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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
23 February 2017 (23.02.2017)

WIPO | PCT

(10) International Publication Number
WO 2017/028308 A1

(51) International Patent Classification:
G01N 30/88 (2006.01) *G01N 30/00* (2006.01)

(21) International Application Number:
PCT/CN2015/087664

(22) International Filing Date:
20 August 2015 (20.08.2015)

(25) Filing Language: English

(26) Publication Language: English

(71) Applicant: **BGI SHENZHEN** [CN/CN]; Main Building, Beishan Industrial Zone, Yantian District, Shenzhen, Guangdong 518083 (CN).

(72) Inventors: **FENG, Qiang**; Main Building, Beishan Industrial Zone, Yantian District, Shenzhen, Guangdong 518083 (CN). **LIU, Zhipeng**; Main Building, Beishan Industrial Zone, Yantian District, Shenzhen, Guangdong 518083 (CN). **WANG, Jun**; Main Building, Beishan Industrial Zone, Yantian District, Shenzhen, Guangdong 518083 (CN).

(74) Agent: **CCPIT PATENT AND TRADEMARK LAW OFFICE**; 8th Floor, Vantone New World Plaza, 2 Fuchengmenwai Street, Xicheng District, Beijing 100037 (CN).

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

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

Published:

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

WO 2017/028308 A1

(54) Title: BIOMARKERS FOR CORONARY HEART DISEASE

(57) Abstract: Biomarkers and methods for predicting the risk of a disease related to metabolites, in particular coronary heart disease.

BIOMARKERS FOR CORONARY HEART DISEASE

CROSS-REFERENCE TO RELATED APPLICATION

None

FIELD

The present invention relates to plasma biomarkers and methods for predicting the risk of a disease related to metabolites, in particular coronary heart disease.

BACKGROUND

Coronary heart disease (CHD) is the top risk factor in modern society with annual mortality rate overpassing the sum of all types of cancers. The majority of cardiovascular deaths occurrence are related to the extent of people's awareness of their own medical conditions and are due to lackness of in-time treatment as demonstrated by a five-year follow-up study by MaGiCAD cohort (Graninger, D.J. & Mosedale, D.E. Metabolomics in coronary heart disease.Heart.Metab.55, 8–12 (2012), incorporated herein by reference). The challenge for early diagnosis and prevention of CHD lies in the lackness of reliable non-invasive biomarkers. The "gold standard" for diagnosis of CHD is still coronary angiography which is invasive and accompanied by many deadly side effects, this limited the large population screening and the CHD risk prediction at early stage.

Many researches point the fatty acids play important roles in the heart metabolism; they are predominant substrates, accounting for 60-90% cardiac ATP synthesis, for cardiac ATP generation by mitochondrial oxidative phosphorylation under normal physiological conditions. Cardiovascular diseases (CVD) like coronary heart disease and cardiac failure undergo a "metabolic shift" as a consequence of both intrinsic and extrinsic perturbations. Increased low-density lipoprotein cholesterol (LDL-C) has previously been considered as one of the major risk factors for CHD. The fact that core defects in cardiovascular disease are lipids metabolism (Fernandez, C. et al. Plasma lipid composition and risk of developing cardiovascular disease. PLoS ONE 8, e71846 (2013), incorporated herein by reference) makes

metabolomics a particularly promising method to study these types of diseases.

Metabolomics is an innovative and high-throughput bioanalytical method aiming to identify and quantify small molecules (molecular weight less than 1500 Daltons) present in any biological system or any specific physiological state. Two major analytical techniques, nuclear magnetic resonance (NMR) and mass spectrometry (MS), have been widely used in endogenous compounds measurement at an exponential increasing rate in last decade. MS-based techniques have made rapid progress and have been used more frequently compared with NMR since 2005 because of the following advantages: higher sensitivity, more coverage of the metabolome, improved metabolites identification and discrimination capability, and modularity to perform compound-class-specific analysis (Griffiths, W.J. et al. Targeted metabolomics for biomarker discovery. Angew.Chem. Int. Ed. 49, 5426–5445 (2010), incorporated herein by reference). MS is mostly used in conjunction with chromatography, such as gas chromatography mass spectrometry (GC–MS) and liquid chromatography mass spectrometry (LC–MS).

Contents of the Invention

The present invention aims to provide a biomarker composition and a method for evaluation of the risk of CHD or diagnosis or early diagnosis of CHD, and specifically comprises the following aspects.

The first aspect of the invention relates to a biomarker composition, which comprises one or more selected from the group consisting of:

Biomarker 1, for which m/z is 518.32±1.00, retention time (RT) is 12.71±1.00min;

Biomarker 2, for which m/z is 480.34±1.00, retention time (RT) is 13.48±1.00min;

Biomarker 3, for which m/z is 482.32±1.00, retention time (RT) is 12.92±1.00min;

Biomarker 4, for which m/z is 496.33±1.00, retention time (RT) is 13.19±1.00min;

Biomarker 5, for which m/z is 468.3±1.00, retention time (RT) is 12.69±1.00min;

Biomarker 6, for which m/z is 494.32±1.00, retention time (RT) is 12.82±1.00min;

Biomarker 7, for which m/z is 524.36±1.00, retention time (RT) is 13.69±1.00min;

Biomarker 8, for which m/z is 516.31±1.00, retention time (RT) is 13.21±1.00min;

Biomarker 9, for which m/z is 590.31±1.00, retention time (RT) is 12.86±1.00min;

Biomarker 10, for which m/z is 546.35±1.00, retention time (RT) is 14.22±1.00min;

Biomarker 11, for which m/z is 570.35±1.00, retention time (RT) is 13.05±1.00min;

Biomarker 12, for which m/z is 518.32±1.00, retention time (RT) is 13.19±1.00min;

Biomarker 13, for which m/z is 524.36±1.00, retention time (RT) is 14.27±1.00min;

Biomarker 14, for which m/z is 522.35±1.00, retention time (RT) is 13.33±1.00min;

Biomarker 15, for which m/z is 181.07±1.00, retention time (RT) is 9.05±1.00min;

Biomarker 16, for which m/z is 544.33±1.00, retention time (RT) is 13.34±1.00min;

Biomarker 17, for which m/z is 175.11±1.00, retention time (RT) is 1.83±1.00min; and

Biomarker 18, for which m/z is 324.04±1.00, retention time (RT) is 9.33±1.00min.

In an embodiment of the invention, Biomarker 1 ~ 18 are shown in Table 1.

In an embodiment of the invention, the biomarker composition comprises Biomarker 1 and one or more selected from the group consisting of Biomarker 2 ~ Biomarker 18.

In an embodiment of the invention, the biomarker composition comprises Biomarker 2 and one or more selected from the group consisting of Biomarker 1 and Biomarker 3 ~ 18.

In an embodiment of the invention, the biomarker composition comprises Biomarker 3 and one or more selected from the group consisting of Biomarker 1 ~ 2 and Biomarker 4 ~ 18.

In an embodiment of the invention, the biomarker composition comprises Biomarker 4 and one or more selected from the group consisting of Biomarker 1 ~ 3 and Biomarker 5 ~ 18.

In an embodiment of the invention, the biomarker composition comprises Biomarker 5 and one or more selected from the group consisting of Biomarker 1 ~ 4 and Biomarker 6 ~ 18.

In an embodiment of the invention, the biomarker composition comprises Biomarker 6 and one or more selected from the group consisting of Biomarker 1 ~ 5 and Biomarker 7 ~ 18.

In an embodiment of the invention, the biomarker composition comprises Biomarker 7 and one or more selected from the group consisting of Biomarker 1 ~ 6 and Biomarker 8 ~ 18.

In an embodiment of the invention, the biomarker composition comprises Biomarker 1 ~ 7 and one or more selected from the group consisting of Biomarker 8 ~ 18.

In an embodiment of the invention, the biomarker composition comprises one or more selected from the group consisting of:

Biomarker 1, for which m/z is 518.32±1.00, retention time (RT) is 12.71±1.00min;

Biomarker 2, for which m/z is 480.34±1.00, retention time (RT) is 13.48±1.00min;

Biomarker 3, for which m/z is 482.32±1.00, retention time (RT) is 12.92±1.00min;

Biomarker 4, for which m/z is 496.33±1.00, retention time (RT) is 13.19±1.00min;

Biomarker 5, for which m/z is 468.3±1.00, retention time (RT) is 12.69±1.00min;

Biomarker 6, for which m/z is 494.32±1.00, retention time (RT) is 12.82±1.00min; and

Biomarker 7, for which m/z is 524.36±1.00, retention time (RT) is 13.69±1.00min.

In an embodiment of the invention, the biomarker composition comprises Biomarker 1 and one or more selected from the group consisting of Biomarker 2 ~ Biomarker 7.

In an embodiment of the invention, the biomarker composition comprises Biomarker 2 and one or more selected from the group consisting of Biomarker 1 and Biomarker 3 ~ Biomarker 7.

In an embodiment of the invention, the biomarker composition comprises Biomarker 3 and one or more selected from the group consisting of Biomarker 1 ~ 2 and Biomarker 4 ~ 7.

In an embodiment of the invention, the biomarker composition comprises Biomarker 4 and one or more selected from the group consisting of Biomarker 1 ~ 3 and Biomarker 5 ~ 7.

In an embodiment of the invention, the biomarker composition comprises Biomarker 5 and one or more selected from the group consisting of Biomarker 1 ~ 4 and Biomarker 6 ~ 7.

In an embodiment of the invention, the biomarker composition comprises Biomarker 6 and one or more selected from the group consisting of Biomarker 1 ~ 5 and Biomarker 7.

In an embodiment of the invention, the biomarker composition comprises Biomarker 7 and one or more selected from the group consisting of Biomarker 1 ~ 6.

In an embodiment of the invention, the biomarker composition comprises Biomarker 1 ~ 7.

In an embodiment of the invention, wherein,

Biomarker 1 is LysoPC(18:3(6Z,9Z,12Z));

Biomarker 2 is LysoPC(P-16:0);

Biomarker 3 is LysoPC(15:0);

Biomarker 4 is 1-Palmitoylglycerophosphocholine;

Biomarker 5 is LysoPC(14:0);

Biomarker 6 is LysoPC(16:1(9Z));

Biomarker 7 is LysoPC(0:0/18:0);

Biomarker 8 is LysoPC(18:4(6Z,9Z,12Z,15Z));

Biomarker 9 is LysoPC(22:6(4Z,7Z,10Z,13Z,16Z,19Z));

Biomarker 10 is LysoPC(20:3(5Z,8Z,11Z));

Biomarker 11 is LysoPC(22:5(4Z,7Z,10Z,13Z,16Z));

Biomarker 12 is LysoPC(18:3(9Z,12Z,15Z));

Biomarker 13 is LysoPC(18:0);

Biomarker 14 is 1-Oleoylglycerophosphocholine;

Biomarker 15 is Paraxanthine;

Biomarker 16 is LysoPC(20:4(5Z,8Z,11Z,14Z));

Biomarker 17 is L-Arginine; and/or

Biomarker 18 is N-Acetyl-D-glucosamine 6-phosphate.

In an embodiment of the invention, the biomarker composition comes from a blood, plasma or serum sample.

In an embodiment of the invention, wherein, as compared with a reference value, the decrease of one or more Biomarkers selected from the group consisting of Biomarker 1-16, preferably, from the group consisting of Biomarker 1-7, indicates that the subject is in the risk of CHD or has CHD.

In an embodiment of the invention, wherein, as compared with a reference value, the increase of Biomarker 17 and/or Biomarker 18 indicates that the subject is in the risk of CHD or has CHD.

In an embodiment of the invention, wherein the identification of each of the biomarker composition is carried out by mass spectrometry, preferably, mass spectrometry in conjunction with chromatography, such as gas chromatography mass spectrometry (GC-MS) or liquid chromatography mass spectrometry (LC-MS).

The second aspect of the invention relates to a reagent composition, which comprises the reagents used for detection of each of the biomarker composition according to any item of the first aspect of the invention.

In an embodiment of the invention, wherein the reagents comprise substances used in mass spectrometry for detection of the biomarker composition according to any item of the

first aspect of the invention.

In an embodiment of the invention, the samples for detection of the biomarkers are selected from blood, plasma and serum.

The third aspect of the invention relates to a kit, which comprises the biomarker composition according to any item of the first aspect of the invention and/or the reagent composition according to any item of the second aspect of the invention.

In an embodiment of the invention, wherein the kit further comprises the training dataset of the level of each of the biomarkers in the biomarker composition according to any item of the first aspect of the invention for CHD patients and healthy controls (for example as shown by Table 7).

The fourth aspect of the invention relates to a use of the biomarker composition according to any item of the first aspect of the invention and/or the reagent composition according to any item of the second aspect of the invention in the preparation of a kit, wherein the kit is used for evaluation of the risk of CHD in a subject, or for use in diagnosis of CHD in a subject.

In an embodiment of the invention, wherein the evaluation or diagnosis comprises the following steps: 1) determining the level of each of the biomarkers of the biomarker composition according to any item of the first aspect of the invention in a sample from the subject; 2) comparing the level of step 1) with a reference dataset or a reference value (for example a reference value of healthy controls); preferably, the reference dataset comprises the level of the biomarker of the biomarker composition according to any item of the first aspect of the invention in a sample from CHD patients and healthy controls.

In an embodiment of the invention, wherein the sample is selected from blood, plasma and serum.

In an embodiment of the invention, wherein the comparing the level of step 1) with a reference dataset further comprises the step of executing a multivariate statistical model to output the probability of illness; preferably, the multivariate statistical model is random forest model.

In an embodiment of the invention, wherein the subject is determined as being at risk of

CHD or having CHD if the probability of illness $\geqslant 0.5$.

In an embodiment of the invention, wherein, when compared with a reference value, the decrease of one or more Biomarkers selected from the group consisting of Biomarker 1-16, preferably, from the group consisting of Biomarker 1-7, indicates that the subject is in the risk of CHD or has CHD.

In an embodiment of the invention, wherein, when compared with a reference value, the increase of Biomarker 17 and/or Biomarker 18 indicates that the subject is in the risk of CHD or has CHD.

In an embodiment of the invention, wherein the step of determining the level of each of the biomarkers are carried out by mass spectrometry, preferably, mass spectrometry in conjunction with chromatography, such as gas chromatography mass spectrometry (GC-MS) or liquid chromatography mass spectrometry (LC-MS).

In an embodiment of the invention, wherein the method further comprises the step of processing the sample before step 1).

In an embodiment of the invention, wherein the kit further comprises the training dataset of the levels of the biomarker composition according to any item of the first aspect of the invention for CHD patients and/or healthy controls (for example as shown by Table 7).

The fifth aspect of the invention relates to a method for evaluation of the risk of CHD in a subject or for diagnosis of CHD in a subject, wherein the method comprises the following steps:1)determining the levels of the biomarkers of the biomarker composition according to any item of the first aspect of the invention in a sample from the subject; 2) comparing the level of step 1) with a reference dataset or a reference value (for example a reference value of healthy controls); preferably, the reference dataset comprises the level of the biomarker of the biomarker composition according to any item of the first aspect of the invention in a sample from CHD patients and healthy controls.

In an embodiment of the invention, wherein the sample is selected from blood, plasma and serum.

In an embodiment of the invention, wherein the comparing the level of step 1) with a reference dataset further comprises the step of executing a multivariate statistical model to

output the probability of illness; preferably, the multivariate statistical model is random forest model.

In an embodiment of the invention, wherein the subject is determined as being at risk of CHD or having CHD if the probability of illness $\geqslant 0.5$.

In an embodiment of the invention, wherein, when compared with a reference value, the decrease of one or more Biomarkers selected from the group consisting of Biomarker 1-16, preferably, from the group consisting of Biomarker 1-7, indicates that the subject is in the risk of CHD or has CHD.

In an embodiment of the invention, wherein, when compared with a reference value, the decrease of Biomarker 17 and/or Biomarker 18 indicates that the compound is a candidate compound for treatment of CHD in a subject or the treatment of CHD in the subject is effective.

In an embodiment of the invention, wherein the step of determining the level of each of the biomarkers are carried out by mass spectrometry, preferably, mass spectrometry in conjunction with chromatography, such as gas chromatography mass spectrometry (GC-MS) or liquid chromatography mass spectrometry (LC-MS).

In an embodiment of the invention, wherein the method further comprises the step of processing the samples before step 1).

In an embodiment of the invention, wherein the method further includes setting up the training dataset of the levels of the biomarker composition according to any item of the first aspect of the invention for CHD patients and/or healthy controls (for example as shown by Table 7).

The present invention further relates to the biomarker composition according to any item of the first aspect of the invention, for use in a method of evaluation of the risk of CHD in a subject or for diagnosis of CHD in a subject.

In an embodiment of the invention, wherein the method of evaluation or diagnosis comprises the following steps: 1) determining the levels of the biomarkers of the biomarker composition according to any item of the first aspect of the invention in a sample from the subject; 2) comparing the level of step 1) with a reference dataset or a reference value (for

example a reference value of healthy controls); preferably, the reference dataset comprises the level of the biomarker of the biomarker composition according to any item of the first aspect of the invention in a sample from CHD patients and healthy controls.

In an embodiment of the invention, wherein the sample is selected from blood, plasma and serum.

In an embodiment of the invention, wherein the comparing the level of step 1) with a reference dataset further comprises the step of executing a multivariate statistical model to output the probability of illness; preferably, the multivariate statistical model is random forest model.

In an embodiment of the invention, wherein the subject is determined as being at risk of CHD or having CHD if the probability of illness $\geqslant 0.5$.

In an embodiment of the invention, wherein, when compared with a reference value, the decrease of one or more Biomarkers selected from the group consisting of Biomarker 1-16, preferably, from the group consisting of Biomarker 1-7, indicates that the subject is in the risk of CHD or has CHD.

In an embodiment of the invention, wherein, when compared with a reference value, the increase of Biomarker 17 and/or Biomarker 18 indicates that the subject is in the risk of CHD or has CHD.

In an embodiment of the invention, wherein the step of determining the level of each of the biomarkers are carried out by mass spectrometry, preferably, mass spectrometry in conjunction with chromatography, such as gas chromatography mass spectrometry (GC-MS) or liquid chromatography mass spectrometry (LC-MS).

In an embodiment of the invention, wherein, the method further comprises the step of processing the samples before step 1).

The present invention further relates to a use of the biomarker composition according to any item of the first aspect of the invention and/or the reagent composition according to any item of the second aspect of the invention in the preparation of a kit, wherein the kit is used for screening candidate compounds for treatment of CHD in a subject, or for evaluating the effect of the treatment of CHD in a subject.

In an embodiment of the invention, wherein the screening or evaluation comprises the following steps: 1)determining the level of each of the biomarkers of the biomarker composition according to any item of the second aspect of the invention in a sample from the subject after administering the candidate compounds or the treatment to the subject; 2) comparing the level of step 1) with the level of the above mentioned biomarker before administering the candidate compounds or the treatment to the subject.

In an embodiment of the invention, wherein the sample is selected from blood, plasma and serum.

In an embodiment of the invention, wherein the increase of one or more Biomarkers selected from the group consisting of Biomarker 1-16, preferably, from the group consisting of Biomarker 1-7, indicates that the compound is a candidate compound for treatment of CHD in a subject or the treatment of CHD in the subject is effective.

In an embodiment of the invention, wherein the decrease of Biomarker 17 and/or Biomarker 18 indicates that the compound is a candidate compound for treatment of CHD in a subject or the treatment of CHD in the subject is effective.

In an embodiment of the invention, wherein the step of determining the level of each of the biomarkers are carried out by mass spectrometry, preferably, mass spectrometry in conjunction with chromatography, such as gas chromatography mass spectrometry (GC-MS) or liquid chromatography mass spectrometry (LC-MS).

In an embodiment of the invention, wherein the method further comprises the step of processing the samples before step 1).

The present invention further relates to a method for screening candidate compounds for treatment of CHD in a subject or for evaluation of the effect of the treatment of CHD in a subject, wherein the method comprises the following steps:1) determining the levels of the biomarkers of the biomarker composition according to any item of the first aspect of the invention in a sample from the subject after administering the candidate compounds or the treatment to the subject; 2) comparing the level of step 1) with the level of the above mentioned biomarker before administering the candidate compounds or the treatment to the subject.

In an embodiment of the invention, wherein the sample is selected from blood, plasma

and serum.

In an embodiment of the invention, wherein the increase of one or more Biomarkers selected from the group consisting of Biomarker 1-16, preferably, from the group consisting of Biomarker 1-7, indicates that the compound is a candidate compound for treatment of CHD in a subject or the treatment of CHD in the subject is effective.

In an embodiment of the invention, wherein the decrease of Biomarker 17 and/or Biomarker 18 indicates that the compound is a candidate compound for treatment of CHD in a subject or the treatment of CHD in the subject is effective.

In an embodiment of the invention, wherein the step of determining the level of each of the biomarkers are carried out by mass spectrometry, preferably, mass spectrometry in conjunction with chromatography, such as gas chromatography mass spectrometry (GC-MS) or liquid chromatography mass spectrometry (LC-MS).

In an embodiment of the invention, wherein the method further includes setting up the training dataset of the levels of the biomarker composition according to any item of the first aspect of the invention for CHD patients and/or healthy controls (for example as shown by Table 7).

In an embodiment of the invention, wherein the method further comprises the step of processing the samples before step 1).

The present invention further relates to the biomarker composition according to any item of the first aspect of the invention, for use in a method of screening candidate compounds for treatment of CHD in a subject or for evaluation of the effect of the treatment of CHD in a subject.

In an embodiment of the invention, wherein the method of screening and evaluation comprises the following steps: 1) determining the levels of the biomarkers of the biomarker composition according to any item of the first aspect of the invention in a sample from the subject after administering the candidate compounds or the treatment to the subject; 2) comparing the level of step 1) with the level of the above mentioned biomarker before administering the candidate compounds or the treatment to the subject.

In an embodiment of the invention, wherein the sample is selected from blood, plasma

and serum.

In an embodiment of the invention, wherein the increase of one or more Biomarkers selected from the group consisting of Biomarker 1-16, preferably, from the group consisting of Biomarker 1-7, indicates that the compound is a candidate compound for treatment of CHD in a subject or the treatment of CHD in the subject is effective.

In an embodiment of the invention, wherein the decrease of Biomarker 17 and/or Biomarker 18 indicates that the compound is a candidate compound for treatment of CHD in a subject or the treatment of CHD in the subject is effective.

In an embodiment of the invention, wherein the step of determining the level of each of the biomarkers are carried out by mass spectrometry, preferably, mass spectrometry in conjunction with chromatography, such as gas chromatography mass spectrometry (GC-MS) or liquid chromatography mass spectrometry (LC-MS).

In an embodiment of the invention, wherein the method further comprises the step of processing the samples before step 1).

The present invention further relates to A method for setting up a mass spectrometry model for evaluation of the risk of CHD in a subject or for diagnosis of CHD in a subject, which comprises the step of identifying the differentially expressed substance in a blood sample between CHD patients and healthy controls, wherein the differentially expressed substance comprises one or more selected from the group consisting of Biomarker 1-18, preferably, from the group consisting of of Biomarker 1-7.

Terms used herein have meanings as commonly understood by a person of ordinary skill in the fields to which the present invention is relevant. However, in order to better understand the invention, the definitions and explanations of the relevant terms are provided as follows.

