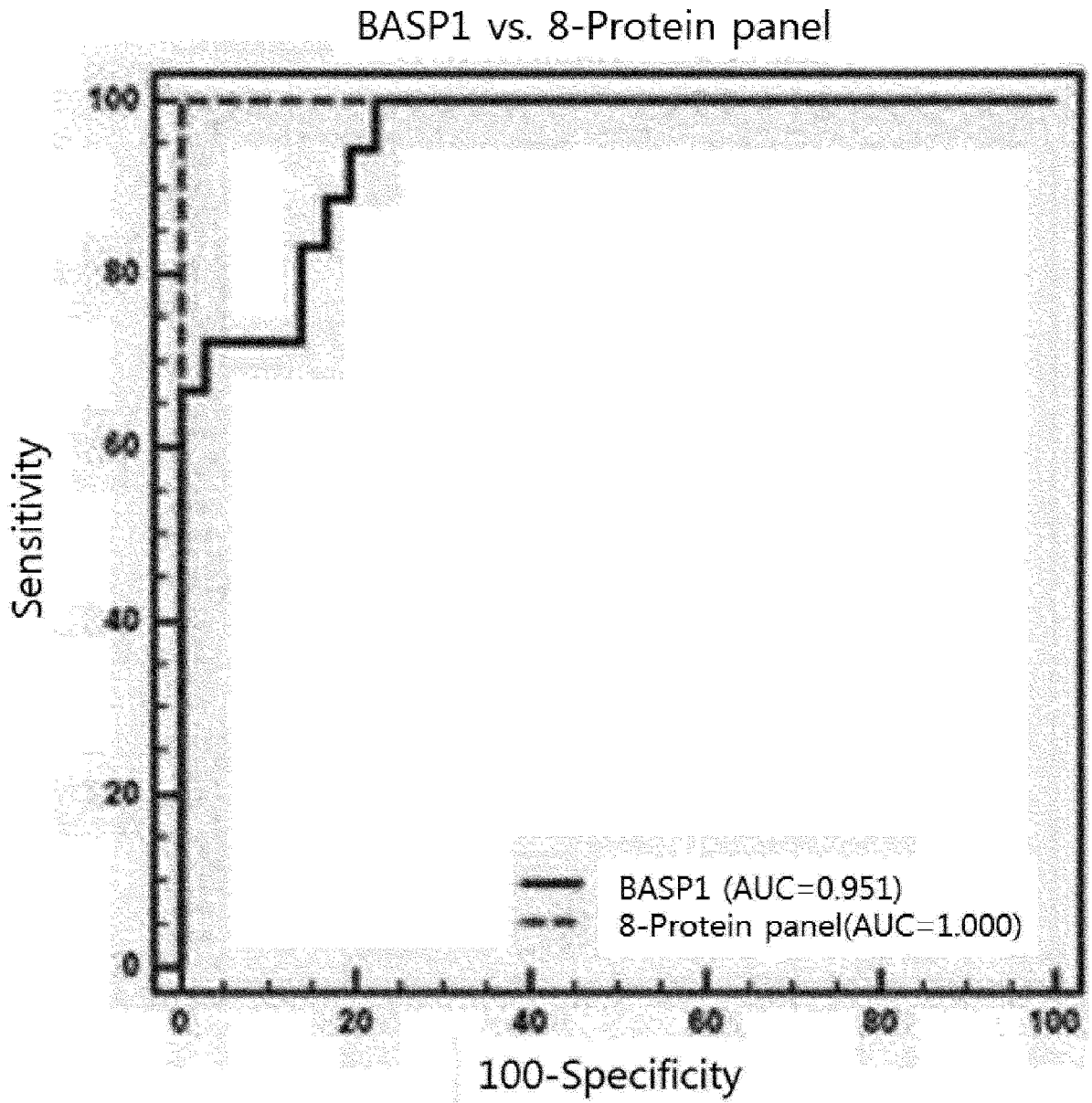


FIG. 6J



INTERNATIONAL SEARCH REPORT

International application No.