According to the invention, the term “coronary heart disease (CHD)”, also known as “coronary artery disease (CAD)”, is a group of diseases that includes: stable angina, unstable angina, myocardial infarction, and sudden coronary death, etc. In the present invention, patients with CHD are diagnosed by coronary angiography techniques.

According to the invention, mass spectrometry (MS) can be classified into ion traps,

quadrupole, orbitrap and time-of-flight mass spectrometry, and the deviations are 0.2amu, 0.4amu, 3ppm, 5ppm, respectively. In the present invention, the MS data is acquired with orbitrap mass spectrometry.

According to the invention, the levels of the biomarkers are indicated by peak intensity in MS.

According to the invention, mass-to-charge (m/z) and retention time (RT) have the same meaning known by prior art. In the present invention, the unit of m/z is amu, which refers to atomic mass unit, also named as Dalton. Dalton is the standard unit that is used for indicating mass on an atomic or molecular scale (atomic mass). One unified atomic mass unit is equivalent to 1 g/mol. It is defined as one twelfth of the mass of an unbound neutral atom of carbon-12 in its nuclear and electronic ground state.

The person skilled in the art knows that the values of m/z and retention time will vary within a certain range when different detection methods or LC-MS instruments are used. For example, m/z can vary in the range of ± 3.00, or ± 2.00, or ± 1.00, and retention time can vary in the range of ± 60s, or ± 45s, or ± 30s, or ± 15s.In an embodiment of the invention, a reference dataset refers to a training dataset.

According to the invention, training datasets and validation datasets have the same meaning known by prior art. In an embodiment of the invention, the training datasets refer to a collection of data of the levels of biomarkers in biological samples for CHD patients and healthy controls. In an embodiment of the invention, the validation datasets refer to a collection of data used to test the performance of the training datasets. In an embodiment of the invention, the levels of biomarkers can be indicated as absolute values or relative values according to the method of determination. For example, when mass spectrometry is used to determine the levels of the biomarkers, the intensity of a peak can represent the levels of the biomarkers, which is a relative value.

In an embodiment of the invention, a reference value refers to a reference value of healthy controls or a normal value.The person skilled in the art knows that range of normal value (absolute value) of each biomarker in sample can be obtained with test and calculation method well known in the art when sample size is large emough. Thus, when other methods except mass spectrometry (for example antibody or ELISA) are applied to detect the levels of

biomarkers, the absolute values of the levels of biomarkers in samples can be directly compared with normal values, in order to evaluate the risk of CHD and diagnosis or early diagnosis of CHD, optionally , statistical methods can be included.

According to the invention, the term "a biomarker", also named "a biological marker", refers to a measurable indicator of a biological state or condition of a subject. Such biomarkers can be any substance in the subject, for example, nucleic acid marker (e.g. DNA), protein marker, cytokine marker, chemokine marker, carbohydrate marker, antigen marker, antibody marker, species marker (species/genus marker) and functions marker (KO/OG marker) and the like, provided that they are in relation with a particular biological state or condition (such as, a disease) of the subject. Biomarkers are often measured and evaluated to examine normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention, and are useful in many scientific fields.

According to the invention, the term "a biomarker set" refers to a set of biomarkers (that is, a combination of two or more biomarkers).

According to the invention, the term "a subject" refers to an animal, particularly a mammal, such as a primate, preferably human.

According to the invention, the term "plasma" refers to the fluid component of the whole blood. Depending on the separation method used, plasma may be completely free of cellular components, or may contain various amounts of platelets and/or small amounts of other cellular components.

According to the invention, the low-molecular-weight metabolites (< 1500Da) in the urine samples are extracted for HPLC-MS experiments. The extraction method is well known in the art. In an embodiment of the invention, organic solvent (for example methanol) is used to precipitate protein, the obtained mixture is centrifuged, and then supernatant is transferred for metabolic profiling by HPLC-MS.

In an embodiment of the invention, Shimadzu Prominence HPLC system (Shimadzu) was coupled to a LTQ Orbitrap Velos instrument (Thermo Fisher Scientific, MA, USA) set at 30000 resolution to acquire HPLC-MS data; an Agilent ZORBAX ODS C18 column (150 mm×2.1 mm, 3.5 μm, Agilent, USA) was used. Sample analysis was performed in positive ion modes with a spray voltage of 4.5 kV and capillary temperature of 350°C; the mass

scanning range was 50−1500 m/z; the flow rates of nitrogen sheath gas and nitrogen auxiliary gas were set to 30 L/min and 10 L/min, respectively; the HPLC−MS system was run in binary gradient mode; solvent A was 0.1% (v/v) formic acid/water, and solvent B was 0.1% (v/v) formic acid/methanol; the gradient was as follows: 5% B at 0 min, 5% B at 5 min, 100% B at 8 min, 100% B at 9 min, 5% B at 18 min, and 5% B at 20 min; the flow rate was 0.2 mL/min. To ensure system equilibrium, the pooled “quality control ”(QC) sample was used to monitor the system stability.

Terms such as “a”, “an” and “the” are not intended to refer to only a singular entity, but include the general class of which a specific example may be used for illustration.

It would be appreciated by those skilled in the art that the terminology herein is provided for better understanding of the present invention, but is not intended to delimit the invention, except as outlined in the claims.

The present invention is based on the following findings by the inventors:

Non-invasive and highly accurate approaches to diagnose and predict CHD are urgently needed. To explore potential characteristic metabolites signatures associated with CHD, non-targeted metabolomics technique is performed to analyze plasma, urine samples and metagenomics technology is applied to further validate the metabolites with potential origin from the fecal metagenomics data of CHD patients and healthy subjects. The workflow is shown in Fig. 1. Statistical and bioinformatics methods are used to identify significantly different mass-to-charge (m/z) that can discriminate CHD cases from healthy controls. Hierarchical cluster analysis (HCA) is performed to identify m/z clusters contributing to phenotype separation and spearman correlation analysis is applied to identify potential biomarkers’ correlations related to abnormal functions. The identified significantly changed metabolites are validated using purchased standards. Several significantly different metabolites are correlated with intestine flora on ECs, KOs and species levels. This study proves the power of potential noninvasive biomarkers discovery in biofluids of patients and the integrated analysis of metabolomics and metagenomics would pave the way to reveal the interactions between host and gut microbiobes.

In this invention, plasma and urine samples from CHD patients and control healthy people

were analyzed using untargeted metabolomics technique. In plasma, 18 significantly changed metabolites (13 LPCs, 2 glycerophosphocholines, L-Arginine, GlcNAc-6-P and paraxanthine) were identified as potential biomarkers. In urine, 4 significantly changed metabolites (GlcNAc-6-P, mannitol, creatine, phytosphigosine) were identified as potential biomarkers in CHD patients. To access the clinical relevance of these potential biomarker, the diagnostic capability of these 22 metabolites was evaluated by ROC. GlcNAc-6-P appears the most discriminative biomarker which shows relatively good diagnostic ability with FN of 0.153 and FP of 0.208. Correlation analysis between potential biomarkers and biochemical clinical data suggest plasma LPCs are significantly positive correlated with cholesterol (CHOL), high-density lipoprotein (HDLC), and total protein (TP), while GlcNAc-6-P and L-arginine exhibit negative correlation with CHOL, HDLC, and TP. This suggests the metabolites may potentially influence the normal metabolic pathways in our body. Among these 59 CHD patients, 32 patients had undergone Percutaneous Coronary Intervention (PCI) before but no difference had been observed between these 32 postoperative patients group and those 27 patients group with no surgery (as shown in PCA score plots in Fig. 2), suggesting the PCI did not influence the whole metabolic pattern in patients.

As estimated, over 30% of metabolites in human body originate from intestinal microbes and may contribute to host diseases. In this study, metabolomics and metagenomics techniques were integrated and evidence that microbial species and their associated metabolites were involved in CHD diseases were uncovered for the first time. Firstly, mannitol was identified as a potential urine biomarker in CHD patients. The fact that no related homo sapiens enzymes are found in the fructose and mannose metabolism so far indicates mannitol should belong to microbial metabolites family. Mannitol was previously reported to be produced by lactic acid bacteria and pseudomonas putida. In current study, spearman correlation analysis of, KOs, species and mannitol indicates that three gut flora species, *Clostridium sp. HGF2*, *Streptococcus sp. M334*, and *Streptococcus sp. M143*, play important roles in metabolism of mannitol. This was further validated by mannose-specific IIB component of PTS system (EC:2.7.1.69) which was found to be the common enzyme in all three CHD enriched gut microbiota species. Secondly, GlcNAc-6-P, an endogenous and microbial metabolites, was identified in both plasma and urine samples of CHD patients.

GlcNAc-6-P participates in sugar metabolism with dual functions in regulating host cardiovascular activity. Previous literature shows that *Escherichia coli* could metabolize anhydro-N-acetylmuramic acid obtained either from the environment or its own cell wall to N-acetylglucosamine-phosphate (Uehara, T., Suefuji, K., Jaeger, T., Mayer, C. &Park, J.T. Mur Q etherase is required by Escherichia coli in order to metabolize anhydro-N-acetylmuramic acid obtained either from the environment or from its own cell wall. J. Bacteriol. 188, 1660-1662 (2006), incorporated herein by reference), which would then be converted to GlcNAc-6-P in amino sugar and nucleotide sugar metabolism. In our study, the metabolism of GlcNAc-6-P was found to be significantly correlated with *Clostridium sp. HGF2* by NagA (EC:3.5.1.25) and N-acylglucosamine-6-phosphate 2-epimerase (EC:5.1.3.9). GlcNAc-6-P can be converted into glucosamine-6-P by NagA (EC:3.5.1.25) enzyme which plays a central role in microbial cell wall synthesis and glycolysis.

The discovery of these two microbial metabolites (Mannitol and GlcNAc-6-P) and their correlated microbiota in CHD patients has two important implications. First, it confirmed that microbial metabolites can be used as potential biomarkers for CHD diagnosis along with other traditional metabolites. For instance, GlcNAc-6-P in urine exhibited relatively strong CHD diagnostic ability with AUC of 0.88 and showed FN of 0.153 and FP of 0.208 in the ROC analysis of validation dataset. Second, microbial metabolites reflect the abnormalities of the host intestine microbiota, so new strategy for CHD treatments can be developed by adjusting patients gut intestine ecosystem. In the future, microbial species and their associated metabolites could be used as new indexes and targets for diagnosis and treatment of CHD.

In summary, non-targeted metabolomics technology and metagenomics technology were integrated for CHD study in this study. This combinational work not only lead to important biological candidate biomarkers discovery in CHD, but also bridge the gap between our human systems and intestine microbiota and provide novel insights into the microbial species related to CHD.

The present invention is further exemplified in the following non-limiting Examples. Unless otherwise stated, parts and percentages are by weight and degrees are Celsius. The agents as used were all commercially available. As apparent to one of ordinary skill in the art,

these Examples, while indicating preferred embodiments of the invention, are given by way of illustration only.

BRIEF DISCRIPTION OF DRAWINGS

These and other aspects and advantages of the present disclosure will become apparent and more readily appreciated from the following descriptions taken in conjunction with the drawings, in which:

Fig. 1 | The workflow in our study. Firstly, potential biomarkers discovery in plasma and urine had been performed. Secondly, pathway analysis and association analysis of potential biomarkers and gut flora had been applied. Lastly, potential biomarkers associated gut flora species had been discovered.

Fig. 2 | PCA Scores plot of 32 postoperative patients and 27 no operative patients with 43 healthy controls. (**a**) The plasma samples of postoperative CHD patients and no operative CHD patients gathered together showing no significant difference among them. (**b**) The scattered distribution of urine samples in postoperative CHD patients and no operative CHD patients also suggested no significant difference among them.

Fig. 3 | Potential biomarkers discovery in plasma metabolomics. Cloud plot of plasma metabolites profiles. Upper and under circles indicated metabolites with increased (*fold change* > 1.2, 196 metabolites) and decreased intensity (*fold change* < 0.8, 319 metabolites) in CHD patients' plasma samples compared with healthy controls. The darkness of color is correlated with adjusted *p.value* (named as *q.value*)*: c*olor from white to dark black indicated smaller adjusted *p.value*. The area of circle is correlated with magnitude of intensity change: In the upper part, the bigger the circle was, the more enriched metabolites were in CHD patients' plasma samples compared with healthy controls' plasma samples. While in the under part, the bigger the circle was, the more enriched metabolites were in healthy controls'.

Fig. 4 | Plasma metabolomics statistical analysis. In Veen-plot, Volcano-plot and S-plot were integrated and there were 202 overlapped m/z that were significantly different between CHD patients' plasma samples and healthy controls' plasma samples.

Fig. 5 | Boxplot of 15 choline metabolites.

Fig. 6 | Cloud plot of urine metabolites profiles.

Fig. 7 | Urine metabolomics statistical analysis. In Veen-plot, the overlapped 391 m/z were significantly different metabolites between CHD urine samples and control urine samples by the integration of Volcano-plot and S-plot analysis.

Fig. 8 | Veen diagram of all significant differential metabolites in plasma and urine.

Fig. 9 | Receiver operating characteristic (ROC) analysis of potential biomarkers. ROC analysis and boxplots of 7 identified plasma potential biomarkers (a-g) and 2 identified urine potential biomarkers (h-i) with good diagnostic capability among 176 additional plasma samples (98 controls vs 78 CHD patients) and 395 additional urine samples (173 controls vs 222 CHD patients) respectively.

Fig.10 | ROC analysis of 7 identified plasma potential biomarkers with good diagnostic capability among plasma validation datasets (78 CHD patients plasma samples VS 98 healthy controls plasma samples).

Fig.11 | ROC analysis of 2 identified urine potential biomarkers with good diagnostic capability among urine validation datasets (222 CHD patients urine samples VS 173 healthy controls urine samples).

Fig 12 | MS/MS spectrums of LysoPC(0:0/18:0), LysoPC(18:3(6Z,9Z,12Z)), LysoPC(16:1(9Z)), LysoPC(15:0), LysoPC(P-16:0) , and LysoPC(14:0).

EXAMPLES

Example 1. Identifying and validating biomarkers for evaluating risk of CHD related diseases

1.1 Materials. Formic acid and methanol (HPLC grade) was purchased from Fisher Scientific Corporation (Loughborough, UK). Water used in the experiments was obtained from a Milli-Q Ultra-pure water system (Millipore, Billerica, MA). An Agilent ZORBAX ODS C-18 column (150 mm×2.1 mm, 3.5 μm, Agilent, USA) was used for all analysis.

1.2 Clinical samples. All patients with CHD diagnosed by coronary angiography techniques were recruited from the Guangdong General Hospital. All control people enrolled in our study were free of clinically evident CHD at medical examination during the same period.

Paired plasma, urine and fecal samples of CHD patients (n=59) and healthy controls

(n=43) were obtained from the Guangdong General Hospital. Coronary angiography techniques were performed to diagnose CHD patients recruited in this study. The healthy control had underwent physical examination in the same hospital. Patients and controls did not receive probiotics or antibiotics within one month before sample collection. Among these 59 CHD patients, 32 patients had undergone Percutaneous Coronary Intervention (PCI) before. The participants' clinical information was provided in Supplementary Table 1. Besides, 176 additional plasma samples (98 controls vs 78 CHD patients) and 395 additional urine samples (173 controls vs 222 CHD patients) were included for potential biomarkers diagnostic capability analysis. Venous blood and midstream urine were collected in the morning before breakfast from all participants. Venous blood collected by Vacuette EDTA blood collection tubes were centrifuged at 2200x g for 5 min at 4 °C to obtain plasma samples. The plasma and urine samples were stored at −80 °C until use.

The corresponding fresh stool samples were collected from CHD patients (n=59) and healthy controls (n=43) on the same day at hospital. Samples were mechanically homogenized with a sterile spatula, and then aliquots containing 1g of stool in a 12ml sterile cryovial were made using the Sarstedt stool sampling system (Sarstedt, Nümbrecht, Germany). The aliquots were then stored in freezers at -20 °C and transported to the laboratory with ice pack within 48 h after collection. After that, fecal samples were stored at −80 °C until use. These protocols were reviewed by the Institutional Review Board of BGI-Shenzhen. Before collecting samples, patients were informed and written consent were obtained from them.

Samples preparations for HPLC-MS experiments. To extract the low-molecular-weight metabolites (<1500Da) in the plasma and urine samples, some modifications were made to the procedures reported previously (Luan, H. et al. Serum metabolomics reveals lipid metabolism variation between coronary artery disease and congestive heart failure: a pilot study. Biomarkers.18, 314−321 (2013), incorporated herein by reference). Before experiment, all plasma and urine samples were thawed on ice and a "quality control" (QC) sample was made by mixing and blending equal volumes (10μL) from each plasma or urine samples which was applied to estimate a "mean" profile representing all the analytes encountered during analysis. Then low molecular weight metabolites (<1500Da) were extracted from the plasma and urine samples using the following procedure: Plasma

samples were mixed with methanol (1:2 v/v) while urine samples were mixed with methanol (1:1 v/v) to precipitate protein. The plasma and urine sample mixture were centrifuged at 14000x g for 10 min at 4°C and then supernatant was transferred into a 1.5 mL polypropylene tube for metabolic profiling by HPLC−MS.

HPLC-MS experiments. Shimadzu Prominence HPLC system (Shimadzu) was coupled to a LTQ Orbitrap Velos instrument (Thermo Fisher Scientific, MA, USA) set at 30000 resolution to acquire HPLC-MS data. An Agilent ZORBAX ODS C18 column (150 mm×2.1 mm, 3.5 μm, Agilent, USA) was used. Sample analysis was performed in positive ion modes with a spray voltage of 4.5 kV and capillary temperature of 350°C. The mass scanning range was 50−1500 m/z. The flow rates of nitrogen sheath gas and nitrogen auxiliary gas were set to 30 L/min and 10 L/min, respectively. The HPLC−MS system was run in binary gradient mode. Solvent A was 0.1% (v/v) formic acid/water, and solvent B was 0.1% (v/v) formic acid/methanol. The gradient was as follows: 5% B at 0 min, 5% B at 5 min, 100% B at 8 min, 100% B at 9 min, 5% B at 18 min, and 5% B at 20 min. The flow rate was 0.2 mL/min. To ensure system equilibrium, the pooled QC sample was injected five times at the beginning. QC sample was injected every five samples during samples detection to further monitor the system stability.

HPLC-MS data analysis. The acquired MS data pretreatments including peak picking, peak grouping, retention time correction, second peak grouping, and annotation of isotopes and adducts was performed using the same method as our previously published work (Luan, H. et al. Serum metabolomics reveals lipid metabolism variation between coronary artery disease and congestive heart failure: a pilot study. Biomarkers.18, 314−321 (2013), incorporated herein by reference). LC−MS raw data files were converted into *mzXML* format and then processed by the XCMS and CAMERA toolbox implemented with the R software(*v3.1.1*). Each ion was identified by combining retention time (RT) and m/z data. Intensities of each peaks were recorded and a three dimensional matrix containing arbitrarily assigned peak indices (retention time-m/z pairs), sample names (*observations*) and ion intensity information (*variables*) was generated.

The obtained matrix was further reduced by removing peaks with more than 80% missing values (*ion intensity = 0*) and those with isotope ions from each groups in order to

obtain consistent results. As a quality assurance strategy in metabolic profiling, all retained peaks were normalized to the QC sample using Robust Loess Signal Correction (R-LSC) based on the periodic analysis of a standard biological quality control sample (QC sample) together with the real plasma and urine samples to ensure that the data are of high quality within an analytical run (Dunn, W. B. et al. Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. Nat.Protoc.6, 1060−1083 (2011), incorporated herein by reference). The relative standard deviation (*RSD*) values of metabolites in the QC samples was set at a threshold of 30% which was accepted as a standard in the assessment of repeatability in metabolomics data sets.

The nonparametric univariate method (Mann−Whitney−Wilcoxon test) was performed to measure and discover the significantly changed metabolites among the CHD patients and control subjects and then corrected by false discovery rate (FDR) to ensure that metabolite peaks were reproducibly detected. And multivariate statistical analysis (PCA, PLS-DA) were performed to discriminate CHD samples from control subjects. A number of metabolites responsible for the difference in the metabolic profile scan of CHD patients and control subjects can be obtained on the basis of variable importance in the projection (*VIP*) threshold of 1 from the 7-fold cross-validated PLS-DA model. The PLS-DA model was validated at a univariate level using FDR test from the R statistical toolbox with the critical *p.value* set to not higher than 0.05. Permutation multivariate analysis of variance (PERMANOVA), a permutation-based version of the multivariate analysis of variance, was performed in R using the “vegan” package to test the statistical significant differences between metabolic profiles and individuals’ phenotypes(Anderson, M. J. A new method for non-parametric multivariate analysis of variance. Aust. Ecol. 26, 32−46 (2001), incorporated herein by reference). Three dimensional PLS-DA analysis was also implemented to show the difference between CHD samples and control subjects. Phenotype analysis was performed to cluster those significantly distributed metabolites. Spearman correlation analysis was implemented among those significantly changed plasma metabolites, urine metabolites and clinical data of CHD patients and control subjects and correlations of metabolites was profiled with Cytoscape software 3.0.2. In addition, receiver operating characteristic (ROC) analysis was used to evaluate

diagnostic capability of identified potential biomarkers with the online tool - ROCCET (*http://www.roccet.ca*) (Xia, J.G., Broadhurst, D.I., Wilson, M. & Wishart, D.S. Translational Biomarker Discovery in Clinical Metabolomics: An Introductory Tutorial . Metabolomics 9, 280-299(2012), incorporated herein by reference).

Metabolites identification and validation. The online HMDB database (*http://www.hmdb.ca*) and KEGG database (*www.genome.jp/kegg/*) were used to identify the metabolites by matching the exact molecular mass data (m/z) of samples with those from database. If a mass difference between observed and the database value was less than 10 ppm, the metabolite would be identified and the molecular formula of metabolites would further be validated by the isotopic distribution measurements. Reference standards were purchased and used to validate and confirm those significantly changed metabolites by comparing their MS/MS spectra and retention time.

Metabolic profiles of plasma and urine samples

Untargeted metabolomics analysis was performed for the plasma and urine samples from 59 CHD patients and 43 healthy controls. The participants' clinical information was listed in Supplementary Table 1. The detailed workflow was illustrated in Fig. 1. A total of 1347 m/z (mass-to-charge ratio, 93.67%) and 2858 m/z (96.68%) were obtained in plasma and urine samples respectively. The stability and reproducibility of current data was evaluated by the QC samples measured during the whole experimental period. Principle component analysis (PCA) scores plot representation of QC samples for plasma and urine samples were obtained respectively. No drift in the metabolites profiles obtained in positive ion modes, were observed demonstrating good stability and reproducibility in our current metabolomics data set.

Results for Plasma Samples

For plasma samples, cloud plot analysis of the total 1347 m/z (Fig. 3) showed that the intensity of 196 m/z (14.55%) were increased in CHD patients' plasma samples (*fold change* > 1.2) and 23.68% of them (319 m/z) were decreased in CHD patients' (*fold change* < 0.8). Both PCA scores plot and three-dimensional partial least squares – discriminant analysis (PLS-DA) (Triba, M.N.et al. PLS/OPLS models in metabolomics: impact of permutation of dataset rows on the K-fold cross-validation quality parameters. Mol. Biosyst. 11, 13-19(2015),

incorporated herein by reference) scores plot of these plasma samples showed that there were significant differences between 59 CHD patient samples and 43 healthy control samples. CHD patients' plasma samples were apart from healthy control's samples with *PC1, PC2, PC3 as 15.32%, 10.62%, 13.73%* respectively. The permutation multivariate analysis of variance (PERMANOVA) was implemented to test the relation of individual's phenotypes with their metabolite characteristics, and we found CHD status had significant impacts on the metabolic profiling ($P < 0.001$, 1000 permutations) in positive ion mode.

S-plot analysis was used for selection of potentially interesting metabolites biomarkers (Miao, H. et al. Urinary Metabolomics on the biochemical profiles in diet-induced hyperlipidemia rat using ultra-performance liquid chromatography coupled with quadrupole time-of-flight SYNAPT high-definition mass spectrometry. J. Anal. Methods.Chem.2014, ID184162 (2014), incorporated herein by reference). Using the criteria that variable importance in the projection (*VIP*) was larger than 1, 230 variables were selected in S-plot. On the condition that *adjusted p.value* < 0.05, *fold change* > 1.2 or < 0.8, 414 variables were retained in Volcano-plot. Combing these two results, 202 shared m/z were obtained (Fig.4). And a total of 109 potential metabolites (20 increased and 89 decreased in CHD patients) were identified by aligning the exactly significant peaks' molecular mass data (m/z) with online database: HMDB and KEGG.

To further identify potential metabolites from 109 m/z, both HMDB and HMDB SERUM databases were searched using accurate mass and mass spectrometric fragmentation patterns (Chen, J. et al. Practical approach for the identification and isomer elucidation of biomarkers detected in a metabonomic study for the discovery of individuals at risk for diabetes by integrating the chromatographic and mass spectrometric information. Anal.Chem. 80, 1280−1289 (2008), incorporated herein by reference). We found 18 matched metabolites from the above database, including 13 Lysophosphatidylcholine (LPCs), 2 glycerophosphocholines, L-Arginine, N-Acetyl-D-glucosamine 6-phosphate (GlcNAc-6-P) and paraxanthine (as listed in Table 1). The MS/MS spectrums of LysoPC(0:0/18:0), LysoPC(18:3(6Z,9Z,12Z)), LysoPC(16:1(9Z)), LysoPC(15:0), LysoPC(P-16:0) , and LysoPC(14:0) are shown in Fig.14.

The intensity of 13 Lysophosphatidylcholine (LPCs) , 2 glycerophosphocholines and

paraxanthine were lower in CHD patients (as shown in Fig.5).

To investigate latent relationships of those 109 significantly changed metabolites, spearman correlation analysis was also performed. Significantly changed plasma metabolites with smaller *adjusted p.value* either in CHD enriched metabolites or in control enriched metabolites had a relatively stronger correlation. Analysis among those 18 identified metabolites showed that LysoPC(18:0) had strong positive correlations with the following metabolites: LysoPC(18:0) vs LysoPC(P-16:0) (*rho = 0.861, q.value = 0*), LysoPC(20:3(5Z,8Z,11Z)) (*rho = 0.831, q.value = 0*), LysoPC(0:0/18:0) (*rho = 0.802, q.value = 0).* LysoPC(16:1(9Z)) had strong positive correlations with LysoPC(14:0) (*rho = 0.854, q.value = 0*) and LysoPC(18:0) (*rho= 0.815, q.value = 0*). On the other hand, L-Arginine negatively correlated with 1-Palmitoylglycerophosphocholine (*rho = -0.558, q.value = 1.07E-08*).

Results for Urine Samples

In the urine cloud plot (Fig.6), there were 870 m/z (30.44%) with increased intensity in CHD patients (*fold change > 1.2*) while the level of 557 m/z (19.49%) were decreased (*fold change < 0.8*). PCA and PLS-DA models were used and the analysis results were shown in PCA scores plot and three-dimensional PLS-DA scores plot (*PC1(4.34%), PC2(8.25%), PC3(2.99%)*). This result indicated the CHD patients' urine samples were significantly different from healthy subjects'. PERMANOVA analysis demonstrated CHD had a significant impact on metabolic profile. Furthermore, S-plot analysis and Volcano-plot analysis were applied for potential biomarkers discovery. Using these criteria (*VIP* > 1, *adjusted p.value* produced by Mann−Whitney−Wilcoxon test after FDR correction < 0.05, *fold change* > 1.2 or < 0.8), 391 m/z were found to be significantly changed in CHD group by intersection of 558 m/z and 559 m/z in S-plot and Volcano-plot, respectively, as is shown in Veen plot (Fig.7).

The 391 m/z were aligned and annotated using the HMDB and KEGG database. Among the 160 annotated metabolites, the intensity of 96 metabolites were increased while that of 64 metabolites were decreased in CHD patients. These 160 metabolites were used to perform phenotype analysis for the 102 samples. The CHD patients' metabolism was obviously different from healthy controls. By comparing MS/MS spectra and retention time with

commercially available reference standards, 4 metabolites were verified and the results were listed in Table 2. The level of GlcNAc-6-P and mannitol were increased with fold change of 165.99 and 8.45 in CHD patients respectively. Meanwhile, the level of creatine and phytosphigosine were decreased with fold changes of 0.41 and 0.39 respectively.

The level of GlcNAc-6-P was found to be increased in both plasma and urine samples. GlcNAc-6-P could be produced by human body enzymes and gut bacterial enzymes, so its origin need to be further determined. Similar pattern of GlcNAc-6-P increasing in urine and blood suggested interactions of host and gut flora were important in the development of CHD. Mannitol is a polyol or sugar alcohol produced by several microorganisms. The fact that no related homo sapiens enzymes are found in the fructose and mannose metabolism so far indicates mannitol might belong to microbial metabolites family. Mannitol could be produced by many microorganisms, such as lactic acid bacteria (Carvalheiro, F., Moniz, P., Duarte, L.C., Esteves, M.P. & Gírio, F.M. Mannitol production by lactic acid bacteria grown in supplemented carob syrup. J. Ind Microbiol. Biotechnol. 38, 221-7 (2011), incorporated herein by reference) and pseudomonas putida (Kets, E.P.Galinski, E.A., de Wit, M., de Bont, J.A. &Heipieper, H.J. Mannitol, a novel bacterial compatible solute in Pseudomonas putida S12.J. Bacteriol.178, 6665-6670 (1996), incorporated herein by reference). Creatine was also found to be decreased in CHD patient's urine sample. It is a nitrogenous organic acid naturally produced by the human body from amino acids. In biosystem, creatine can elevate creatine phosphate levels and improve maintenance of ATP content during tissue oxygen depletion period, and it also has the capacity to scavenge free radicals and reduce oxidative stress. Decreased creatine in CHD patients' urine indicated that obvious energy disturbance, more free radicals and more serious oxidative stress existed in CHD patients. Reduced phytosphingosine was also found in CHD patients' urine. Phytosphingosine is a phospholipid and a major component of mammalian tissue biological membranes. The synthesis of phytosphingosine can be performed by human body and intestinal microbiota in the sphingosine metabolism. Phytosphingosine could induce caspase-independent apoptosis in human T-cell lymphoma and non-small cell lung cancer cells. The decrease in urine phytosphingosine suggested sphingolipid metabolism was abnormal in CHD patients.

To evaluate correlation among 160 annotated urine metabolites, spearman correlation

analysis was performed. The results showed that urine metabolites which were significantly changed (with smaller *adjusted p.value)* had relatively stronger correlations compared with plasma significant metabolites. In addition, among those 4 validated metabolites, mannitol showed a relatively high positive correlation with GlcNAc-6-P (*rho = 0.775, q.value =9.40E-21*) and this could be a signal for disturbed intestine microbiota.

Correlations between plasma and urine significant metabolites.

To apply metabolomics to clinical diagnosis, treatments or pathophysiology research, physiological function of metabolites and relationships between them needed to be illustrated, and correlation networks of all potential biomarkers needed to be built (Liu, P. et al. Biomarkers of primary dysmenorrhea and herbal formula intervention: an exploratory metabonomics study of blood plasma and urine. Mol. BioSyst.9, 77—87 (2013), incorporated herein by reference). For this purpose, the significantly changed plasma and urine metabolites were connected according to spearman correlation analysis profiled with Cytoscape software. These 109 annotated significantly changed plasma metabolites and 160 annotated significantly changed urine metabolites were involved in different pathways and can be divided into 8 categories: carbohydrate metabolism, lipids metabolism, amino acids metabolism, bile acids metabolism, purine/pyrimidine metabolism, vitamins metabolism, microbial related metabolism and others. Lipids metabolism showed significantly negatively correlations with microbial related metabolism while other 6 metabolism categories were in strong positive correlation with microbial related metabolism.

Seven significantly changed metabolites (Supplementary Table 2), including GlcNAc-6-P, were found both in plasma and urine on the condition that retention time error was less than 1 min and m/z error was less than 0.01 Dalton with MS/MS comparison. A Veen diagram exhibiting the common metabolites among plasma and urine significantly changed metabolites is provided in Fig.8. Two metabolites (m/z : 185.04, 202.04) were decreased in CHD patients while other five metabolites (m/z : 125.01, 309.05, 310.04, 311.05, 324.04) were increased in CHD patients.

To evaluate the correlation among 7 common metabolites, spearman correlation analysis was implemented using the criteria that the coefficient was larger than 0.90. First, correlations among plasma metabolites were obtained. m/z 311.05 showed strong correlation

with m/z 309.05 (*rho = 0.929, q.value = 9.31E-45*), m/z 310.04 (*rho = 0.911, q.value = 5.70E-40*), m/z 324.04 (*rho = 0.900, q.value = 1.53E-37*); m/z 309.05 also strongly correlated with m/z 310.04 (*rho = 0.929, q.value = 9.31E-45*). Second, correlations among urine metabolites were obtained， GlcNAc-6-P (m/z 324.04) was strongly correlated with m/z 310.04 (*rho = 0.933, q.value = 1.26E-45*), m/z 311.05 (*rho = 0.910, q.value = 1.17E-39*), m/z 125.01 (*rho = 0.903, q.value = 3.43E-38*), while m/z 125.01 also showed strong correlation with urine metabolite (m/z 310.04 (*rho = 0.918, q.value = 1.28E-41*). In addition, correlations of these metabolites in plasma and urine were also evaluated. The results showed that plasma metabolites have strong positive correlations with the same metabolites in urine (Supplementary Table3). Among them, validated GlcNAc-6-P (324.04) showed very strong positive correlation with itself (*rho = 0.747, q.value = 5.60E-19*).

Clinical relevance of plasma and urine potential metabolites

Receiver operating characteristic analysis

To evaluate the potential of the identified metabolites (18 plasma and 4 urine ones) as biomarkers, receiver operating characteristic analysis (ROC) was applied to 176 additional plasma samples (98 healthy controls vs 78 CHD patients) and 395 additional urine samples (173 healthy controls vs 222 CHD patients).

In plasma validation datasets, 6 LPCs and 1 glycerophosphocholine metabolites showed area under curve (*AUC*) larger than 0.80 and were significantly different in CHD patients (Table 3, Fig.12A-F). As shown in Fig. 9a-g, The levels of LysoPC(18:3(6Z,9Z,12Z)), LysoPC(P-16:0), LysoPC(15:0), 1-Palmitoylglycerophosphocholine, LysoPC(14:0), LysoPC(16:1(9Z)), LysoPC(0:0/18:0) were decreased in CHD patients with *fold change* at 0.26, 0.58, 0.51, 0.65, 0.49, 0.62, 0.42 respectively and *AUC* of 0.91, 0.88, 0.88, 0.88, 0.84, 0.83, 0.83 respectively. On the other hand, other 9 plasma potential biomarkers exhibited the same enrichment direction except that LysoPC(20:3(5Z,8Z,11Z)) became normal and GlcNAc-6-P even became undetected (data shown in Table 3). These results validated that LPCs could become biomarkers and targets for CHD diagnosis and therapies in the future.

In urine validation datasets, GlcNAc-6-P and mannitol exhibited *AUC* of 0.88, 0.81 and *fold change* at 36.91 and 2.62 respectively (as shown in Fig. 9h、9i and Table 4). However, creatine and phytosphingosine did not show good diagnostic ability in both training and

validation datasets. The existence of GlcNAc-6-P and mannitol in urine indicates the interaction of gut flora activity and host metabolism. These results also validated that GlcNAc-6-P and mannitol could become biomarkers and targets for CHD diagnosis and therapies in the future.

Among these 7 choline metabolites and 2 urine metabolites with *AUC* larger than 0.80, GlcNAc-6-P appeared the most discriminative biomarker which showed relatively good diagnostic ability with false negative (*FN*) of 0.051, 0.153 and false positive (*FP*) of 0.047, 0.208 in the training datasets and validation datasets respectively. LysoPC(18:3(6Z,9Z,12Z)), LysoPC(P-16:0), LysoPC(15:0), 1-Palmitoylglycerophosphocholine, LysoPC(14:0), LysoPC(16:1(9Z)), LysoPC(0:0/18:0) and mannitol exhibited diagnostic ability with *FN* of 0.271, 0.169, 0.136, 0.068, 0.119, 0.119, 0.085, 0.153 and *FP* of 0.233, 0.163, 0.256, 0.233, 0.209, 0.140, 0.279, 0.093 respectively in the training datasets, showed diagnostic ability with *FN* of 0.013, 0, 0, 0.013, 0.013, 0.013, 0, 0.135 and *FP* of 0.582, 0.755, 0.673, 0.694, 0.612, 0.684, 0.714, 0.416 respectively in the validation datasets. The *FN* and *FP* of 7 choline metabolites and 2 urine metabolites were all analysed with R using the "randomForest" and "pROC" packages based on the intensity of training datasets shown in Table 7 and Table 8 respectively. The randomForest model classification output prediction results (probability of illness; cutoff was 0.5, and if the probability of illness $\geqslant$0.5, the subject was at risk of CHD).

7 potential plasma biomarkers were combined to perform ROC analysis in plasma training datasets (59 CHD patients plasma samples VS 43 healthy control plasma samples) and plasma validation datasets (78 CHD patients plasma samples VS 98 healthy controls plasma samples, Fig.10, Table 9, Table 11) with R package – pROC. The FN and FP were (0,0) and (0,0.724) in plasma training and validation datasets respectively. 2 potential urine biomarkers were combined to perform ROC analysis in urine training datasets (59 CHD patients urine samples VS 43 healthy control urine samples) and urine validation datasets (222 CHD patients urine samples VS 173 healthy controls urine samples, Fig.11, Table 10, Table 11) with R package - pROC. The FN and FP were (0.016,0) and (0.121,0.225) in urine training and validation datasets respectively. The FN and FP of 7 choline metabolites combination and 2 urine metabolites combination were all analysed with R using the "randomForest" and "pROC" packages based on the intensity of training datasets shown in

Table 7 and Table 8 respectively. The randomForest model classification output prediction results (probability of illness; cutoff was 0.5, and if the probability of illness $\geqslant$0.5, the subject was at risk of CHD).

Table 7 Ion intensity of 7 plasma biomarkers in training datasets

102 samples(CD:CHD patients;N:healthy controls)	LysoPC(14:0)	LysoPC(P-16:0)	LysoPC(15:0)	LysoPC(16:1(9Z))	1-Palmitoylglycerophosphocholine	LysoPC(18:3(6Z,9Z,12Z))	LysoPC(0:0/18:0)
CD6582	1.042127541	1.208737598	1.059574898	1.173205968	0.889190052	0.880587135	0.725244112
CD6584	0.585457489	1.2218839	0.863245781	0.784451075	0.934003832	0.797798668	0.894604604
CD6586	0.754401051	1.173551821	0.608456925	0.793428778	0.972553931	1.005087799	0.997822897
CD6591	0.925819606	0.824273154	0.89561544	0.801804465	1.117545963	0.513786927	1.104934224
CD6594	0.7750205	0.638422814	0.997973893	0.565017812	0.715874411	0.553600588	0.338796763
CD6595	0.727783442	1.673317344	1.72226422	1.08308191	1.139624411	0.680697505	1.191273461
CD6598	0.885322709	1.785177295	1.238386104	1.062626005	1.22321139	0.849432582	1.231429677
CD6601	0.98738124	1.318177668	1.412582055	1.173876529	1.195912193	0.784581216	1.393532295
CD6604	0.384427926	0.824003959	0.527495274	0.73293692	0.711575131	0.330228649	0.419482463
CD6608	0.766958651	1.260338626	1.077527819	0.963272581	1.158262076	1.002930714	1.001835912
CD6610	1.334870141	1.055327451	0.673935468	1.40436822	1.388892022	1.755096444	0.444391047
CD6620	0.425644078	0.945142684	0.755444414	0.742756908	0.929056603	0.331255954	0.824626728
CD6636	0.734272503	0.848497181	1.189941134	0.626327482	0.866579463	0.913066847	0.723191444
CD6643	0.826018091	1.095699157	1.00389615	0.786208169	1.026640481	0.685517831	0.753589967
CD6644	0.775671551	0.903558342	0.93756113	1.124040474	1.013692341	1.336955282	0.717079016
CD6645	0.771487085	0.991562899	1.226891899	0.868823723	0.883231341	1.144283179	0.885128339
CD6647	0.885001623	0.779400038	1.025587955	0.974781456	0.872074511	0.3554807	0.848641093
CD6649	0.646395473	1.559075425	1.058749648	0.848424013	1.037164171	1.599950433	0.66383702
CD6650	1.034107367	1.109170011	0.918848746	1.02621704	0.734988131	0.793103569	0.873765794
CD6651	0.541759998	0.670514632	0.60489938	0.592296406	0.95552446	0.655151981	0.244816878

CD6654	1.591832799	1.111739861	0.999739818	1.457178204	1.050858093	1.617693429	1.268062473
CD6657	0.992142257	0.571218792	0.830926041	0.863869862	0.873212549	1.27618041	0.881058941
CD6700	0.588454892	0.600097626	0.524591242	0.618840965	0.642807719	0.584464257	0.502367343
CD6703	1.883308489	1.391762866	1.472522211	1.685274113	1.365291077	1.938890943	1.137688671
CD6716	1.420322994	1.036417711	1.221009363	1.272621726	1.174123547	0.965766482	1.108037761
CD6722	0.467208606	1.059638677	0.583586286	0.705480061	0.881165215	0.611159491	1.153215258
CD6733	1.760585708	1.09659694	1.915289336	1.026096267	1.021551321	0.886358542	0.911577403
CD8049	0.443065066	1.07316899	0.699382648	0.990393703	1.119307312	0.787325688	1.263154249
CD8051	0.501855893	0.370564652	0.306700987	0.411953013	0.482780056	0.393890074	0.967013397
CD8059	1.59063762	0.978063894	1.358492672	1.360320878	1.243191869	1.675300747	1.890643506
CD8117	0.687451375	0.550692305	0.641297334	0.917253937	0.61527712	0.775509112	0.646742606
CD8118	0.903161894	0.689325727	0.570031453	0.60345769	0.426929161	0.977577996	0.714416154
CD8119	0.132745819	0.534962851	0.271578567	0.389660553	0.487855475	0.140796057	0.104885331
CD8122	1.095615654	1.035582539	0.972209623	0.881584863	0.970524415	1.06604249	1.832521311
CD8144	0.369558278	0.654335863	0.296023025	0.62424516	0.676585	0.401916551	0.459727402
CD8145	0.610702555	0.947299645	0.543962077	0.847275928	0.819390169	0.60670942	0.715433852
CD8149	0.23248202	0.425229856	0.38298945	0.443535123	0.661453559	0.194710221	0.684065023
CD8185	0.691938104	1.10714853	0.920055549	0.787715687	0.895700203	0.881958906	0.748573865
CD8196	1.193354658	1.094940834	1.180044104	0.84963318	1.129466874	1.261832648	0.878941972
CD8197	1.215571232	1.046186889	1.168599365	0.978450042	1.239230629	1.414279939	1.391258483
CD8203	1.130785605	1.820334591	1.850076822	0.984938485	1.018106865	1.903976303	1.293821259
CD8211	1.105068999	0.844853636	0.811573305	1.078937497	1.008277324	1.690867823	1.045855929
CD8212	0.426719811	1.083216532	0.647663115	0.698535165	0.90884984	0.788159693	0.832229318
CD8234	0.539801935	1.025453938	0.765880062	0.961917247	0.992227783	0.601017421	1.027137793
CD8239	0.957753797	0.9259964	0.993557179	1.09553555	1.251578206	0.781789583	1.349955481
CD8243	0.835429899	0.769432974	0.777062489	0.723962313	0.859378653	0.604270006	1.152437684

CD8257	0.651163457	1.69782061	1.600830042	0.875788536	1.117781612	0.465805266	0.989487339
CD8264	0.696768555	0.906042498	0.520902572	0.808099195	0.670839217	0.418755525	0.511460376
CD8305	0.290258602	0.761380141	0.479892863	0.608288618	0.726659248	0.590744201	0.685215737
CD8313	1.187789962	0.719629439	1.275547047	1.486350135	1.006871736	1.710568741	1.373562043
CD8321	0.74073159	0.804077081	0.799887826	1.025370527	0.877334666	0.772551111	0.429237067
CD8330	0.667723105	1.343793683	1.209481377	1.006067183	1.15498797	0.764110102	1.82050866
CD8331	0.738537339	1.082435244	0.641465294	1.069472602	0.927371816	1.038786618	1.179091285
CD8350	0.617731094	1.22837042	0.888603376	0.842253649	1.346601146	0.930264672	1.004622288
CD8355	0.727776228	1.220668943	1.327010623	1.079970961	1.060743575	1.233941319	1.316189714
CD8367	1.42659392	1.240710024	1.260970828	1.41567668	1.347550149	2.749090084	1.664474414
CD8368	0.824021708	0.933013706	0.915189277	0.482553393	0.900277381	0.628167661	0.668553538
CD8376	1.085331682	0.747205642	1.07921017	0.952096414	0.957988036	1.390849179	1.022692458
CD8380	1.130851248	0.922055863	0.927063504	0.86566868	1.007185381	1.026500314	0.777887782
N1260	2.046320754	1.700307054	1.464240729	1.779928694	1.513044159	1.938544307	1.883127366
N1261	1.924092291	1.293484891	1.060862408	1.600014635	1.238191542	1.455747583	1.540435679
N1262	1.995964767	1.991706828	2.132467088	2.724426091	1.440382627	2.313239451	1.645754373
N1263	1.794035998	1.874215062	2.234165954	2.662380147	1.379986663	1.68371937	1.923225988
N1264	1.88882092	1.66941781	1.773773538	2.108199957	1.444691355	1.826052065	2.286571797
N1265	4.507382889	1.82142689	2.150878313	2.888518282	1.937721496	3.41779712	3.567635349
N1266	1.052027347	1.485232624	1.215138802	1.568559685	1.096725839	1.539341157	1.388851996
N1267	2.794687238	2.695569404	2.213073127	3.351189981	1.779007668	2.444158133	3.263204823
N1268	2.212700753	1.635220917	2.735288382	1.824096021	1.451193528	1.037072675	1.754497191
N1269	1.691987981	1.31207241	1.771928279	1.729717854	1.18696777	1.199303562	1.060428391
N1270	2.876453248	1.946343146	0.927718311	2.916851749	1.446978344	3.8946178	3.366974603
N1271	0.857064185	1.986401954	1.339919891	1.340313797	1.114013069	0.734173409	1.614795311
N1272	1.616774463	1.249752801	1.963735632	1.57688309	1.281657269	1.490700743	0.972921278