PCT/KR2014/002209

5	<p>A. CLASSIFICATION OF SUBJECT MATTER</p> <p><i>G01N 33/574(2006.01)</i>, <i>G01N 33/53(2006.01)</i>, <i>G01N 33/48(2006.01)</i></p> <p>According to International Patent Classification (IPC) or to both national classification and IPC</p>																												
	<p>B. FIELDS SEARCHED</p> <p>Minimum documentation searched (classification system followed by classification symbols)</p> <p>G01N 33/574; G01N 27/62; C07K 16/32; C12Q 1/37; C12N 15/09; C12N 15/11; G01N 33/53; A01K 67/027; C12Q 1/68; G01N 33/48</p> <p>Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched                  Korean Utility models and applications for Utility models: IPC as above                  Japanese Utility models and applications for Utility models: IPC as above</p> <p>Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)                  eKOMPASS (KIPO internal) &amp; Keywords: ACADVL, ANLN, BASP1, MTHFD1, CAPN1, C4A, FLNB, PABPC1, liver cancer marker, liver cancer diagnosis</p>																												
10																													
15																													
	<p>C. DOCUMENTS CONSIDERED TO BE RELEVANT</p> <table border="1"> <thead> <tr> <th style="width: 10%;">Category*</th> <th style="width: 70%;">Citation of document, with indication, where appropriate, of the relevant passages</th> <th style="width: 20%;">Relevant to claim No.</th> </tr> </thead> <tbody> <tr> <td>X</td> <td>JP 2006-308533 A (MCBI LTD et al.) 09 November 2006 See claims 1, 6, 8, 10.</td> <td>1,3-9,18</td> </tr> <tr> <td>Y</td> <td></td> <td>10</td> </tr> <tr> <td>A</td> <td></td> <td>2</td> </tr> <tr> <td>Y</td> <td>KR 10-2012-0021518 A (KIM, Hyun Kee) 09 March 2012 See claims 1, 2.</td> <td>10</td> </tr> <tr> <td>X</td> <td>WO 2009-105154 A2 (THE JACKSON LABORATORY) 27 August 2009 See claims 1, 13, 38.</td> <td>1</td> </tr> <tr> <td>X</td> <td>US 2003-0087316 A1 (HUGLI, Tony E. et al.) 08 May 2003 See claims 1, 15, 16, 24.</td> <td>1</td> </tr> <tr> <td>A</td> <td>EP 2126121 B1 (ROCHE DIAGNOSTICS GMBH et al.) 27 April 2011 See abstract.</td> <td>1-10,18</td> </tr> <tr> <td>A</td> <td>JP 2009-034092 A (NATIONAL YANG MING UNIV) 19 February 2009 See abstract.</td> <td>1-10,18</td> </tr> </tbody> </table>		Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.	X	JP 2006-308533 A (MCBI LTD et al.) 09 November 2006 See claims 1, 6, 8, 10.	1,3-9,18	Y		10	A		2	Y	KR 10-2012-0021518 A (KIM, Hyun Kee) 09 March 2012 See claims 1, 2.	10	X	WO 2009-105154 A2 (THE JACKSON LABORATORY) 27 August 2009 See claims 1, 13, 38.	1	X	US 2003-0087316 A1 (HUGLI, Tony E. et al.) 08 May 2003 See claims 1, 15, 16, 24.	1	A	EP 2126121 B1 (ROCHE DIAGNOSTICS GMBH et al.) 27 April 2011 See abstract.	1-10,18	A	JP 2009-034092 A (NATIONAL YANG MING UNIV) 19 February 2009 See abstract.	1-10,18
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.																											
X	JP 2006-308533 A (MCBI LTD et al.) 09 November 2006 See claims 1, 6, 8, 10.	1,3-9,18																											
Y		10																											
A		2																											
Y	KR 10-2012-0021518 A (KIM, Hyun Kee) 09 March 2012 See claims 1, 2.	10																											
X	WO 2009-105154 A2 (THE JACKSON LABORATORY) 27 August 2009 See claims 1, 13, 38.	1																											
X	US 2003-0087316 A1 (HUGLI, Tony E. et al.) 08 May 2003 See claims 1, 15, 16, 24.	1																											
A	EP 2126121 B1 (ROCHE DIAGNOSTICS GMBH et al.) 27 April 2011 See abstract.	1-10,18																											
A	JP 2009-034092 A (NATIONAL YANG MING UNIV) 19 February 2009 See abstract.	1-10,18																											
20																													
25																													
30																													
35																													
40	<p><input type="checkbox"/> Further documents are listed in the continuation of Box C.      <input checked="" type="checkbox"/> See patent family annex.</p> <p>* Special categories of cited documents:</p> <p>"A" document defining the general state of the art which is not considered to be of particular relevance      "I" 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</p> <p>"E" earlier application or patent but published on or after the international filing date      "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</p> <p>"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)      "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art</p> <p>"O" document referring to an oral disclosure, use, exhibition or other means      "G" document member of the same patent family</p> <p>"P" document published prior to the international filing date but later than the priority date claimed</p>																												
45																													
50	<table border="1"> <tr> <td style="width: 50%;">Date of the actual completion of the international search</td> <td style="width: 50%;">Date of mailing of the international search report</td> </tr> <tr> <td style="text-align: center;">16 JULY 2014 (16.07.2014)</td> <td style="text-align: center;"><b>17 JULY 2014 (17.07.2014)</b></td> </tr> </table>		Date of the actual completion of the international search	Date of mailing of the international search report	16 JULY 2014 (16.07.2014)	<b>17 JULY 2014 (17.07.2014)</b>																							
Date of the actual completion of the international search	Date of mailing of the international search report																												
16 JULY 2014 (16.07.2014)	<b>17 JULY 2014 (17.07.2014)</b>																												
55	<table border="1"> <tr> <td style="width: 50%;">                 Name and mailing address of the ISA/KR                  Korean Intellectual Property Office                  Government Complex-Daejeon, 189 Seonsa-ro, Daejeon 302-701,                  Republic of Korea                  Facsimile No. 82-42-472-7140             </td> <td style="width: 50%;">                 Authorized officer                   Telephone No.             </td> </tr> </table>		Name and mailing address of the ISA/KR Korean Intellectual Property Office Government Complex-Daejeon, 189 Seonsa-ro, Daejeon 302-701, Republic of Korea Facsimile No. 82-42-472-7140	Authorized officer  Telephone No.																									
Name and mailing address of the ISA/KR Korean Intellectual Property Office Government Complex-Daejeon, 189 Seonsa-ro, Daejeon 302-701, Republic of Korea Facsimile No. 82-42-472-7140	Authorized officer  Telephone No.																												