N1273	1.879024102	1.279908834	2.16294089	1.363843268	1.272776005	1.442085434	1.902380935
N1274	2.853474057	1.353523111	1.742905374	2.294757698	1.795092341	2.1541488	0.609632285
N1275	1.064733611	2.421759704	1.278768879	0.895738092	1.361756529	1.56806713	1.936265924
N1276	0.845398654	1.333893977	0.827840897	1.257295971	1.075905426	0.559640908	1.034940505
N1277	1.734815992	2.258098074	1.862456172	1.621100216	1.519581007	1.468551597	2.641548554
N1279	1.407328997	1.355987944	1.547188274	1.19329358	1.139223667	2.085947924	1.928764885
N1280	1.349036986	1.223411469	1.419178944	1.318711604	1.87503374	1.180289933	1.229448015
N1282	3.916886286	1.552552431	1.276833946	2.564044798	1.753284316	1.402588775	2.580263876
N1283	2.136788877	2.054017047	1.70305136	2.233168036	1.46303553	1.648455773	2.722062288
N1284	0.929019044	1.250443511	0.849181077	0.91678578	1.171963595	0.61359045	1.069918419
N1285	1.468459745	1.592522901	1.520577167	1.629024315	1.322574181	1.682972845	1.203567515
N1286	0.814143319	1.146830648	1.238560853	1.260127629	0.995352236	0.888088341	1.619585736
N1287	1.362490657	1.359433848	0.724850489	1.197097301	1.478871295	1.25390384	1.49750818
N1288	1.50448877	1.304781875	1.542497022	1.78414391	1.107031308	1.21067673	1.616090286
N1290	1.987569891	1.670441453	2.656716171	1.583175078	1.36392868	1.795122726	1.946159214
N1291	0.855578117	1.246587833	0.867953024	1.000715218	1.574841357	0.927106635	0.817086544
N1292	1.528391541	1.37811599	2.454712757	1.364816558	1.277494788	2.711350378	1.614085543
N1293	0.893275574	1.372648944	0.794361335	1.044908555	1.365133592	1.041637494	1.556359143
N1294	1.273210663	1.250050374	1.3225874	1.296704796	1.377082506	1.413047215	1.449927241
N1295	1.454492113	1.058567097	1.689091635	1.236820253	1.107396514	0.655295442	0.982385493
N1296	0.997219626	1.450176971	1.288939857	1.655156343	1.043402576	1.280591356	1.14820249
N1297	1.923165716	1.481051464	1.67562426	1.250365285	1.90347951	1.788374001	1.841531677
N1298	1.28116715	2.127293557	1.675255124	1.492634015	1.26661256	1.361678226	2.593424387
N1299	3.220748423	1.267094676	3.403846728	1.867421051	1.674035428	3.548918886	1.919565925
N1300	1.341491359	0.986538531	2.241108794	1.498997202	1.648538928	2.173814793	0.379959601
N1301	1.428761111	1.008724733	1.089398601	1.092856751	2.486750015	0.849850939	1.134545291

N1302	1.601623593	1.829949721	1.849709716	1.594649047		1.890136966	2.126923877	2.784367605
N1307	1.248639234	1.30591334	0.932272625	1.250592437		1.170090489	0.757947138	1.132525821
N1308	1.605738806	1.150441023	1.822618567	1.670746425		1.302906753	1.137880435	1.375562823
N1309	1.297769219	1.184934328	1.646806911	1.146180362		1.040626277	2.456553254	1.318963141

Table 9 Ion intensity of 7 plasma biomarkers in two samples of validation datasets

samples	label (1:CHD patients;0:healthy individuals)	LysoPC (14:0)	LysoPC (15:0)	LysoPC(16:1(9Z))	LysoPC(18:3(6Z,9Z,12Z))	1-Palmitoylglycerophosphocholine	LysoPC(P-16:0)	LysoPC(0:0/18:0)
N190E_2	0	1.789680525	1.528504271	1.564425402	0.928544121	1.332653635	1.383823769	1.201800254
CD5751_1	1	1.212552792	1.070354256	0.841407844	0.455679314	0.882024927	0.790375507	0.615492081

Table 10 Ion intensity of 2 urine biomarkers in two samples of validation datasets

samples	label (1:CHD patients;0:healthy individuals)	Mannitol	N-Acetyl-D-glucosamine-6-phosphate
N003_6_6	0	0.3165	0
ZSL001_9_8	1	0.403944	0.309672

Table 11 Probability of illness of validation datasets predicted by 7 plasma biomarkers and 2 urine biomarkers respectively

plasma		urine			
samples'name(CD:CHD patients;N:healthy individuals)	probability of illness	samples'name (ZSL:CHD patients;N:healthy individuals)	probability of illness	samples'name (ZSL:CHD patients;N:healthy individuals)	probability of illness
CD5751_1	0.984	N001_5_5	0.252	ZSL001_9_8	0.97
CD5762_1	1	N002_7_7	0.516	ZSL002_3_3	0.06
CD5767_1	1	N003_6_6	0.144	ZSL003_1_1	0.4
CD5778_1	1	N005_14_13	0.262	ZSL004_1_1	1
CD5779_1	0.866	N006_7_7	0.4	ZSL005_1_1	0.206
CD5782_1	0.978	N009_7_7	0.144	ZSL006_2_2	0.786
CD5783_1	0.994	N010_13_12	0	ZSL007_2_2	0.97
CD5788_1	1	N011_7_7	0	ZSL008_3_3	0.942
CD5789_2	0.752	N013_13_12	0.588	ZSL009_1_1	0.974
CD5793_1	1	N017_5_5	0.002	ZSL010_13_12	0.946
CD5795_1	1	N018_5_5	0.01	ZSL012_12_11	1
CD5796_1	1	N019_6_6	0.132	ZSL013_11_10	1
CD5797_1	1	N020_11_10	0.002	ZSL014_12_11	0.146
CD5802_1	1	N021_13_12	0.246	ZSL014_12_130827012100_11	0.908
CD5805_1	1	N022_13_12	0	ZSL016_12_11	1
CD5814_1	1	N023_12_11	0.004	ZSL018_12_11	0.768
CD5816_1	0.97	N024_11_10	0.13	ZSL021_14_13	1
CD5819_1	1	N025_11_10	0.012	ZSL022_11_10	1

CD5822_1	0.992	N026_13_12	0.162	ZSL023_9_8	0.338
CD5828_1	1	N027_7_7	0.024	ZSL024_9_8	0.996
CD5829_1	1	N028_10_9	0.28	ZSL025_12_11	0.966
CD5831_1	1	N029_13_12	0.278	ZSL026_9_8	0.916
CD5832_1	0.876	N031_13_12	0.128	ZSL027_11_10	0.97
CD5833_1	1	N032_11_10	0.348	ZSL028_12_11	0.832
CD5835_1	1	N033_12_11	0.002	ZSL029_10_9	0.516
CD5838_1	1	N034_12_11	0.28	ZSL030_13_12	0.974
CD5839_1	1	N035_11_10	0.152	ZSL031_12_11	1
CD5845_1	1	N036_7_7	0.002	ZSL032_12_11	1
CD5847_1	0.998	N037_6_6	0.516	ZSL033_12_11	0.984
CD5848_1	1	N038_13_12	0.262	ZSL034_14_13	0.076
CD5851_1	0.996	N039_13_12	0.002	ZSL035_9_8	0.862
CD5859_1_1	1	N040_11_10	0.086	ZSL036_9_8	1
CD5859_2_2	1	N041_13_12	0.016	ZSL037_7_7	0.26
CD5860_1	1	N042_13_12	0.4	ZSL038_14_13	1
CD5861_1	1	N043_10_9	0.016	ZSL039_12_11	0.246
CD5862_1	0.994	N044_13_12	0.4	ZSL040_11_10	0.976
CD5863_1	1	N045_13_12	0.356	ZSL041_12_11	0.028
CD5864_1	0.998	N046_13_12	0.142	ZSL042_14_13	0.846
CD5867_1	0.998	N047_5_5	0.132	ZSL043_11_10	0.972
CD5871_1	1	N049_5_5	0.16	ZSL043_9_8	0.996
CD5873_1	1	N051_10_9	0.12	ZSL044_14_13	1
CD5875_1	1	N052_12_11	0.348	ZSL045_14_13	1
CD5876_2	0.984	N054_14_13	0.01	ZSL046_10_9	1
CD5877_2	1	N055_9_8	0.028	ZSL048_7_7	1
CD5881_2	0.962	N056_9_8	0.002	ZSL049_12_11	0.926
CD5884_2	0.984	N057_11_10	0.036	ZSL050_9_8	0.508
CD5888_2	0.998	N058_13_12	0.004	ZSL052_9_8	1
CD5889_2	1	N062_13_12	0.048	ZSL053_7_7	1
CD5891_2	0.98	N063_7_7	0.002	ZSL054_10_9	0.516
CD5892_2	1	N064_5_5	0.28	ZSL056_1_1	1
CD5898_2	1	N065_10_9	0.516	ZSL057_11_10	1
CD5900_2	0.968	N066_9_8	0.262	ZSL058_6_6	0.988
CD5906_2	1	N069_11_10	0.028	ZSL059_2_2	0.746
CD5916_2	0.996	N070_9_8	0.002	ZSL060_10_9	0.992
CD5923_2	1	N071_10_9	0.516	ZSL061_12_11	1
CD5925_2	1	N072_11_10	0.332	ZSL062_7_7	0.646
CD5926_2	1	N073_11_10	0.13	ZSL063_9_8	0.044
CD5931_2	0.998	N075_6_130820152027_6	0.016	ZSL064_1_1	1

CD5934_2	0.98	N075_6_6	0.008	ZSL065_11_10	1
CD5935_2	1	N076_6_6	0.002	ZSL066_9_8	0.136
CD5936_2	0.982	N077_10_9	0.018	ZSL068_3_3	1
CD5937_2	1	N078_9_8	0.356	ZSL069_10_9	1
CD5941_2	1	N079_11_10	0.03	ZSL070_2_2	0.942
CD5979_2	1	N081_12_11	0	ZSL071_10_9	1
CD5982_2	1	N082_10_9	0.398	ZSL073_10_9	1
CD5988_2	0.964	N083_10_9	0.016	ZSL074_2_2	1
CD5990_2	1	N084_11_10	0.13	ZSL075_2_2	0.94
CD6011_2	1	N085_7_7	0.136	ZSL076_1_1	0.884
CD6017_2	1	N086_5_5	0.412	ZSL077_2_2	1
CD6024_2	1	N087_9_8	0.004	ZSL078_2_2	0.17
CD6033_2	1	N088_9_8	0.338	ZSL079_2_2	0.916
CD6039_2	1	N089_10_9	0.042	ZSL080_3_3	1
CD6043_2	1	N094_12_11	0.212	ZSL081_2_2	0.996
CD6046_2	1	N097_14_13	0.01	ZSL082_2_2	0.994
CD6050_2	1	N101_11_10	0	ZSL083_1_1	1
CD6054_2	0.994	N103_12_11	0	ZSL084_2_2	0.888
CD6059_2	1	N104_11_10	0.516	ZSL085_14_13	0.856
CD6060_2	0.994	N105_7_7	0.002	ZSL086_13_12	0.944
N162E_2	1	N106_13_12	0.004	ZSL087_1_1	1
N163E_2	0.97	N107_5_5	0	ZSL088_3_3	0.984
N164E_2	1	N108_9_8	0.206	ZSL089_1_1	0.93
N165E_2	0.978	N109_10_9	0.144	ZSL090_1_1	0.986
N166E_2	0.598	N114_13_12	0.13	ZSL091_2_2	1
N167E_2	0.994	N137_11_10	0.028	ZSL092_3_3	0.974
N168E_2	0.78	N162_11_10	1	ZSL093_3_3	1
N170E_2	0.996	N165_11_10	0.516	ZSL094_1_1	1
N171E_2	0.998	N166_11_10	0.28	ZSL095_7_7	0.172
N175_2	0.618	N167_14_13	0.996	ZSL096_3_3	1
N176_2	0.476	N168_6_6	0.246	ZSL098_3_3	1
N178_2	0.198	N170_5_5	1	ZSL099_1_1	0.96
N185E_2	0.082	N171_10_9	0.946	ZSL100_10_9	0.982
N186E_2	0.996	N172_7_7	0.898	ZSL101_7_7	0.916
N187E_2	0.3	N173_7_7	0.208	ZSL101_9_8	1
N188E_2	0.642	N174_10_9	0.146	ZSL103_10_9	0.794
N189E_2	0.526	N174_3_3	0.862	ZSL104_12_11	0.4
N190E_2	0.228	N176_7_7	0.994	ZSL105_9_8	1
N191E_2	0.632	N179_10_9	0.976	ZSL106_7_7	0.114
N192E_2	0.836	N180_11_10	0.934	ZSL107_14_13	0
N193E_2	0.204	N181_1_1	0.01	ZSL108_12_11	0.142

N194E_2	0.196	N182_2_2	0.146	ZSL109_11_10	0.516
N195E_2	0.85	N183_1_1	0.172	ZSL110_13_12	0.064
N196E_2	0.518	N184_1_1	0.122	ZSL110_5_5	0.992
N197E_2	0.934	N185_3_3	0.288	ZSL111_14_13	1
N198gan_huruilian_2	0.108	N186_2_2	0.338	ZSL113_14_13	0.96
N199gan_linrufang_2	0.34	N187_1_1	0.82	ZSL114_9_8	0.4
N200E_2	0.958	N189_1_1	0.964	ZSL115_12_11	0.996
N201gan_lvhuiX_2	0.568	N190_2_2	0.264	ZSL116_10_9	0.916
N202gan_zhangjing_1	0.292	N191_2_2	0.4	ZSL117_11_10	1
N203gan_1	0.602	N195_2_2	0.73	ZSL118_14_13	0.996
N204gan_wangmiaorong_1	0.574	N197_13_12	0.85	ZSL119_12_11	1
N205gan_liuqifang_1	0.304	N197_3_3	0.03	ZSL120_11_10	1
N206gan_liuguoying_1	0.452	N198_1_1	0.968	ZSL121_14_13	1
N207E_2	0.994	N198_130814023846_1_1	0.944	ZSL122_9_8	0.96
N208gan_zhengshuX_2	0.222	N199_13_12	0	ZSL123_14_13	1
N209gan_wangxin_2	0.606	N200_2_2	0.028	ZSL124_14_13	1
N211E_1	1	N201_2_2	0.138	ZSL125_14_13	0.898
N212E_1	0.994	N203_13_12	0	ZSL126_6_6	1
N213E_1	0.866	N204_1_1	0.4	ZSL127_12_11	0.912
N214E_1	0.986	N205_1_1	0.028	ZSL129_10_9	0.996
N215E_1	0.998	N206_2_2	0.964	ZSL130_14_13	1
N216gan_zhaominling_1	0.264	N207_2_2	0.144	ZSL131_12_11	0.852
N217E_1	0.728	N208_1_1	0.028	ZSL132_10_9	1
N218E_1	0.866	N209_2_2	0.028	ZSL133_14_13	1
N219E_2	0.72	N212_1_1	0	ZSL134_14_13	1
N220E_1	0.998	N213_1_1	0.28	ZSL135_14_13	0.96
N221E_1	0.746	N214_3_3	0.036	ZSL136_13_12	1
N222E_1	0.894	N215_2_2	0.136	ZSL137_10_9	1
N223E_1	0.856	N216_1_1	0.81	ZSL138_12_11	1
N224E_1	0.712	N217_2_2	0.28	ZSL140_13_12	0.976
N225E_1	1	N218_1_1	0.4	ZSL141_9_8	0.982
N226E_1	0.88	N220_3_3	0.986	ZSL142_14_13	0.278
N227E_1	1	N222_1_1	0.556	ZSL143_14_13	0.988
N228E_1	1	N223_1_1	0.516	ZSL144_9_8	1
N229E_1	0.984	N224_3_3	0.14	ZSL145_!2_11	1
N230E_1	0.834	N226_2_2	0.946	ZSL146_9_8	1
N231E_1	0.596	N227_2_2	0.038	ZSL147_10_9	1
N232E_1	0.258	N228_5_5	0.974	ZSL148_14_13	0.892
N233E_1	0.982	N229_6_6	0.956	ZSL149_10_9	1

N234E_1	0.98	N230_5_5	0.988	ZSL150_11_10	1
N235E_1	0.992	N231_9_8	0.824	ZSL151_7_7	1
N236E_1	0.992	N232_6_6	0.002	ZSL152_12_11	0.894
N237E_1	0.824	N233_5_5	0.988	ZSL153_9_8	1
N238E_1	0.992	N234_4_4	0.994	ZSL154_13_12	0.986
N239E_1	0.998	N235_6_6	0.16	ZSL155_13_12	0.962
N240E_1	1	N236_5_5	0.178	ZSL226_3_3	1
N241E_1	0.516	N237_4_4	0.956	ZSL227_2_2	1
N242E_1	0.998	N238_4_4	0.996	ZSL228_4_4	0.946
N243E_1	0.982	N239_4_4	0.994	ZSL229_2_2	0.956
N244E_1	0.778	N240_9_8	0	ZSL230_4_4	1
N245E_1	0.042	N241_4_4	0.158	ZSL234_1_1	0.97
N246E_1	0.08	N242_3_3	0.332	ZSL235_2_2	1
N247E_2	0.466	N243_5_5	0.998	ZSL235_3_3	0.956
N248E_2	0.334	N244_14_13	0	ZSL236_3_3	1
N249E_2	0.318	N245_5_5	0	ZSL237_4_4	1
N250E_2	0.566	N247_6_6	0	ZSL238_3_3	1
N251E_2	0.374	N248_5_5	0.002	ZSL239_6_6	1
N252E_2	0.936	N249_4_4	0.136	ZSL240_6_6	0.3
N253E_2	0.212	N250_4_4	0.102	ZSL241_5_5	1
N254E_2	0.142	N251_3_3	0.394	ZSL248_5_5	0.96
N256E_2	0.402	N253_4_4	0.4	ZSL249_6_6	0.956
N257E_2	0.282	N255_9_8	0.144	ZSL250_5_5	1
N258E_2	1	N256_3_3	0.028	ZSL252_6_6	1
N259E_2	0.242	N257_4_130818010306_4	0.332	ZSL253_14_13	1
N261E_2	0.946	N257_4_4	0.356	ZSL255_5_5	1
N262E_130320005321_2	0.788	N259_4_4	0.066	ZSL259_14_13	0.986
N262E_2_2	0.914	N267_10_9	0.204	ZSL261_6_6	1
N263E_2	1	N268_4_4	0.048	ZSL265_6_6	0.802
N264E_2	0.97	N275_13_12	0.116	ZSL266_14_13	0.548
N265E_2	0.998	N281_4_4	0.01	ZSL267_5_5	1
N266E_2	1	N284_4_4	0.228	ZSL268_5_5	1
N267E_2	0.802	N285_4_4	0.008	ZSL270_5_5	1
N268E_2	0.33	N286_4_4	0.016	ZSL271_5_5	1
N285gan_2	1	N288_4_4	0.932	ZSL272_6_6	0.976
N286E_2	0.838			ZSL275_6_6	0.938
N288E_2	0.642			ZSL277_6_6	1
N288gan_2	0.998			ZSL281_13_12	1
				ZSL282_7_7	1
				ZSL289_7_7	1

ZSL290_14_13	1
ZSL293_7_7	1
ZSL296_7_7	1
ZSL297_6_6	0.848
ZSL300_6_6	1
ZSL301_6_6	0.95
ZSL302_5_5	1
ZSL304_6_6	0.208
ZSL310_13_12	0.984
ZSL310_7_7	0.952
ZSL312_5_5	0.516
ZSL314_6_6	0.956
ZSL315_6_6	1
ZSL317_14_13	1
ZSL318_6_6	1
ZSL320_5_5	1
ZSL322_7_7	1
ZSL330_14_13	1
ZSL332_5_5	1
ZSL334_6_6	0.138
ZSL335_6_6	0.138
ZSL336_7_7	1
ZSL337_7_7	0.516
ZSL338_14_13	1
ZSL340_13_12	1
ZSL341_13_12	1
ZSL342_7_7	1
ZSL349_7_7	0.976
ZSL351_4_4	0.044
ZSL353_4_4	0.996
ZSL356_4_4	1
ZSL360_4_4	0.12
ZSL362_3_3	1
ZSL363_4_4	0.912
ZSL365_4_4	1
ZSL368_4_4	1
ZSL369_4	0.996
ZSL370_3_3	0.974
ZSL374_14_13	1
ZSL377_13_12	0.992
ZSL378_4_4	0.944

ZSL379_3_3	0.982
ZSL379_5_5	0.978
ZSL380_4_4	0.128

Example 2. Association of potential metabolic biomarkers with clinical phenotypes

To access the effects of patients' covariates (such as age and clinical biochemical factors) on metabolic profiles, PERMANOVA analysis was performed. Albumin (ALB), alanine aminotransferase (ALT), total protein (TP), low-density lipoprotein (LDLC), cholesterol (CHOL), high-density lipoprotein (HDLC), apolipoprotein b (APOB) and apolipoprotein a (APOA) were found to be significantly different in CHD patients (Supplementary Table 1).

Besides, spearman correlation analysis was performed among 18 potential plasma biomarkers and 4 potential urine biomarkers with individual phenotypes. CHOL, HDLC and TP showed significantly positive correlations with plasma LPCs (Supplementary Table 4).

LysoPC (18:0) was correlated with CHOL (*rho = 0.518, q.value = 7.89E-07*), HDLC (*rho = 0.548, q.value = 1.29E-07*) and TP (*rho = 0.573, q.value = 5.16E-08*). LysoPC(P-16:0) was positively correlated with HDLC (*rho = 0.561, q.value* = 7.39E-08). Meanwhile, the two potential urine biomarkers, GlcNAc-6-P and mannitol, exhibited strong negative correlations with CHOL, HDLC, TP and APOB (*q.value* < 0.01). These results indicated significantly abnormal LPCs metabolism in CHD patients, and thus we speculated that it could be beneficial to reduce CHD occurrence by properly increasing intake of these extra LPCs which were significantly decreased in CHD patients.