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/KR2014/002209

**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.: **11-17**  
because they relate to subject matter not required to be searched by this Authority, namely:  
Claims 11-17 pertain to a method for treatment of the human body by surgery or therapy, or to a diagnostic method, and thus pertain to subject matter on which the International Searching Authority is not required to carry out an international search under the provisions of PCT Article 17(2)(a)(i) and PCT Rule 39.1(iv).
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:

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 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.:

**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.

EP 2 977 760 A1

INTERNATIONAL SEARCH REPORT  
Information on patent family members

International application No.

PCT/KR2014/002209

5  
10  
15  
20  
25  
30  
35  
40  
45  
50  
55

Patent document cited in search report	Publication date	Patent family member	Publication date
JP 2006-308533 A	09/11/2006	JP 4801930 B2	26/10/2011
KR 10-2012-0021518 A	09/03/2012	NONE	
WO 2009-105154 A2	27/08/2009	WO 2009-105154 A3	14/01/2010
US 2003-0087316 A1	08/05/2003	NONE	
EP 2126121 B1	27/04/2011	AT 507310 T	15/05/2011
		CA 2679290 A1	12/09/2008
		CA 2679290 C	29/04/2014
		CN 101675171 A	17/03/2010
		DE 602008006533 D1	09/06/2011
		EP 2126121 A2	02/12/2009
		ES 2365829 T3	11/10/2011
		JP 2008-245635 A	16/10/2008
		JP 5378687 B2	25/12/2013
		US 2010-0304372 A1	02/12/2010
		WO 2008-107134 A2	12/09/2008
		WO 2008-107134 A3	20/11/2008
JP 2009-034092 A	19/02/2009	CN 101407815 A	15/04/2009
		CN 101407815 B	16/03/2011
		EP 2020443 A2	04/02/2009
		EP 2020443 A3	20/01/2010
		EP 2471913 A1	04/07/2012
		JP 2011-160805 A	25/08/2011
		JP 4855433 B2	18/01/2012
		TW 200906441 A	16/02/2009
		TW 201141525 A	01/12/2011
		TW 1345983 I	01/08/2011
		US 2009-0035290 A1	05/02/2009
		US 2009-0148428 A1	11/06/2009
		US 2010-0279897 A1	04/11/2010
		US 7759542 B2	20/07/2010

## REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

## Patent documents cited in the description

- US 2010030437 A [0005]
- KR 1161789 [0006]
- US 5599695 A [0042]
- US 5556752 A [0042]
- US 5631734 A [0042]
- US 6308170 B [0059]
- US 6228575 B [0059]