Supplementary Table 1 | The characteristics of CHD samples and control samples in the study

Characteristics	CHD patients	Control subjects	*p.value**	permuted *p.value*†
Sample number	59	43	-	-
Age (median, range), year	61, 35-79	59, 50-70	0.07	0.12
CKMB (median, range), U/L	7.40, 0.50-32.60	7.13, 2.30-14.70	0.15	0.21
ALB (median, range), g/L	37.90, 26.30-47.20	40.68, 35.40-44.10	4.06E-05‡	8.40E-03‡
ALT (median, range), U/L	26, 7-86	24.47, 10.00-56.00	0.02‡	0.02‡
TP (median, range), g/L	64.80, 55.60-75.10	73.52, 66.70-82.10	1.10E-14‡	1.00E-04‡
AST (median, range), U/L	24.30, 14.95-77.50	26.48, 19.00-45.00	0.17	0.18
CREA (median, range), μmol/L	84.20, 47.00-288.00	80.19, 50.00-123.00	0.09	0.46
HBDH (median, range), U/L	132, 91-544	151.48, 94.00-206.00	0.47	0.48
TRIG (median, range), mmol/L	1.57, 0.67-4.83	1.79, 0.37-5.09	0.96	0.18
LDLC (median, range), mmol/L	2.83, 0.87-4.99	3.32, 1.79-5.64	0.01‡	0.01‡
CHOL (median, range), mmol/L	4.48, 2.42-7.26	5.46, 3.08-8.26	4.69E-05‡	1.00E-04‡
HDLC (median, range), mmol/L	1.01, 0.66-1.59	1.32, 0.90-2.21	1.25E-07‡	1.00E-04‡
APOB (median, range), g/L	0.81, 0.52-1.21	0.92, 0.59-1.38	1.17E-03‡	5.00E-04‡
APOA (median, range), g/L	1.06, 0.56-1.90	1.30, 0.83-2.33	0.01‡	8.10E-03‡
LPA (median, range), mg/L	123.24, 23.93-1386.11	240.68, 18.00-678.00	0.92	0.48

CKMB, creatine kinase MB; ALB, albumin; ALT, Alanine aminotransferase; TP, Total Protein; AST, Aspartate transaminase; CREA, creatinine; HBDH, hydroxybutyrate dehydrogenase; TRIG, triglyceride; LDLC, low-density lipoprotein; CHOL, cholesterol; HDLC, high-density lipoprotein; APOB, apolipoprotein (b); APOA, apolipoprotein (a); LPA, lipoprotein (a).

* *P.value* calculated by Student's t-test. †Permuted *p.value* calculated by PERMANOVA analysis (permutation=1000), permuted *p.value* less than 0.05 indicated significant impacts on metabolic profiles. ‡ Significant clinical biochemical indicators affected metabolic profiles of CHD patients (*p.value* and permutated *p.value* less than 0.05)

Table 1 | Potential plasma biomarkers for discriminating CHD patients from control subjects

m/z	RT(min)*	FC(CHD/ control)†	Adjusted *p.value*‡	*VIP*§	Adduct ion	Formula	Metabolites	Pathways
175.11	1.83	2.14	1.36E-11	1.61	H+	$C_6H_{14}N_4O_2$	L-Arginine||	Arginine and proline metabolism
181.07	9.05	0.43	2.39E-02	1.34	H+	$C_7H_8N_4O_2$	Paraxanthine||	Caffeine metabolism
496.33	13.19	0.67	2.25E-11	1.36	NAN	$C_{24}H_{51}NO_7P$	1-Palmitoylglycerophosphocholine||	Unknown
522.35	13.33	0.72	1.02E-07	1.15	NAN	$C_{26}H_{53}NO_7P$	1-Oleoylglycerophosphocholine||	Unknown
324.04	9.33	8.58	2.08E-12	2.16	Na+	$C_8H_{16}NO_9P$	N-Acetyl-D-glucosamine 6-phosphate||¶	Amino sugar and nucleotide sugar metabolism
468.3	12.69	0.48	1.30E-10	1.8	H+	$C_{22}H_{46}NO_7P$	LysoPC(14:0)||	Glycerophospholipid metabolism
480.34	13.48	0.65	4.12E-10	1.4	H+	$C_{24}H_{50}NO_6P$	LysoPC(P-16:0)||	Glycerophospholipid metabolism
482.32	12.92	0.57	1.37E-08	1.58	H+	$C_{23}H_{48}NO_7P$	LysoPC(15:0)||	Glycerophospholipid metabolism
494.32	12.82	0.54	6.12E-12	1.74	H+	$C_{24}H_{48}NO_7P$	LysoPC(16:1(9Z))||	Glycerophospholipid metabolism
516.31	13.21	0.67	5.12E-07	1.19	H+	$C_{26}H_{46}NO_7P$	LysoPC(18:4(6Z,9Z,12Z,15Z))||	Glycerophospholipid metabolism
518.32	13.19	0.78	7.15E-11	1.04	H+	$C_{26}H_{48}NO_7P$	LysoPC(18:3(9Z,12Z,15Z))||	Glycerophospholipid metabolism
518.32	12.71	0.57	1.75E-06	1.43	H+	$C_{26}H_{48}NO_7P$	LysoPC(18:3(6Z,9Z,12Z))||	Glycerophospholipid metabolism
524.36	14.27	0.54	2.19E-08	1.65	H+	$C_{26}H_{54}NO_7P$	LysoPC(18:0)||	Glycerophospholipid metabolism
524.36	13.69	0.61	2.25E-11	1.64	H+	$C_{26}H_{54}NO_7P$	LysoPC(0:0/18:0)||	Glycerophospholipid metabolism
544.33	13.34	0.76	2.18E-07	1.02	H+	$C_{28}H_{50}NO_7P$	LysoPC(20:4(5Z,8Z,11Z,14Z))||	Glycerophospholipid metabolism
546.35	14.22	0.5	1.11E-09	1.81	H+	$C_{28}H_{52}NO_7P$	LysoPC(20:3(5Z,8Z,11Z))||	Glycerophospholipid metabolism
570.35	13.05	0.67	8.07E-05	1.14	H+	$C_{30}H_{52}NO_7P$	LysoPC(22:5(4Z,7Z,10Z,13Z,16Z))||	Glycerophospholipid metabolism

590.31	12.86	0.6	9.89E-05	1.31	Na+	$C_{30}H_{50}NO_7P$	LysoPC(22:6(4Z,7Z,10Z,13Z,16Z,19Z))‖	Glycerophospholipid metabolism

*Retention time.†Fold change (FC=mean of metabolite ion intensity in CHD patients/ mean of metabolite ion intensity in healthy controls). When FC was larger than 1, it represented that the biomarker was enriched in CHD patients. When FC was lower than 1, it represented that the biomarker was enriched in healthy individuals. ‡Adjusted *p.value* calculated by the two-tailed Wilcoxon rank-sum tests after false discovery rate correction.§VIP (Variable Importance for Projection), one indicator reflecting the capability of the variables to explain Y.‖Metabolites matched with the online database.¶Metabolites which have matched characteristic peaks but mismatched retention time with commercial available reference standards.

Table 2 | Potential urine biomarkers for discriminating CHD patients from control subjects

m/z	RT(min)*	FC(CHD/ control)†	Adjusted *p.value*‡	*VIP*§	Adduct ion	Formula	Metabolites	Pathways
132.07	1.9	0.41	2.36E-02	1.41	H+	$C_4H_9N_3O_2$	Creatine#	Arginine and proline metabolism
205.06	1.79	8.45	9.69E-12	2.87	Na+	$C_6H_{14}O_6$	Mannitol#	Fructose and mannose metabolism
318.29	11.37	0.39	1.65E-05	2.51	H+	$C_{18}H_{39}NO_3$	Phytosphingosine#	Sphingolipid metabolism
324.04	10.25	165.99	4.28E-14	2.3	Na+	$C_8H_{16}NO_9P$	N-Acetyl-D-glucosamine 6-phosphate¶#	Amino sugar and nucleotide sugar metabolism

*Retention time.†Fold change (FC=mean of metabolite ion intensity in CHD patients/ mean of metabolite ion intensity in healthy controls). When FC was larger than 1, it represented that the biomarker was enriched in CHD patients. When FC was lower than 1, it represented that the biomarker was enriched in healthy individuals .‡Adjusted *p.value* calculated by the two-tailed Wilcoxon rank-sum tests after false discovery rate correction.§VIP (Variable Importance for Projection), one indicator reflecting the capability of the variables to explain Y.#Metabolites matched with commercial available reference standards.¶Metabolites which have matched characteristic peaks but mismatched retention time with commercial available reference standards.

Supplementary Table 2 | Seven common potential biomarkers in plasma and urine samples

Category	m/z	RT(min)*	FC(CHD/ control)†	Adjusted *p.value*‡	*VIP* §	NAME	PATHWAYS
Potential biomarkers in plasma	185.04	10.24	0.33	3.69E-07	1.57	unknown	unknown
	311.05	10.46	29.38	1.94E-13	2.34	unknown	unknown
	310.04	10.46	28.41	1.65E-13	2.35	unknown	unknown
	309.05	10.46	26.40	9.08E-14	2.35	unknown	unknown
	202.04	9.72	0.38	2.70E-04	1.41	unknown	unknown
	324.04	9.33	8.58	2.08E-12	2.16	N-Acetyl-D-glucosamine 6-phosphate‖¶	Amino sugar and nucleotide sugar metabolism
	125.01	10.46	7.86	1.50E-11	2.31	unknown	unknown
Potential biomarkers in urine	185.04	10.30	0.34	1.70E-03	1.62	unknown	unknown
	311.05	10.51	59.41	3.75E-14	2.26	unknown	unknown
	310.04	11.25	110.51	5.70E-14	2.47	unknown	unknown
	309.05	11.28	12.65	1.28E-07	2.03	unknown	unknown
	202.04	9.77	0.58	8.10E-04	1.06	unknown	unknown
	324.04	10.25	165.99	4.28E-14	2.30	N-Acetyl-D-glucosamine 6-phosphate‖¶	Amino sugar and nucleotide sugar metabolism
	125.01	11.18	158.93	5.11E-14	2.42	unknown	unknown

* Retention time. † Fold change (FC=mean of metabolite ion intensity in CHD patients/ mean of metabolite ion intensity in healthy controls). When FC was larger than 1, it represented that the biomarker was enriched in CHD patients. When FC was lower than 1, it represented that the biomarker was enriched in healthy individuals. ‡ Adjusted *p.value* calculated by the two-tailed Wilcoxon rank-sum tests after false discovery rate correction. §VIP (Variable Importance for Projection), one indicator reflecting the capability of the variables to explain Y. || Metabolites matched with commercial available reference standards. ¶ Metabolites matched characteristic peaks but mismatching retention time with commercial available reference standards.

Supplementary Table 3 | Spearman correlation analysis of the 7 common metabolites

Plasma	Urine	Coefficient	*p.value**	Adjusted *p.value* †	Enrichment
311.05	311.05	0.812	3.18E-25	3.12E-24	CHD
310.04	310.04	0.803	3.00E-24	1.63E-23	CHD
324.04	324.04	0.747	1.83E-19	5.60E-19	CHD
125.01	125.01	0.687	1.38E-15	3.22E-15	CHD
202.04	202.04	0.642	0	0	Control
309.05	309.05	0.556	1.96E-09	4.18E-09	CHD
185.04	185.04	0.547	4.48E-09	8.78E-09	Control

* *P.value* calculated by spearman correlation analysis.

†Adjusted *p.value* calculated by false discovery rate correction.

Table 3 | AUC results of plasma training and validation datasets

Name	Plasma training datasets			Plasma validation datasets		
	*AUC**	*p.value*†	FC(CHD/control)‡	*AUC**	*p.value*†	FC(CHD/control)‡
LysoPC(18:3(6Z,9Z,12Z))	0.79	2.72E-07	0.57	0.91	6.72E-09	0.26
LysoPC(P-16:0)	0.88	2.35E-11	0.65	0.88	2.57E-21	0.58
LysoPC(15:0)	0.85	6.56E-11	0.57	0.88	6.94E-19	0.51
1-Palmitoylglycerophosphocholine	0.91	4.56E-14	0.67	0.88	1.45E-20	0.65
LysoPC(14:0)	0.89	2.86E-11	0.48	0.84	5.94E-15	0.49
LysoPC(16:1(9Z))	0.92	7.78E-14	0.54	0.83	5.57E-16	0.62
LysoPC(0:0/18:0)	0.91	7.98E-15	0.61	0.83	6.05E-13	0.42
LysoPC(18:4(6Z,9Z,12Z,15Z))	0.81	3.81E-08	0.67	0.79	1.65E-12	0.55
LysoPC(22:6(4Z,7Z,10Z,13Z,16Z,19Z))	0.74	1.30E-05	0.6	0.76	7.34E-10	0.69

L-Arginine	0.91	1.52E-11	2.14	0.75	7.63E-10	1.64
LysoPC(22:5(4Z,7Z,10Z,13Z,16Z))	0.74	4.70E-06	0.67	0.73	4.00E-06	0.73
LysoPC(18:3(9Z,12Z,15Z))	0.9	9.90E-14	0.78	0.71	5.42E-06	0.77
LysoPC(18:0)	0.84	2.98E-10	0.54	0.7	1.54E-06	0.83
1-Oleoylglycerophosphocholine	0.83	5.50E-09	0.72	0.64	1.24E-03	0.88
Paraxanthine	0.62	7.33E-03	0.43	0.63	4.41E-04	0.33
LysoPC(20:4(5Z,8Z,11Z,14Z))	0.82	1.38E-08	0.76	0.63	1.39E-03	0.85
LysoPC(20:3(5Z,8Z,11Z))	0.87	5.76E-11	0.5	0.5	8.67E-01	1.01
N-Acetyl-D-glucosamine 6-phosphate	0.93	7.55E-10	8.58	-	-	-

**AUC* calculated by online tool – ROCCET (http://www.roccet.ca).†*P.value* calculated by T-test.‡*Fold change.*

Table 4 | AUC results of urine training and validation datasets

Name	Urine training datasets			Urine validation datasets		
	*AUC**	*p.value* †	FC(CHD/control) ‡	*AUC**	*p.value* †	FC(CHD/control) ‡
N-Acetyl-D-glucosamine 6-phosphate	0.96	4.44E-09	165.99	0.88	8.29E-16	36.91
Mannitol	0.92	2.11E-11	8.45	0.81	3.99E-08	2.62
Creatine	0.66	2.85E-02	0.41	0.57	7.59E-01	0.94
Phytosphingosine	0.78	1.55E-05	0.39	0.55	2.31E-01	1.05

**AUC* calculated by online tool – ROCCET (http://www.roccet.ca).†*P.value* calculated by T-test.‡*Fold change.*

Supplementary Table 4 | Spearman correlation analysis of clinical data and identified biomarkers

Potential Biomarkers	Clinical biochemical indicators	Coefficient	*p.vlaue**	Adjusted *p.value* †
LysoPC(18:0)	CHOL	0.52	2.39E-08	7.89E-07
LysoPC(P-16:0)	HDLC	0.56	8.45E-10	7.39E-08
LysoPC(18:0)	HDLC	0.55	2.56E-09	1.29E-07
LysoPC(14:0)	TP	0.59	5.19E-11	1.71E-08
LysoPC(18:0)	TP	0.57	3.13E-10	5.16E-08
LysoPC(16:1(9Z))	TP	0.56	9.08E-10	7.39E-08
LysoPC(15:0)	TP	0.56	1.12E-09	7.39E-08
1-Palmitoylglycerophosphocholine	TP	0.55	2.74E-09	1.29E-07
LysoPC(18:3(9Z,12Z,15Z))	TP	0.54	6.40E-09	2.64E-07
LysoPC(20:3(5Z,8Z,11Z))	TP	0.53	8.31E-09	3.05E-07
LysoPC(18:4(6Z,9Z,12Z,15Z))	TP	0.51	3.63E-08	1.09E-06

* *P.value* calculated by spearman correlation analysis.

†Adjusted *p.value* calculated by false discovery rate correction.

Thus the inventors have identified and validated 7 plasma metabolites and 2 urine metabolites for early and non-invasive diagnosis of CHD by a random forest model based on the associated metabolites. And the inventors have constructed a method to evaluate the risk of CHD based on these associated metabolites.

Although explanatory embodiments have been shown and described in detailed, it would be appreciated by those skilled in the art that the above embodiments are illustrative, and are not intended to limit the present disclosure in any way, and that changes, alternatives, and modifications can be made to the embodiments without departing from the spirit, principles and scope of the present disclosure.

WHAT IS CLAIMED IS:

1. A biomarker composition, which comprises one or more selected from the group consisting of :
Biomarker 1, for which m/z is 518.32±1.00, retention time (RT) is 12.71±1.00min;
Biomarker 2, for which m/z is 480.34±1.00, retention time (RT) is 13.48±1.00min;
Biomarker 3, for which m/z is 482.32±1.00, retention time (RT) is 12.92±1.00min;
Biomarker 4, for which m/z is 496.33±1.00, retention time (RT) is 13.19±1.00min;
Biomarker 5, for which m/z is 468.3±1.00, retention time (RT) is 12.69±1.00min;
Biomarker 6, for which m/z is 494.32±1.00, retention time (RT) is 12.82±1.00min;
Biomarker 7, for which m/z is 524.36±1.00 retention time (RT) is 13.69±1.00min;
Biomarker 8, for which m/z is 516.31±1.00, retention time (RT) is 13.21±1.00min;
Biomarker 9, for which m/z is 590.31±1.00, retention time (RT) is 12.86±1.00min;
Biomarker 10, for which m/z is 546.35±1.00, retention time (RT) is 14.22±1.00min;
Biomarker 11, for which m/z is 570.35±1.00, retention time (RT) is 13.05±1.00min;
Biomarker 12, for which m/z is 518.32±1.00, retention time (RT) is 13.19±1.00min;
Biomarker 13, for which m/z is 524.36±1.00, retention time (RT) is 14.27±1.00min;
Biomarker 14, for which m/z is 522.35±1.00, retention time (RT) is 13.33±1.00min;
Biomarker 15, for which m/z is 181.07±1.00, retention time (RT) is 9.05±1.00min;
Biomarker 16, for which m/z is 544.33±1.00, retention time (RT) is 13.34±1.00min;
Biomarker 17, for which m/z is 175.11±1.00, retention time (RT) is 1.83±1.00min; and
Biomarker 18, for which m/z is 324.04±1.00, retention time (RT) is 9.33±1.00min.

2. The biomarker composition according to claim 1, which comprises one or more selected from the group consisting of:
Biomarker 1, for which m/z is 518.32±1.00, retention time (RT) is 12.71±1.00min;
Biomarker 2, for which m/z is 480.34±1.00, retention time (RT) is 13.48±1.00min;
Biomarker 3, for which m/z is 482.32±1.00, retention time (RT) is 12.92±1.00min;
Biomarker 4, for which m/z is 496.33±1.00, retention time (RT) is 13.19±1.00min;
Biomarker 5, for which m/z is 468.3±1.00, retention time (RT) is 12.69±1.00min;
Biomarker 6, for which m/z is 494.32±1.00, retention time (RT) is 12.82±1.00min; and
Biomarker 7, for which m/z is 524.36±1.00, retention time (RT) is 13.69±1.00min;

3. The biomarker composition according to claim 1, wherein,
Biomarker 1 is LysoPC(18:3(6Z,9Z,12Z));
Biomarker 2 is LysoPC(P-16:0);
Biomarker 3 is LysoPC(15:0);
Biomarker 4 is 1-Palmitoylglycerophosphocholine;
Biomarker 5 is LysoPC(14:0);
Biomarker 6 is LysoPC(16:1(9Z));
Biomarker 7 is LysoPC(0:0/18:0);
Biomarker 8 is LysoPC(18:4(6Z,9Z,12Z,15Z));

Biomarker 9 is LysoPC(22:6(4Z,7Z,10Z,13Z,16Z,19Z));
Biomarker 10 is LysoPC(20:3(5Z,8Z,11Z));
Biomarker 11 is LysoPC(22:5(4Z,7Z,10Z,13Z,16Z));
Biomarker 12 is LysoPC(18:3(9Z,12Z,15Z));
Biomarker 13 is LysoPC(18:0);
Biomarker 14 is 1-Oleoylglycerophosphocholine;
Biomarker 15 is Paraxanthine;
Biomarker 16 is LysoPC(20:4(5Z,8Z,11Z,14Z));
Biomarker 17 is L-Arginine; and/or
Biomarker 18 is N-Acetyl-D-glucosamine 6-phosphate.

4. A reagent composition, which comprises the reagents used for detection of each of the biomarker composition according to any item of claims 1-3.

5. The reagent composition according to claim 4, wherein the reagents comprise substances used in mass spectrometry for detection of the biomarker composition according to any item of claims 1-3.

6. A kit, which comprises the biomarker composition according to any item of claims 1-3 and/or the reagent composition according to claim 4 or 5.

7. A use of the biomarker composition according to any item of claims 1-3 and/or the reagent composition according to claim 4 or 5 in the preparation of a kit, wherein the kit is used for evaluation of the risk of CHD in a subject, or for use in diagnosis of CHD in a subject.

8. The use according to claim 7, wherein the evaluation or diagnosis comprises the following steps: 1) determining the level of each of the biomarkers of the biomarker composition according to any item of claims 1-3 in a sample from the subject; 2) comparing the level of step 1) with a reference dataset or a reference value (for example a reference value of healthy controls); preferably, the reference dataset comprises the level of the biomarker of the biomarker composition according to any item of claims 1-3 in a sample from CHD patients and healthy controls.

9. The use according to claim 8, wherein the sample is selected from blood, plasma and serum.

10. The use according to claim 8 or 9, wherein the comparing the level of step 1) with a reference dataset further comprises the step of executing a multivariate statistical model to output the probability of illness; preferably, the multivariate statistical model is random forest model.

11. The use according to claim 10, wherein the subject is determined as being at risk of CHD or having CHD if the probability of illness $\geqslant 0.5$.

12. The use according to claim 8 or 9, wherein, when compared with a reference value, the decrease of one or more Biomarkers selected from the group consisting of Biomarker 1-16, preferably, from the group consisting of Biomarker 1-7, indicates that the subject is in the risk of CHD or has CHD.

13. The use according to claim 8 or 9, wherein, when compared with a reference value, the increase of Biomarker 17 and/or Biomarker 18 indicates that the subject is in the risk of CHD or has CHD.

14. The use according to claim 8 or 9, wherein the step of determining the level of each of the biomarkers are carried out by mass spectrometry, preferably, mass spectrometry in conjunction with chromatography, such as gas chromatography mass spectrometry (GC-MS) or liquid chromatography mass spectrometry (LC-MS).

15. The use according to claim 8, wherein the method further comprises the step of processing the sample before step 1).

16. A method for evaluation of the risk of CHD in a subject or for diagnosis of CHD in a subject, wherein the method comprises the following steps: 1) determining the level of each of the biomarkers of the biomarker composition according to any item of claims 1-3 in a sample from the subject; 2) comparing the level of step 1) with a reference dataset or a reference value (for example a reference value of healthy controls); preferably, the reference dataset comprises the level of the biomarker of the biomarker composition according to any item of claims 1-3 in a sample from CHD patients and healthy controls.

17. The method according to claim 16, wherein the sample is selected from blood, plasma and serum.

18. The method according to claim 16 or 17, wherein the comparing the level of step 1) with a reference dataset further comprises the step of executing a multivariate statistical model to output the probability of illness; preferably, the multivariate statistical model is random forest model.

19. The method according to claim 16 or 17, wherein the subject is determined as being at risk of CHD or having CHD if the probability of illness $\geqslant 0.5$.

20. The method according to claim 16 or 17, wherein, when compared with a reference value, the decrease of one or more Biomarkers selected from the group consisting of Biomarker 1-16, preferably, from the group consisting of Biomarker 1-7, indicates that the subject is in the risk of CHD or has CHD.

21. The method according to claim 16 or 17, wherein, when compared with a reference

value, the increase of Biomarker 17 and/or Biomarker 18 indicates that the subject is in the risk of CHD or has CHD.

22. The method according to claim 16 or 17, wherein the step of determining the level of each of the biomarkers are carried out by mass spectrometry, preferably, mass spectrometry in conjunction with chromatography, such as gas chromatography mass spectrometry (GC-MS) or liquid chromatography mass spectrometry (LC-MS).

23. The method according to claim 16 or 17, wherein the method further comprises the step of processing the samples before step 1).