## Non-patent literature cited in the description

- **STEFANIUK P et al.** *World J Gastroenterol*, 2010, vol. 16, 418-424 [0002]
- **LEE et al.** *Nature Medicine*, 2006, vol. 12, 410-416 [0020]
- **ROESSLER et al.** *Cancer Research*, 2010, vol. 70, 10202-10212 [0020]
- **SATOW et al.** *Clinical Cancer Research*, 2010, vol. 16, 2518-2528 [0020]
- **CHEN et al.** *Hepatology*, 2010, vol. 52, 1690-1701 [0020]
- **CHOCHI et al.** *J. of Pathology*, 2009, vol. 217, 677-684 [0020]
- **HASHIMOTO et al.** *Modern Pathology*, 2004, vol. 17, 617-622 [0020]
- **ACEVEDO et al.** *Cancer Research*, 2008, vol. 68, 2641-2651 [0020]
- **HERNANDEZ-VARGAS et al.** *PLoS ONE*, 2010, vol. 5 (3), e9749 [0020]
- **ZHU.** *Cell Research*, 2005, vol. 15, 272-280 [0020]
- **COLYER et al.** *J Chromatogr A.*, 1997, vol. 781 (1-2), 271-6 [0033]
- **PETRICOIN et al.** *Lancet*, 2002, vol. 359, 572-77 [0033]
- **CHAN-HUI et al.** *Clinical Immunology*, 2004, vol. 111, 162-174 [0033]
- **MONTAGNE et al.** *Eur J Clin Chem Clin Biochem.*, 1992, vol. 30, 217-22 [0033]
- **KIM et al.** *J Proteome Res.*, 2010, vol. 9, 689-99 [0034]
- **ANDERSON, L et al.** *Mol Cell Proteomics*, 2006, vol. 5, 573-88 [0034]
- **GILLETTE et al.** *Nature Methods*, 2013, vol. 10, 28-34 [0034]
- Enzyme Immunoassay. CRC Press, 1980 [0037]
- Enzyme-linked immunosorbent assay(ELISA). **GAASTRA, W.** *Methods in Molecular Biology*. Humana Press, 1984, vol. 1 [0037]
- **HAN, H. et al.** *Cancer Res.*, 2002, vol. 62, 2890-6 [0039]
- **SCHENA et al.** *Proc Natl Acad Sci USA.*, 1996, vol. 93 (20), 10614-9 [0042]
- **SCHENA et al.** *Science*, 1995, vol. 270 (5235), 467-70 [0042]
- **BEN-DOR et al.** *J. Comput. Biol.*, 2007, vol. 7, 559-83 [0060]
- **NGUYEN et al.** *Bioinformatics*, 2002, vol. 18, 39-50 [0060]
- **WANG et al.** *BMC Bioinformatics*, 2003, vol. 4, 60 [0060]
- **LIU et al.** *Genome Inform. Ser. Workshop Genome Inform.*, 2001, vol. 12, 14-23 [0060]
- **YEANG et al.** *Bioinformatics*, 2001, vol. 17 (1), 316-22 [0060]
- **XIONG.** *Biotechniques*, 2000, vol. 29 (6), 1264-8, 1270 [0060]
- **RUCZINSKI.** *Journal of Computational and Graphical Statistics*, 2003, vol. 12, 475-512 [0061]
- **TIBSHIRANI.** *PNAS*, 2002, vol. 99, 6567-72 [0061]
- **BREIMAN.** *Machine Learning*, 2001, vol. 45, 5-32 [0061]
- **HASTIE.** *The Elements of Statistical Learning*. Springer, 2001 [0061]
- **STERGACHIS AB et al.** *Nat Methods*, 2011, vol. 8, 1041-1043 [0078]
- **KIRKPATRICK DS et al.** *Methods*, 2005, vol. 35, 265-273 [0083]

专利名称(译)	用于诊断肝癌的生物标志物		
公开(公告)号	<a href="#">EP2977760A4</a>	公开(公告)日	2017-02-15
申请号	EP2014767707	申请日	2014-03-17
[标]申请(专利权)人(译)	首尔大学校产学协力团		
申请(专利权)人(译)	首尔国立大学R & DB基金会		
当前申请(专利权)人(译)	首尔国立大学R & DB基金会		
[标]发明人	KIM YOUNGSOO YOON JUNGHWAN KIM HYUNSOO JIN JONGHWA		
发明人	KIM, YOUNGSOO YOON, JUNGHWAN KIM, HYUNSOO JIN, JONGHWA		
IPC分类号	G01N33/574 G01N33/53 G01N33/48 G01N33/50		
CPC分类号	G01N33/57438 G01N33/5079 G01N2333/9015 G01N2333/90206 G01N2333/96413		
优先权	1020130029626 2013-03-20 KR		
其他公开文献	EP2977760A1 EP2977760B1		
外部链接	<a href="#">Espacenet</a>		

#### 摘要(译)

本发明涉及用于肝癌的诊断或预后的生物标志物，其为选自ACADVL，ANLN，BASP1，MTHFD1，CAPN1，C4A，FLNB和PABPC1中的至少一种，以及测定试剂盒，组合物一种用于肝癌的诊断或预后的方法，其包括所述方法。根据本公开的生物标志物以高精度和灵敏度使得能够早期诊断或监测肝细胞癌和有效预后或评估疾病进展。另外，通过使用血液，尿等的非侵入性试验，可以简单且有效地诊断或诊断肝癌，因此本发明在家庭或一般诊所的肝癌的早期检测中将非常有效，通过开发例如，护理点测试(POCT)。