24. The biomarker composition according to any item of claims 1-3 or the reagent composition according to claim 4 or 5, for use in a method of evaluation of the risk of CHD in a subject or for diagnosis of CHD in a subject.

25. The biomarker composition according to claim 24, wherein the method of evaluation or diagnosis comprises the following steps: 1) determining the level of each of the biomarkers of the biomarker composition according to any item of claims 1-3 in a sample from the subject; 2) comparing the level of step 1) with a reference dataset or a reference value (for example a reference value of healthy controls); preferably, the reference dataset comprises the level of the biomarker of the biomarker composition according to any item of claims 1-3 in a sample from CHD patients and healthy controls.

26. The biomarker composition according to claim 25, wherein the sample is selected from blood, plasma and serum.

27. The biomarker composition according to claim 25 or 26, wherein the comparing the level of step 1) with a reference dataset further comprises the step of executing a multivariate statistical model to output the probability of illness; preferably, the multivariate statistical model is random forest model.

28. The biomarker composition according to claim 25 or 26, wherein the subject is determined as being at risk of CHD or having CHD if the probability of illness $\geqslant$ 0.5.

29. The biomarker composition according to claim 25 or 26, wherein, when compared with a reference value, the decrease of one or more Biomarkers selected from the group consisting of Biomarker 1-16, preferably, from the group consisting of Biomarker 1-7, indicates that the subject is in the risk of CHD or has CHD.

30. The biomarker composition according to claim 25 or 26, wherein, when compared with a reference value, the increase of Biomarker 17 and/or Biomarker 18 indicates that the subject is in the risk of CHD or has CHD.

31. The biomarker composition according to claim 25 or 26, wherein the step of determining the level of each of the biomarkers are carried out by mass spectrometry, preferably, mass spectrometry in conjunction with chromatography, such as gas chromatography mass spectrometry (GC-MS) or liquid chromatography mass spectrometry (LC-MS).

32. The biomarker composition according to claim 25 or 26, wherein the method further comprises the step of processing the samples before step 1).

33. A use of the biomarker composition according to any item of claims 1-3 and/or the reagent composition according to claim 4 or 5 in the preparation of a kit, wherein the kit is used for screening candidate compounds for treatment of CHD in a subject, or for evaluating the effect of the treatment of CHD in a subject.

34. The use according to claim 33, wherein the screening or evaluation comprises the following steps: 1) determining the level of each of the biomarkers of the biomarker composition according to any item of claims 1-3 in a sample from the subject after administering the candidate compounds or the treatment to the subject; 2) comparing the level of step 1) with the level of the above mentioned biomarker before administering the candidate compounds or the treatment to the subject.

35. The use according to claim 34, wherein the sample is selected from blood, plasma and serum.

36. The use according to claim 34 or 35, wherein the increase of one or more Biomarkers selected from the group consisting of Biomarker 1-16, preferably, from the group consisting of Biomarker 1-7, indicates that the compound is a candidate compound for treatment of CHD in a subject or the treatment of CHD in the subject is effective.

37. The use according to claim 34 or 35, wherein the decrease of Biomarker 17 and/or Biomarker 18 indicates that the compound is a candidate compound for treatment of CHD in a subject or the treatment of CHD in the subject is effective.

38. The use according to claim 34 or 35, wherein the step of determining the level of each of the biomarkers are carried out by mass spectrometry, preferably, mass spectrometry in conjunction with chromatography, such as gas chromatography mass spectrometry (GC-MS) or liquid chromatography mass spectrometry (LC-MS).

39. The use according to claim 34 or 35, wherein the method further comprises the step of processing the samples before step 1).

40. A method for screening candidate compounds for treatment of CHD in a subject or for evaluation of the effect of the treatment of CHD in a subject, wherein the method

comprises the following steps: 1) determining the level of each of the biomarkers of the biomarker composition according to any item of claims 1-3 in a sample from the subject after administering the candidate compounds or the treatment to the subject; 2) comparing the level of step 1) with the level of the above mentioned biomarker before administering the candidate compounds or the treatment to the subject.

41. The method according to claim 40, wherein the sample is selected from blood, plasma and serum.

42. The method according to claim 40 or 41, wherein the increase of one or more Biomarkers selected from the group consisting of Biomarker 1-16, preferably, from the group consisting of Biomarker 1-7, indicates that the compound is a candidate compound for treatment of CHD in a subject or the treatment of CHD in the subject is effective.

43. The method according to claim 40 or 41 the decrease of Biomarker 17 and/or Biomarker 18 indicates that the compound is a candidate compound for treatment of CHD in a subject or the treatment of CHD in the subject is effective.

44. The method according to claim 40 or 41, wherein the step of determining the level of each of the biomarkers are carried out by mass spectrometry, preferably, mass spectrometry in conjunction with chromatography, such as gas chromatography mass spectrometry (GC-MS) or liquid chromatography mass spectrometry (LC-MS).

45. The method according to claim 40 or 41, wherein the method further comprises the step of processing the samples before step 1).

46. The biomarker composition according to any item of claims 1-3 or the reagent composition according to claim 4 or 5, for use in a method of screening candidate compounds for treatment of CHD in a subject or for evaluation of the effect of the treatment of CHD in a subject.

47. The biomarker composition according to claim 46, wherein the method of screening and evaluation comprises the following steps: 1) determining the level of each of the biomarkers of the biomarker composition according to any item of claims 1-3 in a sample from the subject after administering the candidate compounds or the treatment to the subject; 2) comparing the level of step 1) with the level of the above mentioned biomarker before administering the candidate compounds or the treatment to the subject.

48. The biomarker composition according to claim 47, wherein the sample is selected from blood, plasma and serum.

49. The biomarker composition according to claim 47 or 48, wherein the increase of one

or more Biomarkers selected from the group consisting of Biomarker 1-16, preferably, from the group consisting of Biomarker 1-7, indicates that the compound is a candidate compound for treatment of CHD in a subject or the treatment of CHD in the subject is effective.

50. The biomarker composition according to claim 47 or 48, wherein the decrease of Biomarker 17 and/or Biomarker 18 indicates that the compound is a candidate compound for treatment of CHD in a subject or the treatment of CHD in the subject is effective.

51. The biomarker composition according to claim 47 or 48, wherein the step of determining the level of each of the biomarkers are carried out by mass spectrometry, preferably, mass spectrometry in conjunction with chromatography, such as gas chromatography mass spectrometry (GC-MS) or liquid chromatography mass spectrometry (LC-MS).

52. The biomarker composition according to claim 47 or 48, wherein the method further comprises the step of processing the samples before step 1).

53. A method for setting up a mass spectrometry model for evaluation of the risk of CHD in a subject or for diagnosis of CHD in a subject, which comprises the step of identifying the differentially expressed substance in a blood sample between CHD patients and healthy controls, wherein the differentially expressed substance comprises one or more selected from the group consisting of Biomarker 1-18, preferably, from the group consisting of of Biomarker 1-7.

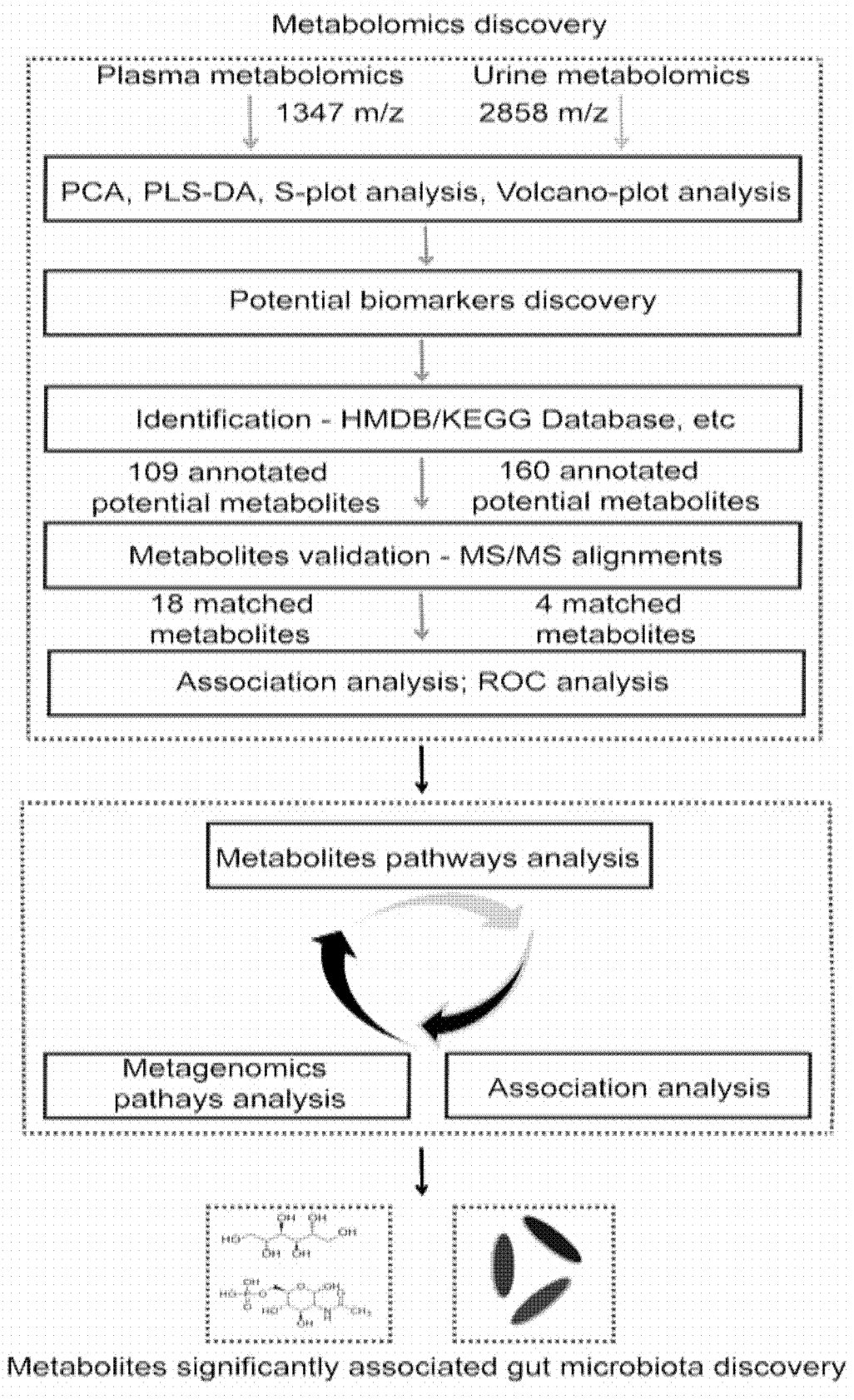

Metabolomics discovery
Plasma metabolomics
1347 m/z
Urine metabolomics
2858 m/z
PCA, PLS-DA, S-plot analysis, Volcano-plot analysis
Potential biomarkers discovery
Identification - HMDB/KEGG Database, etc
109 annotated potential metabolites
160 annotated potential metabolites
Metabolites validation - MS/MS alignments
18 matched metabolites
4 matched metabolites
Association analysis; ROC analysis
Metabolites pathways analysis
Metagenomics pathays analysis
Association analysis
Metabolites significantly associated gut microbiota discovery

Fig.1

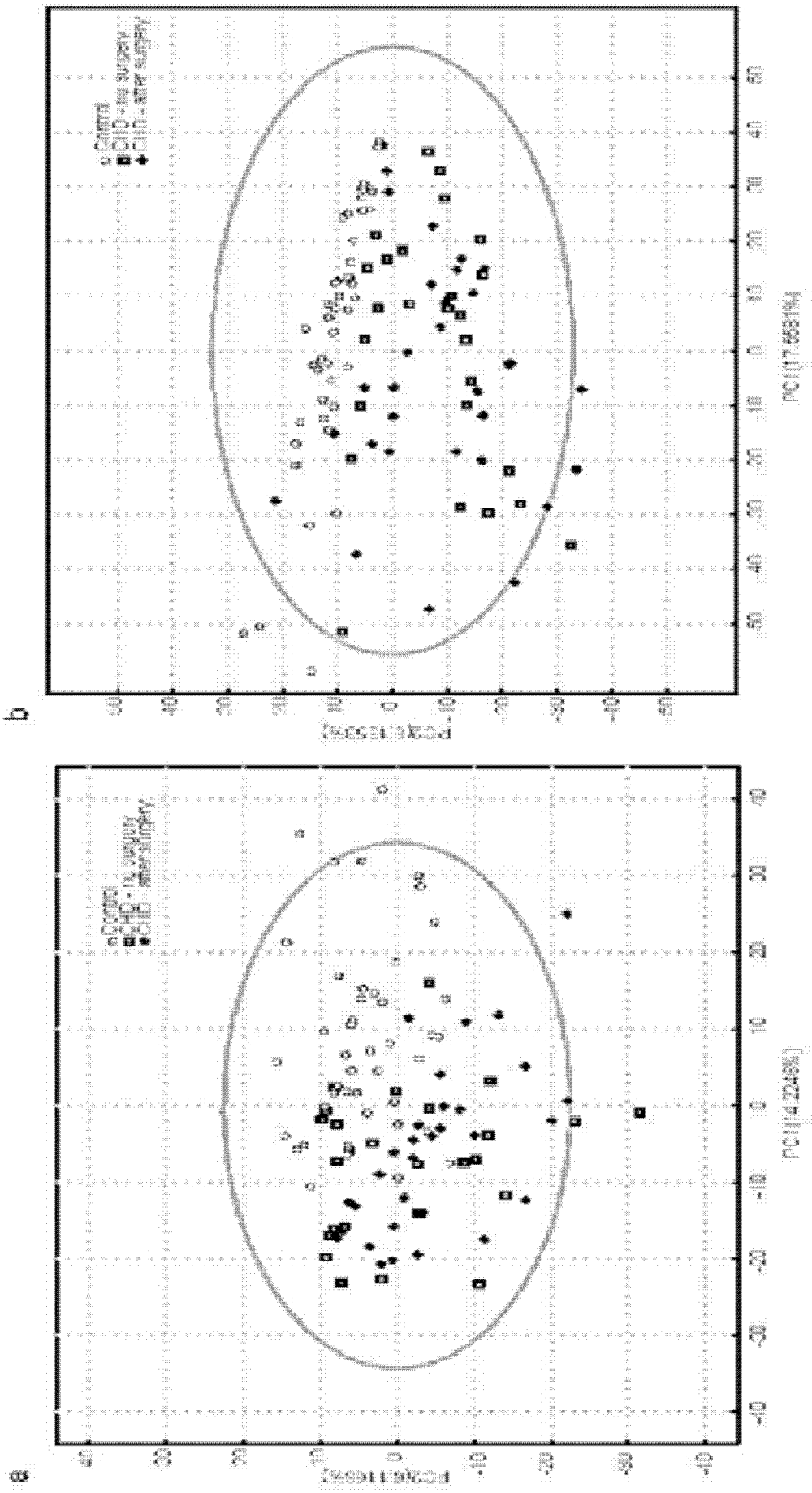

b
Control
CHD - the severity
CHD - other severity
PC2(6.1653%)
PC1(17.5581%)
a
Control
CHD - the severity
CHD - other severity
PC2(6.1166%)
PC1(14.2246%)

Fig.2

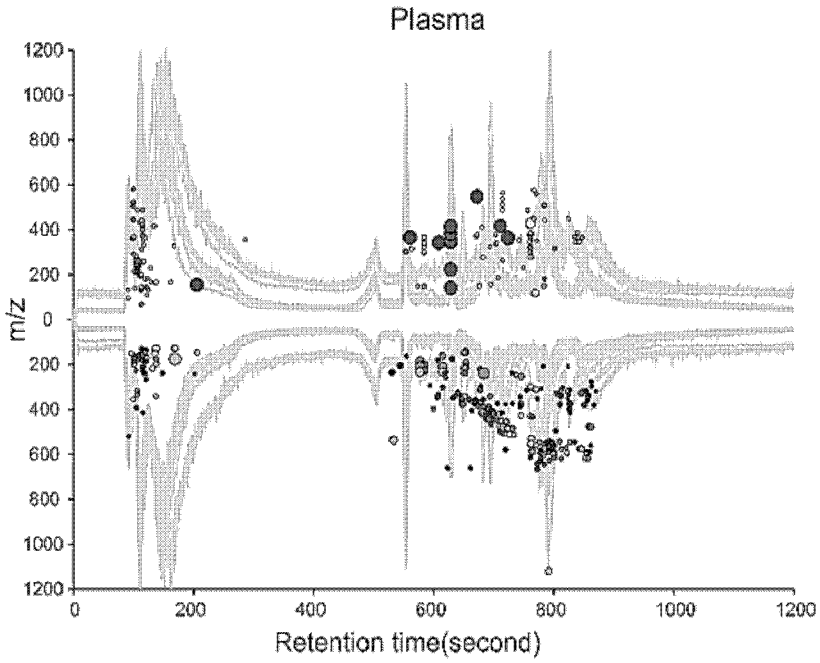

Plasma
m/z
1200
1000
800
600
400
200
0
200
400
600
800
1000
1200
0
200
400
600
800
1000
1200
Retention time(second)

Fig.3

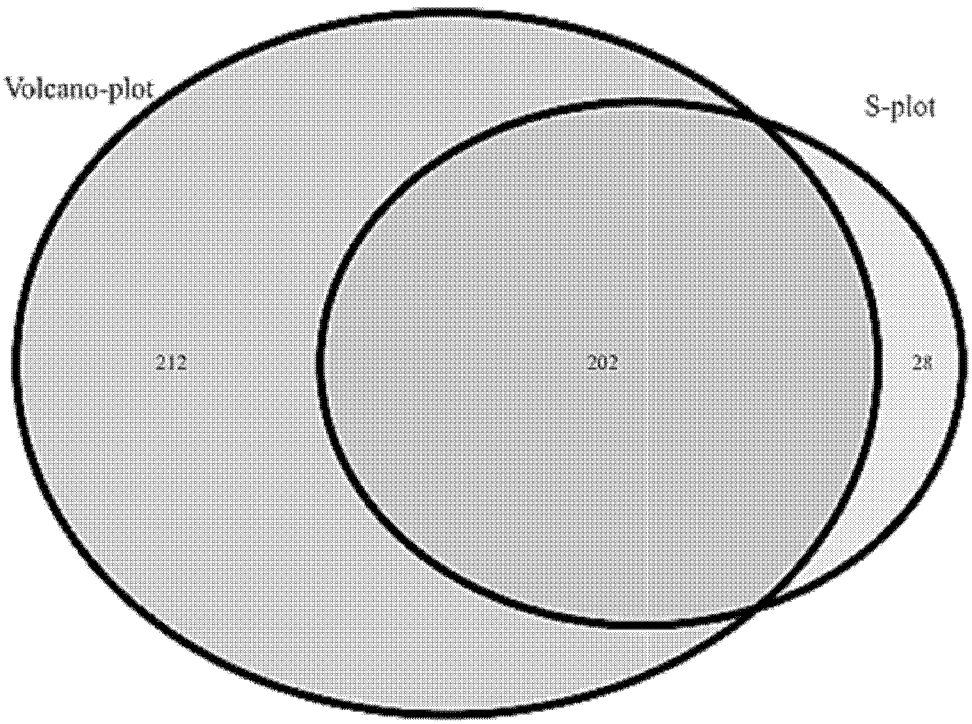

Volcano-plot
S-plot
212
202
28

Fig.4

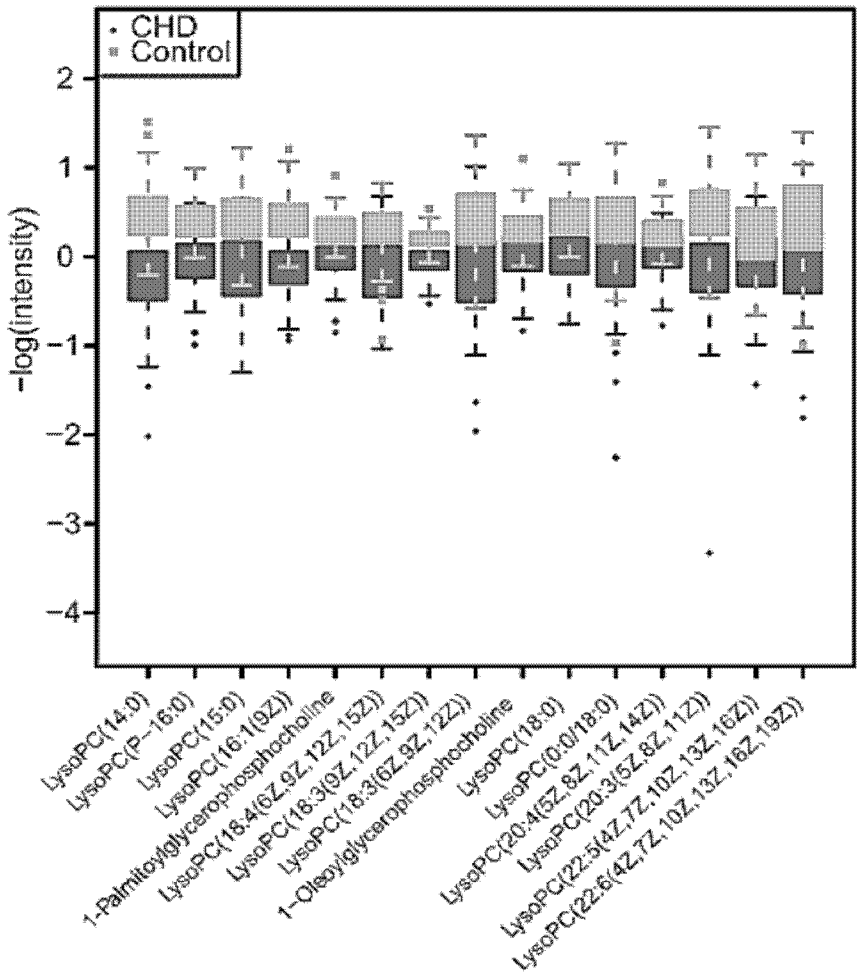

CHD
Control
-log(intensity)
2
1
0
−1
−2
−3
−4
LysoPC(14:0)
LysoPC(P−16:0)
LysoPC(15:0)
LysoPC(16:1(9Z))
1-Palmitoylglycerophosphocholine
LysoPC(18:4(6Z,9Z,12Z,15Z))
LysoPC(18:3(9Z,12Z,15Z))
LysoPC(18:3(6Z,9Z,12Z))
1-Oleoylglycerophosphocholine
LysoPC(18:0)
LysoPC(0:0/18:0)
LysoPC(20:4(5Z,8Z,11Z,14Z))
LysoPC(20:3(5Z,8Z,11Z))
LysoPC(22:5(4Z,7Z,10Z,13Z,16Z))
LysoPC(22:6(4Z,7Z,10Z,13Z,16Z,19Z))

Fig.5

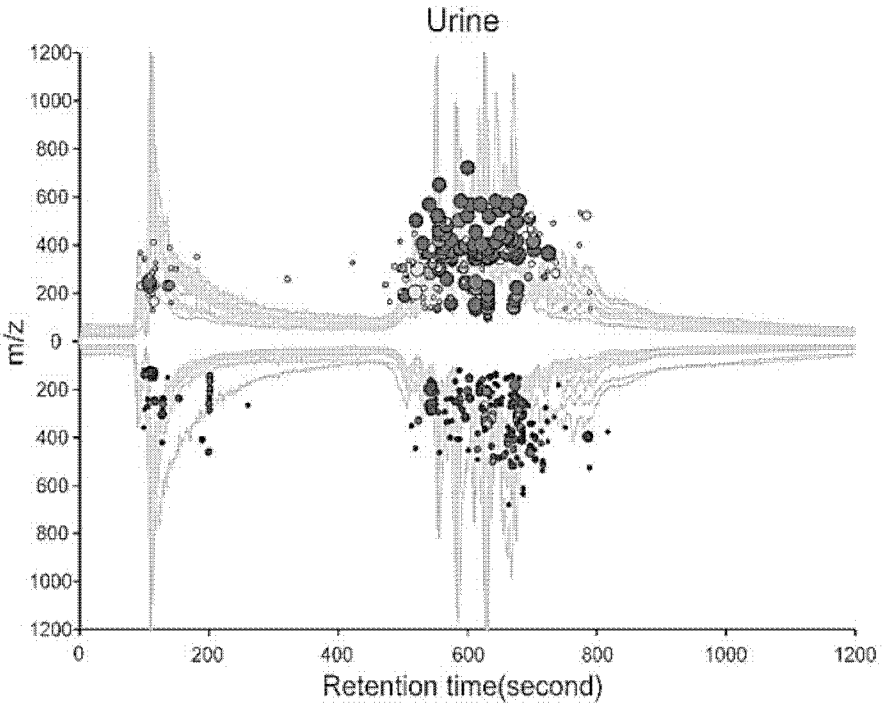

Urine
m/z
1200
1000
800
600
400
200
0
200
400
600
800
1000
1200
0
200
400
600
800
1000
1200
Retention time(second)

Fig.6

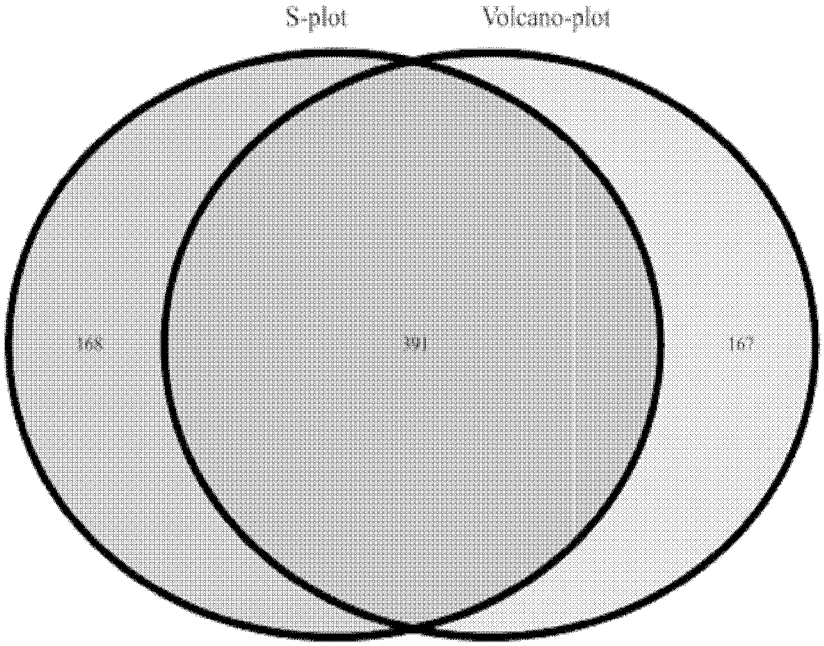


Fig.7

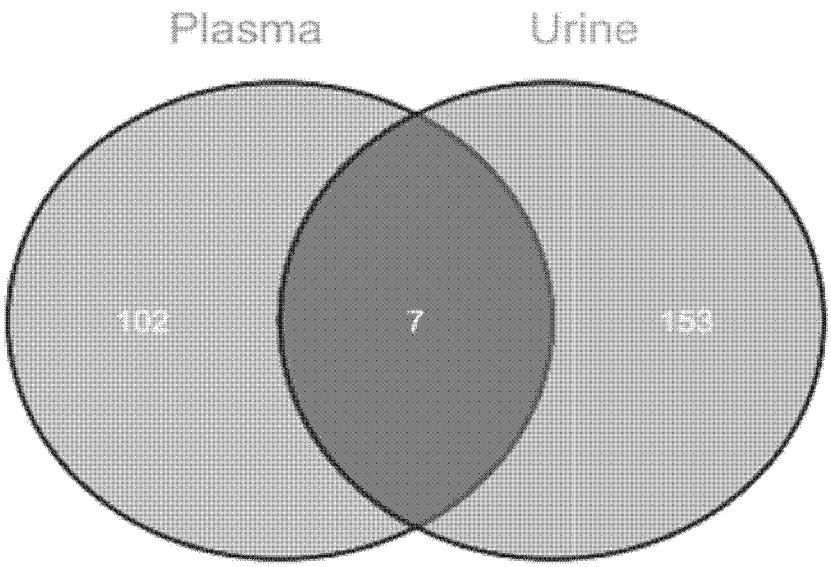
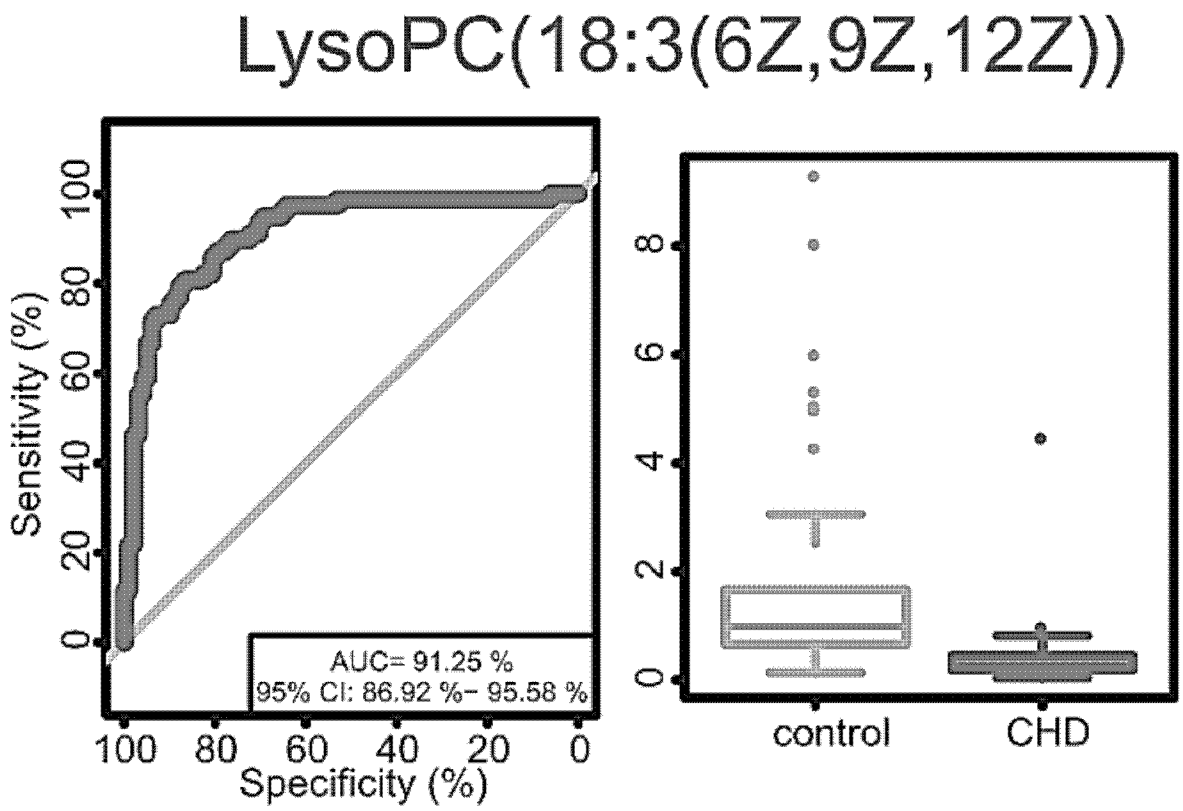
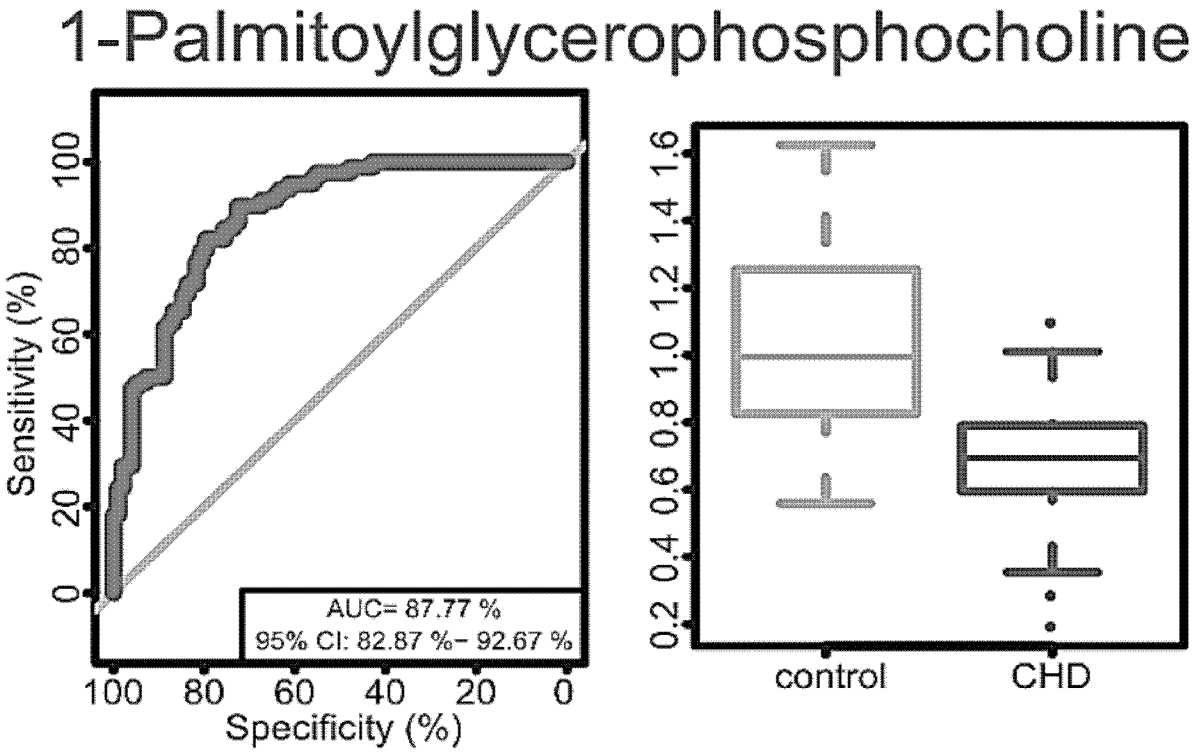


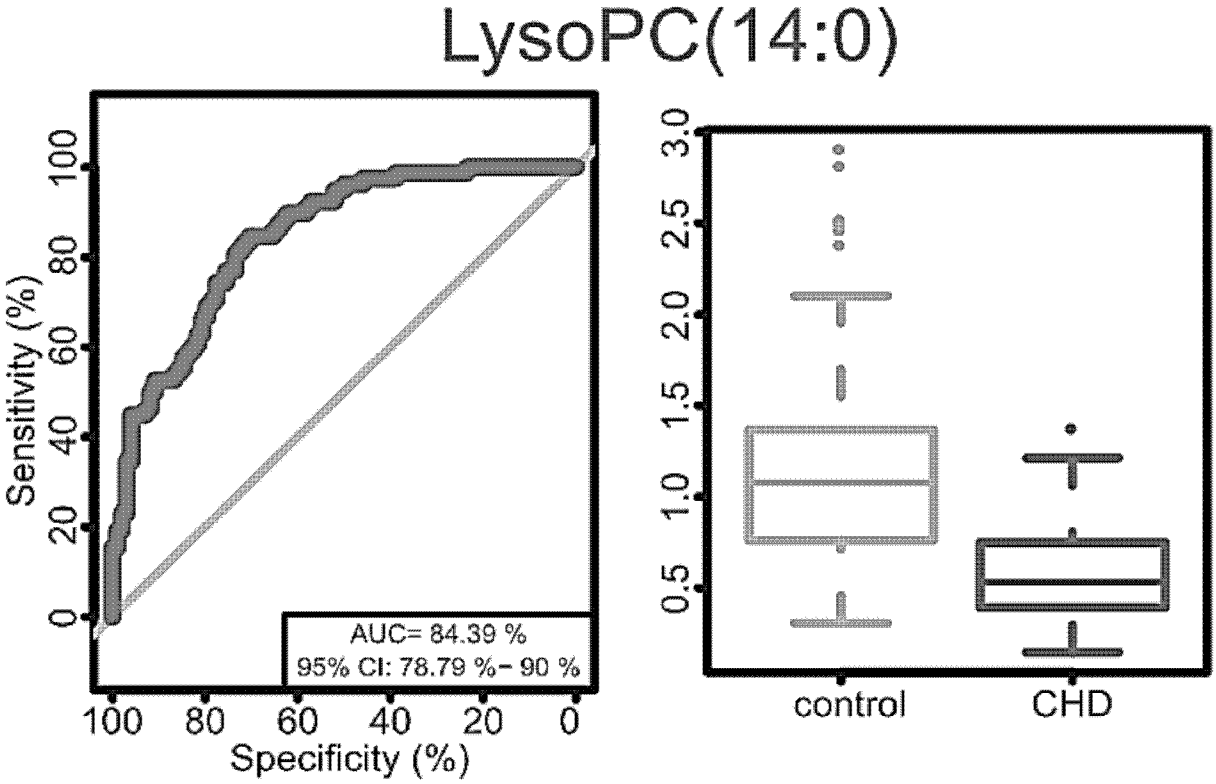
Fig.8


LysoPC(18:3(6Z,9Z,12Z))
Sensitivity (%)
100
80
60
40
20
0
AUC= 91.25 %
95% CI: 86.92 %- 95.58 %
100 80 60 40 20 0
Specificity (%)
0 2 4 6 8
control
CHD

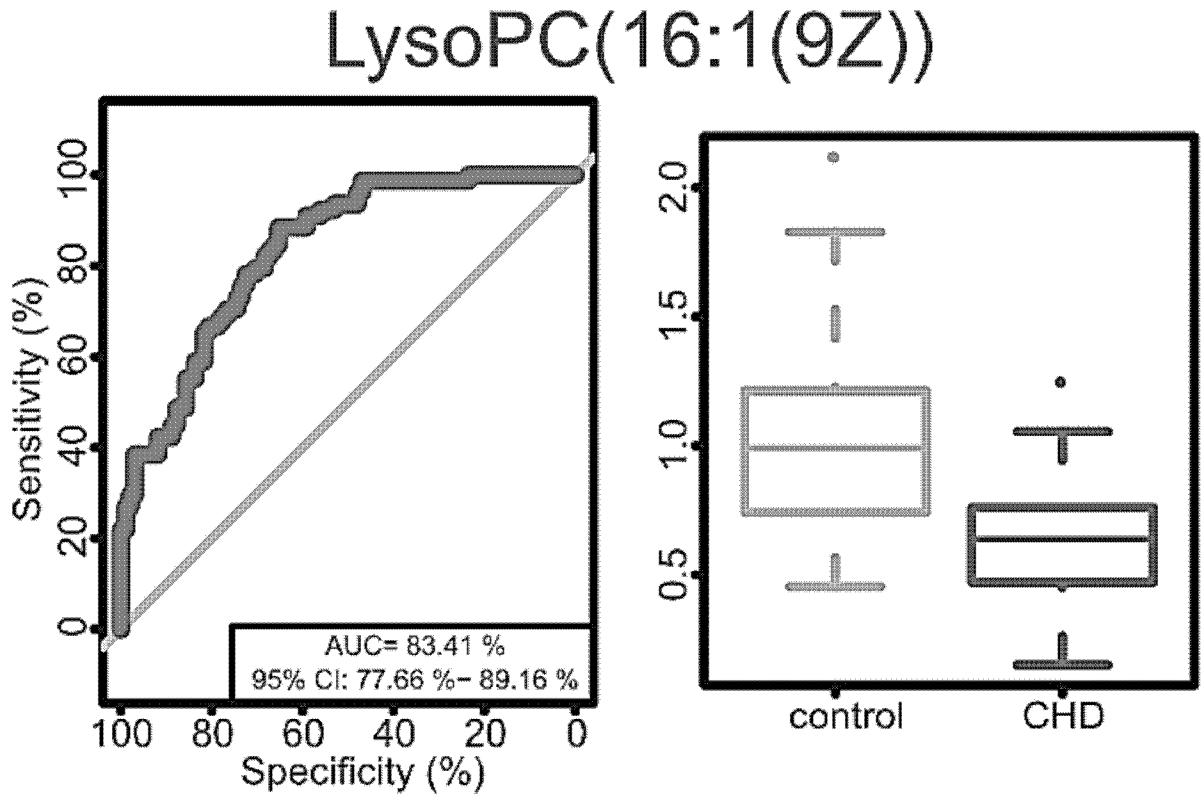
a


1-Palmitoylglycerophosphocholine
Sensitivity (%)
100
80
60
40
20
0
AUC= 87.77 %
95% CI: 82.87 %- 92.67 %
100 80 60 40 20 0
Specificity (%)
0.2 0.4 0.6 0.8 1.0 1.2 1.4 1.6
control
CHD

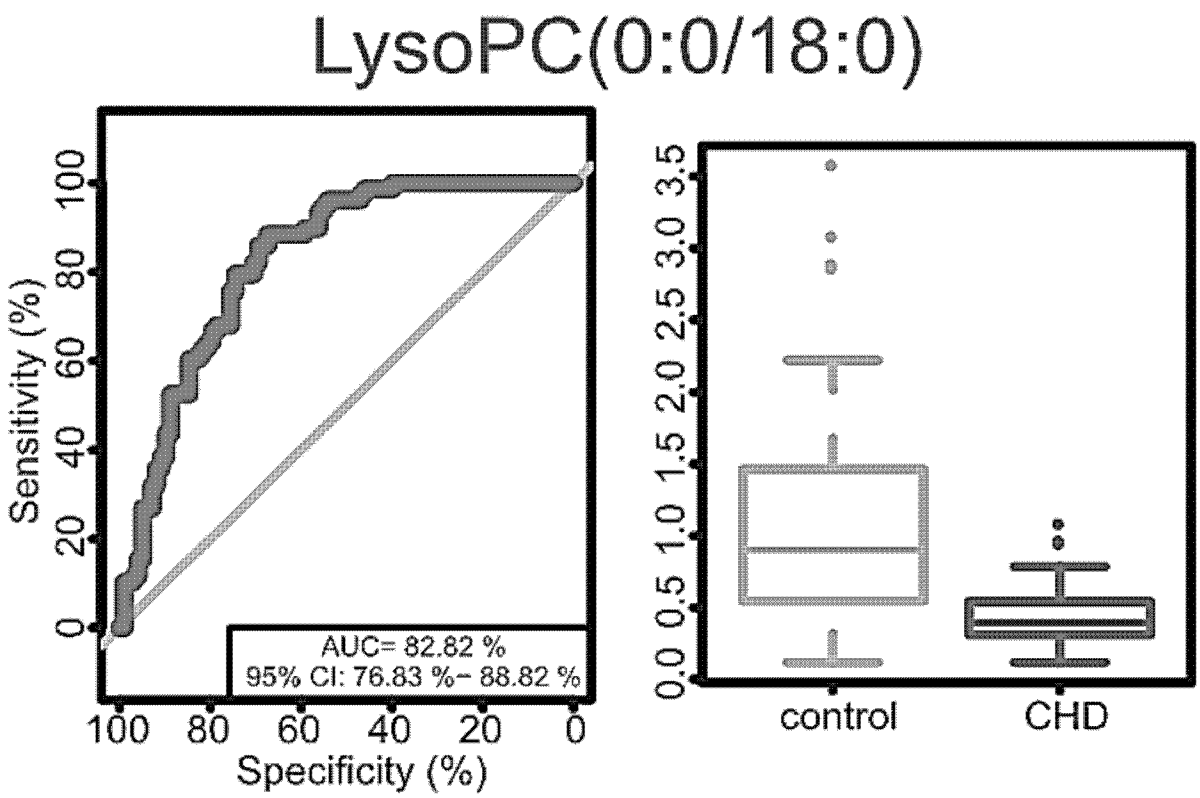
b


LysoPC(14:0)
Sensitivity (%)
100
80
60
40
20
0
AUC= 84.39 %
95% CI: 78.79 %- 90 %
100 80 60 40 20 0
Specificity (%)
3.0
2.5
2.0
1.5
1.0
0.5
control
CHD

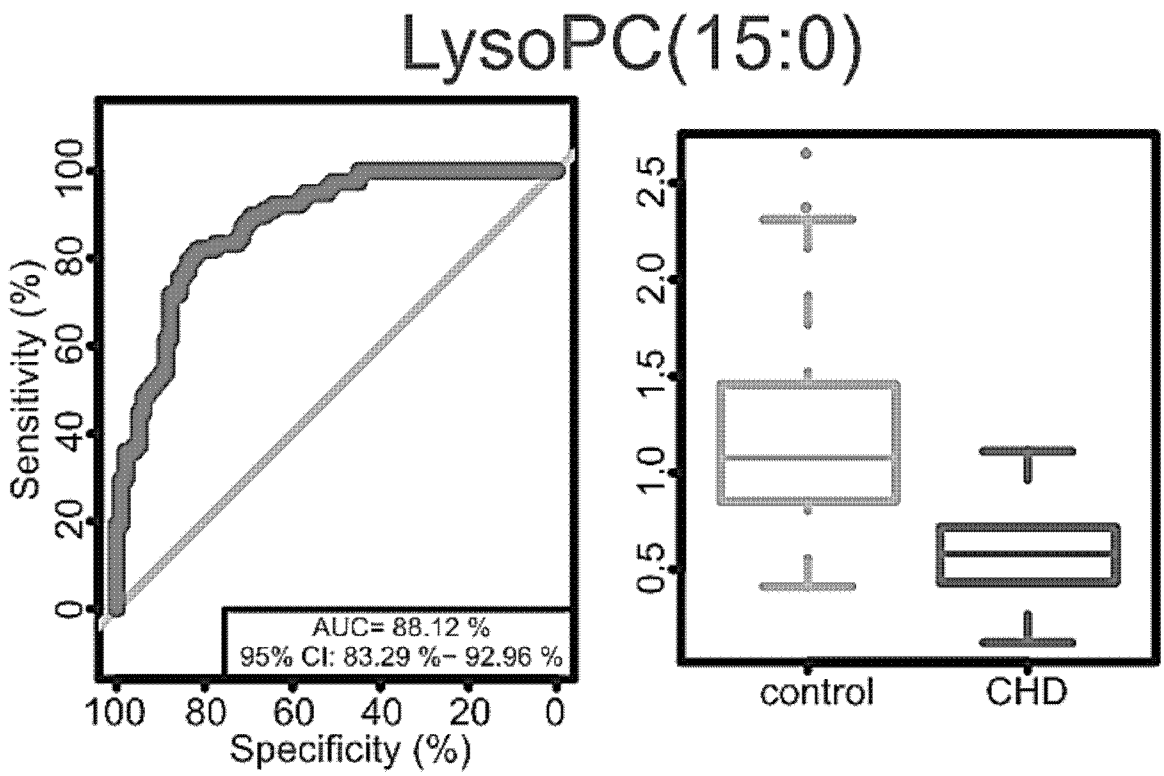
c


LysoPC(16:1(9Z))
Sensitivity (%)
100
80
60
40
20
0
AUC= 83.41 %
95% CI: 77.66 %- 89.16 %
100 80 60 40 20 0
Specificity (%)
2.0
1.5
1.0
0.5
control
CHD

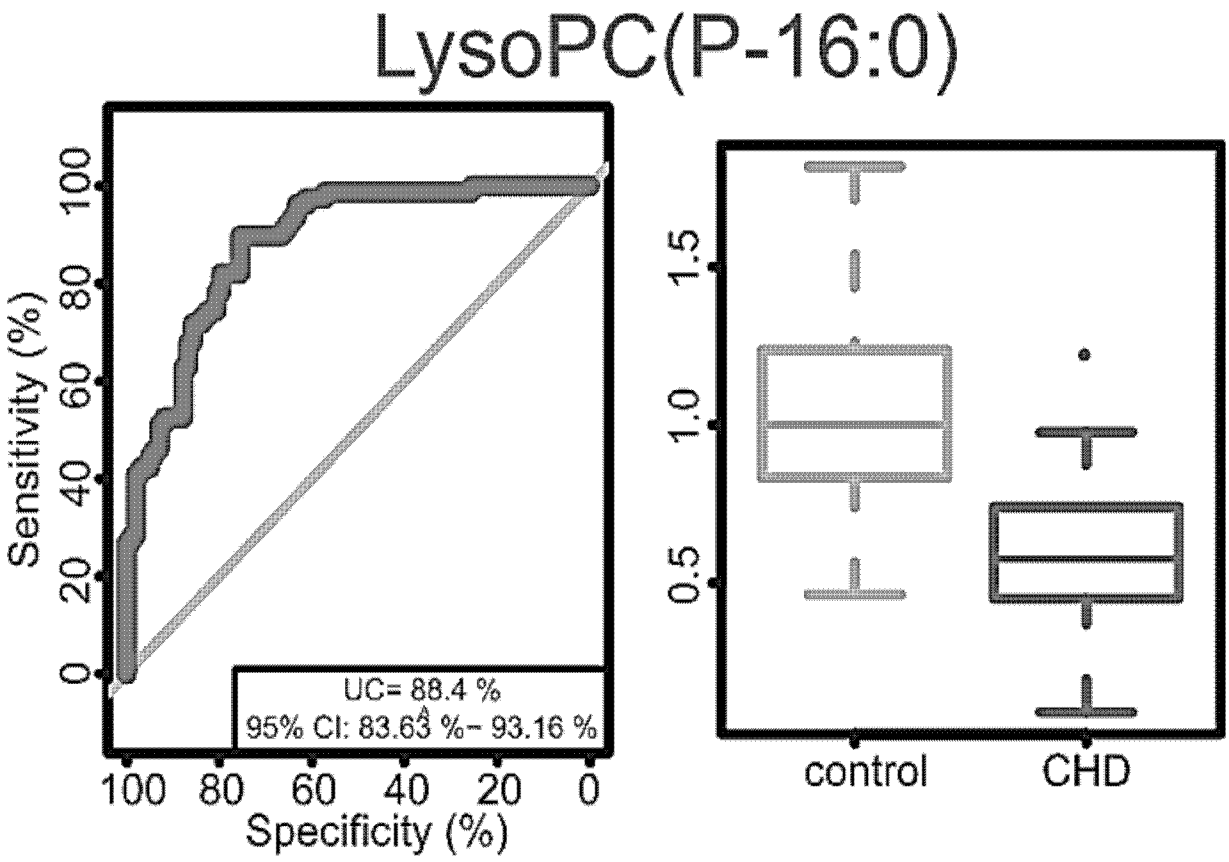
d


LysoPC(0:0/18:0)
Sensitivity (%)
0
20
40
60
80
100
AUC= 82.82 %
95% CI: 76.83 %- 88.82 %
100 80 60 40 20 0
Specificity (%)
0.0 0.5 1.0 1.5 2.0 2.5 3.0 3.5
control
CHD

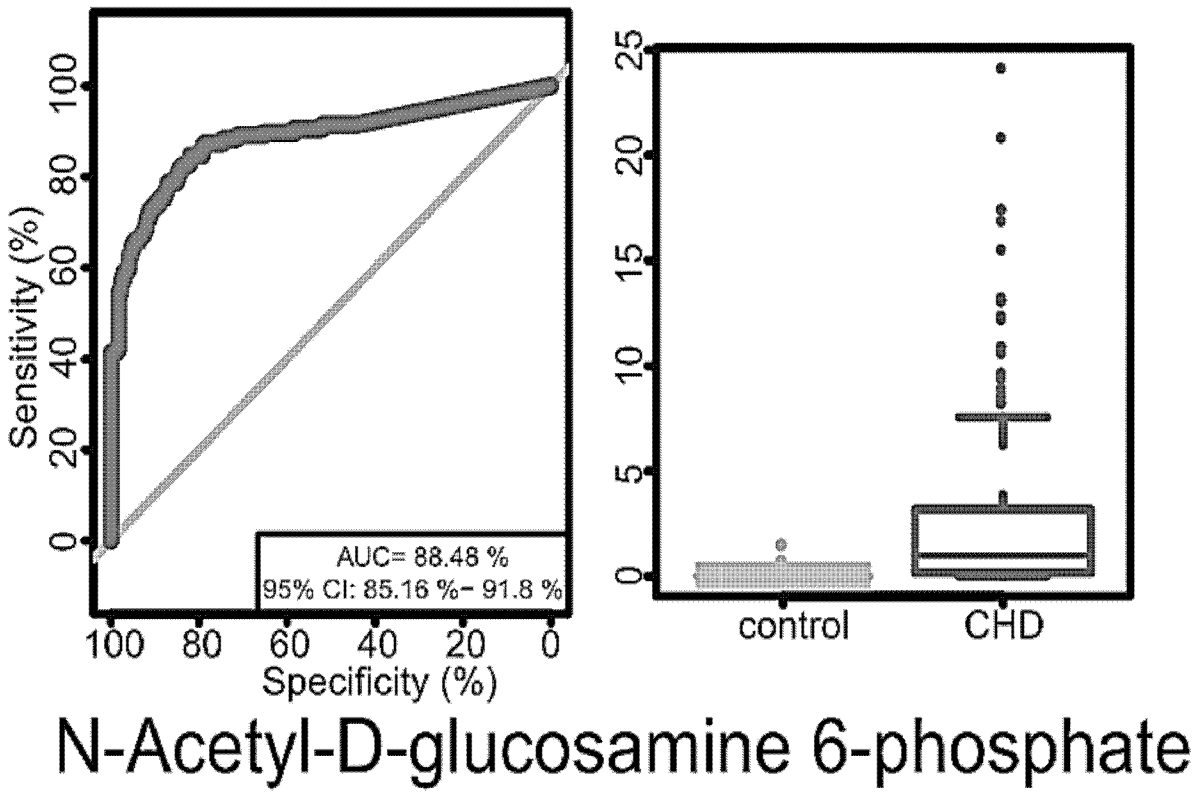
e


LysoPC(15:0)
Sensitivity (%)
0
20
40
60
80
100
AUC= 88.12 %
95% CI: 83.29 %- 92.96 %
100 80 60 40 20 0
Specificity (%)
0.5 1.0 1.5 2.0 2.5
control
CHD

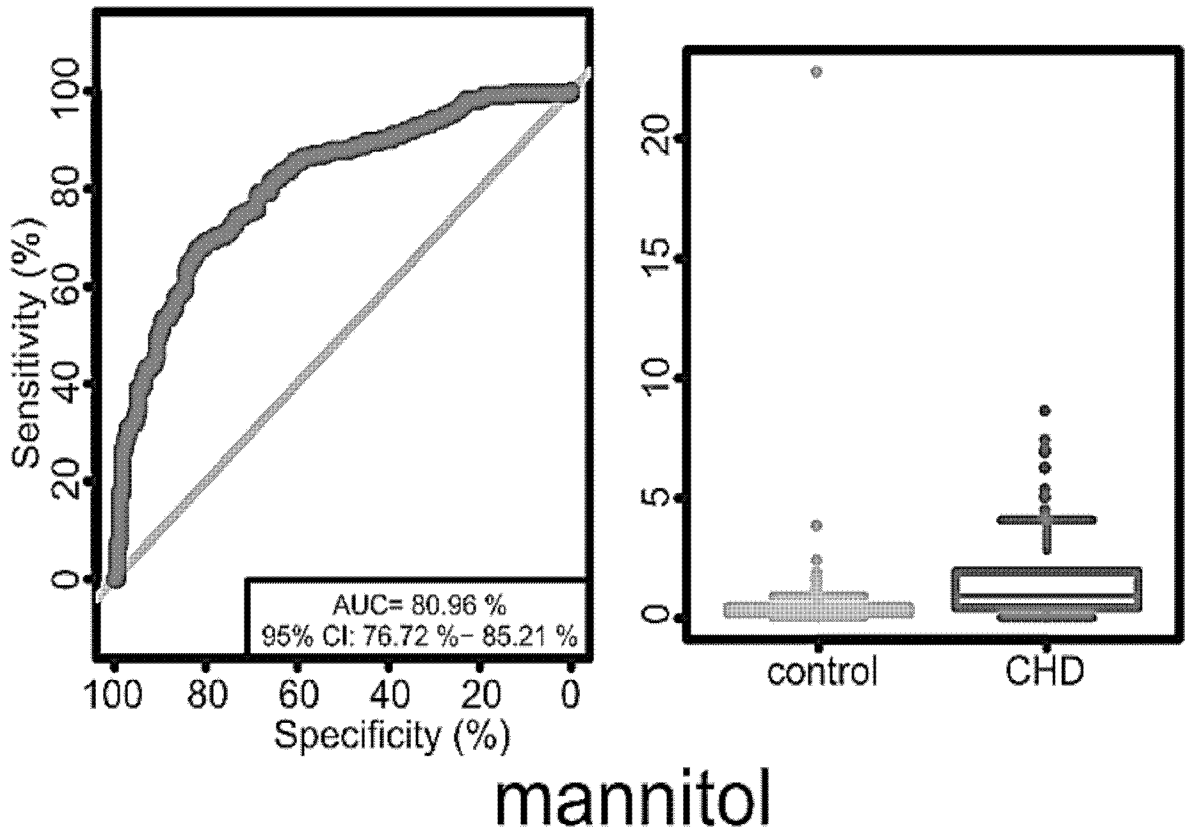
f


LysoPC(P-16:0)
Sensitivity (%)
100
80
60
40
20
0
UC= 88.4 %
95% CI: 83.63 %- 93.16 %
100 80 60 40 20 0
Specificity (%)
1.5
1.0
0.5
control
CHD

g


Sensitivity (%)
100
80
60
40
20
0
AUC= 88.48 %
95% CI: 85.16 %- 91.8 %
100 80 60 40 20 0
Specificity (%)
25
20
15
10
5
0
control
CHD
N-Acetyl-D-glucosamine 6-phosphate

h


Sensitivity (%)
0
20
40
60
80
100
AUC= 80.96 %
95% CI: 76.72 %- 85.21 %
100
80
60
40
20
0
Specificity (%)
0
5
10
15
20
control
CHD
mannitol

i

Fig.9

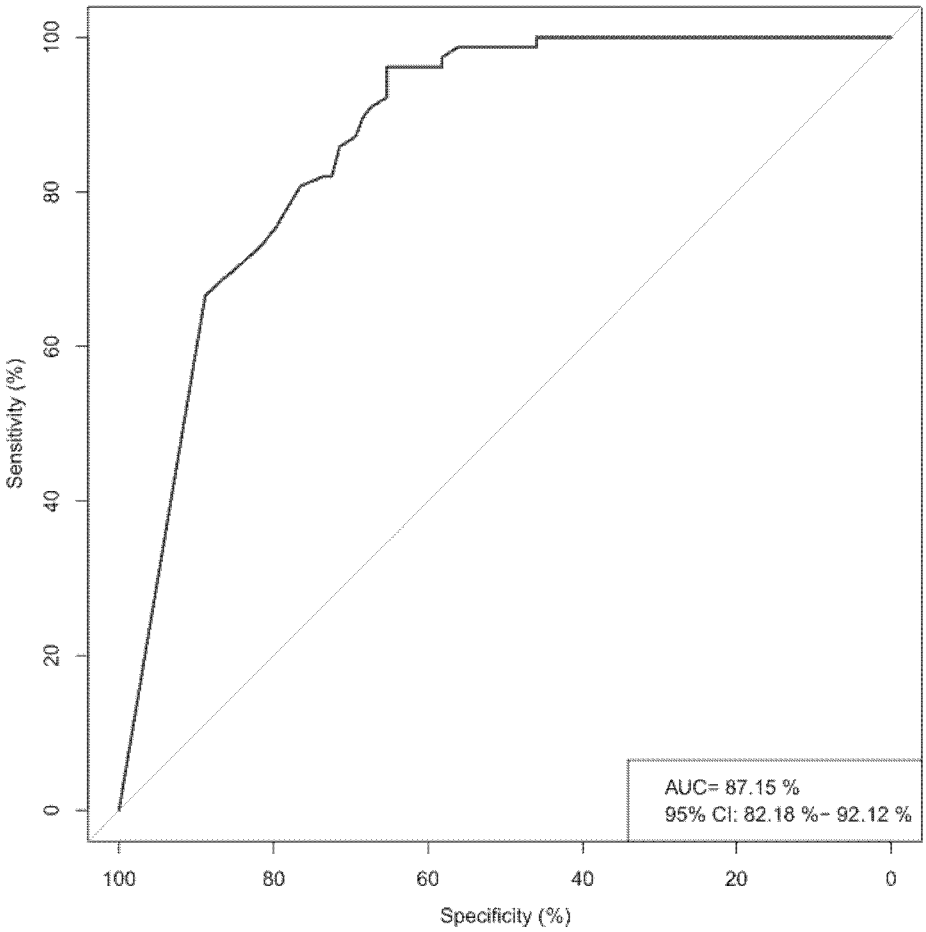

Sensitivity (%)
100
80
60
40
20
0
Specificity (%)
100
80
60
40
20
0
AUC= 87.15 %
95% CI: 82.18 %- 92.12 %

Fig.10

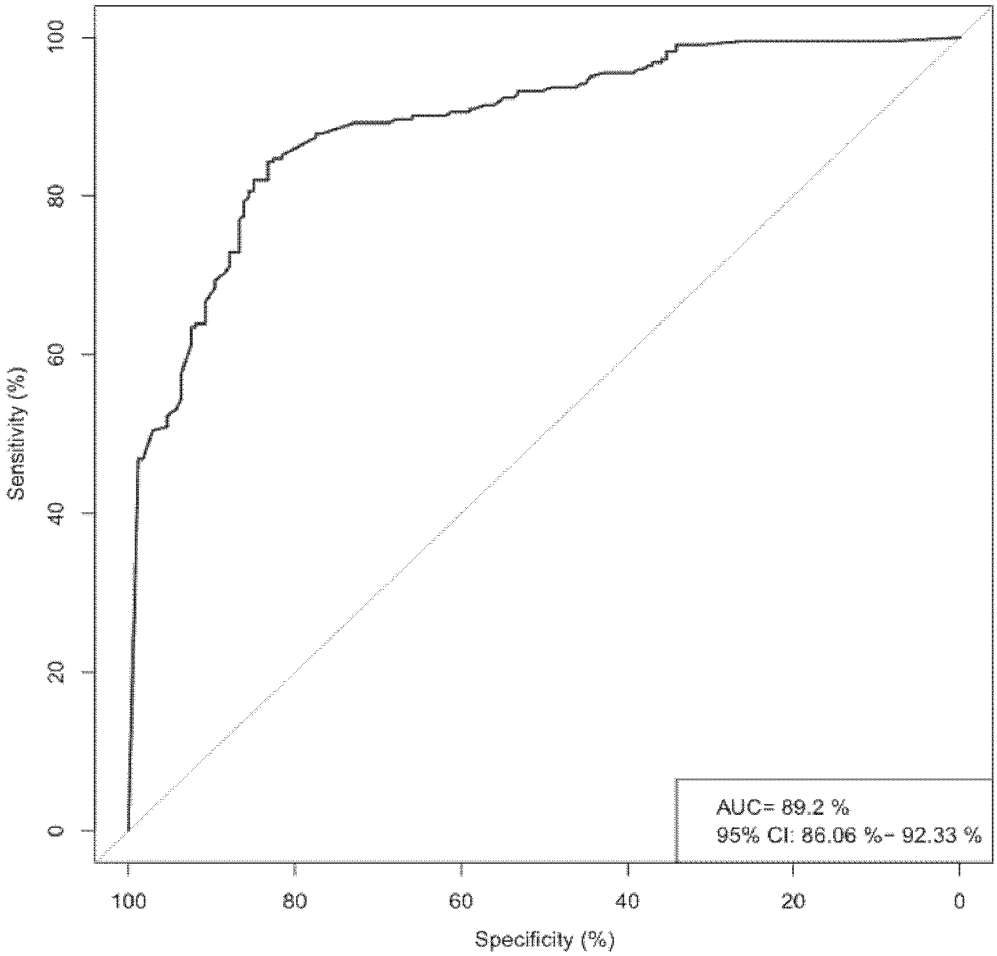
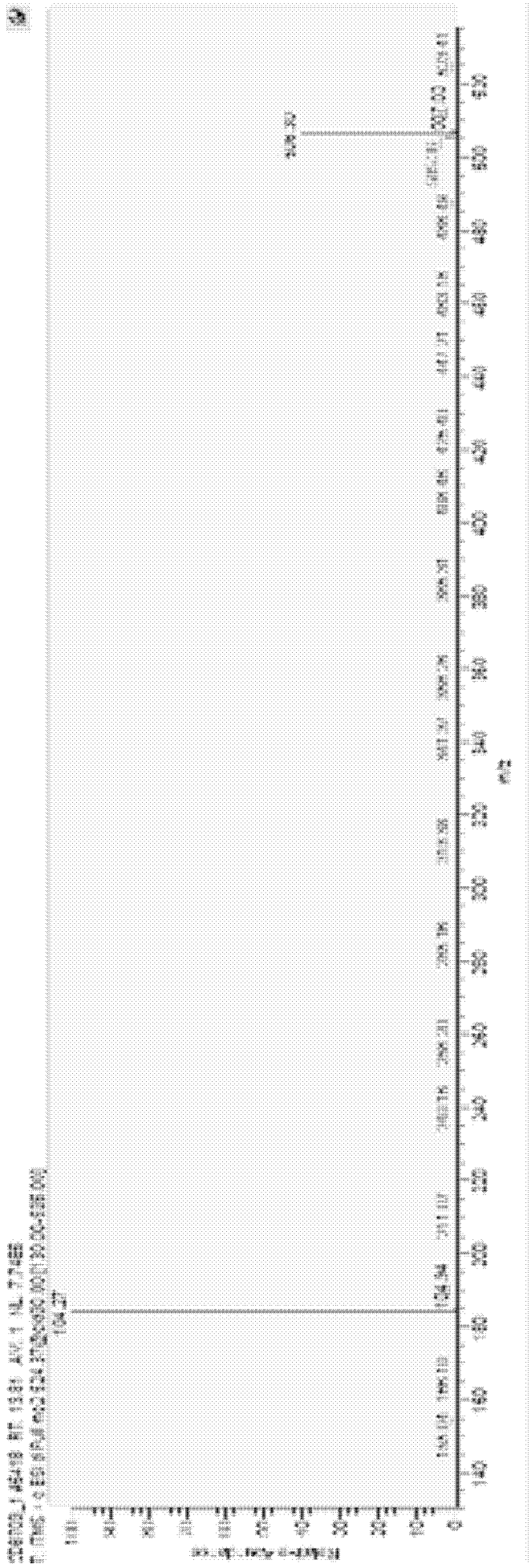
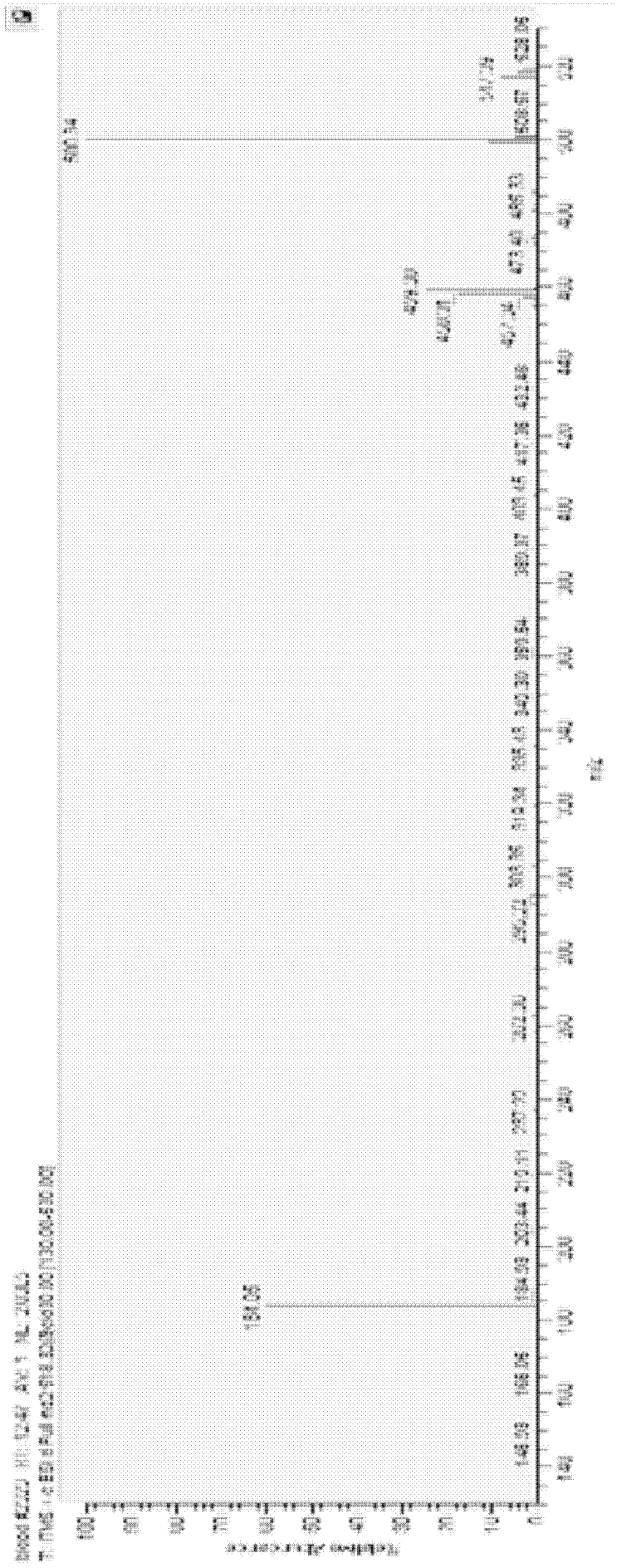

Sensitivity (%)
100
80
60
40
20
0
Specificity (%)
100
80
60
40
20
0
AUC= 89.2 %
95% CI: 86.06 %- 92.33 %

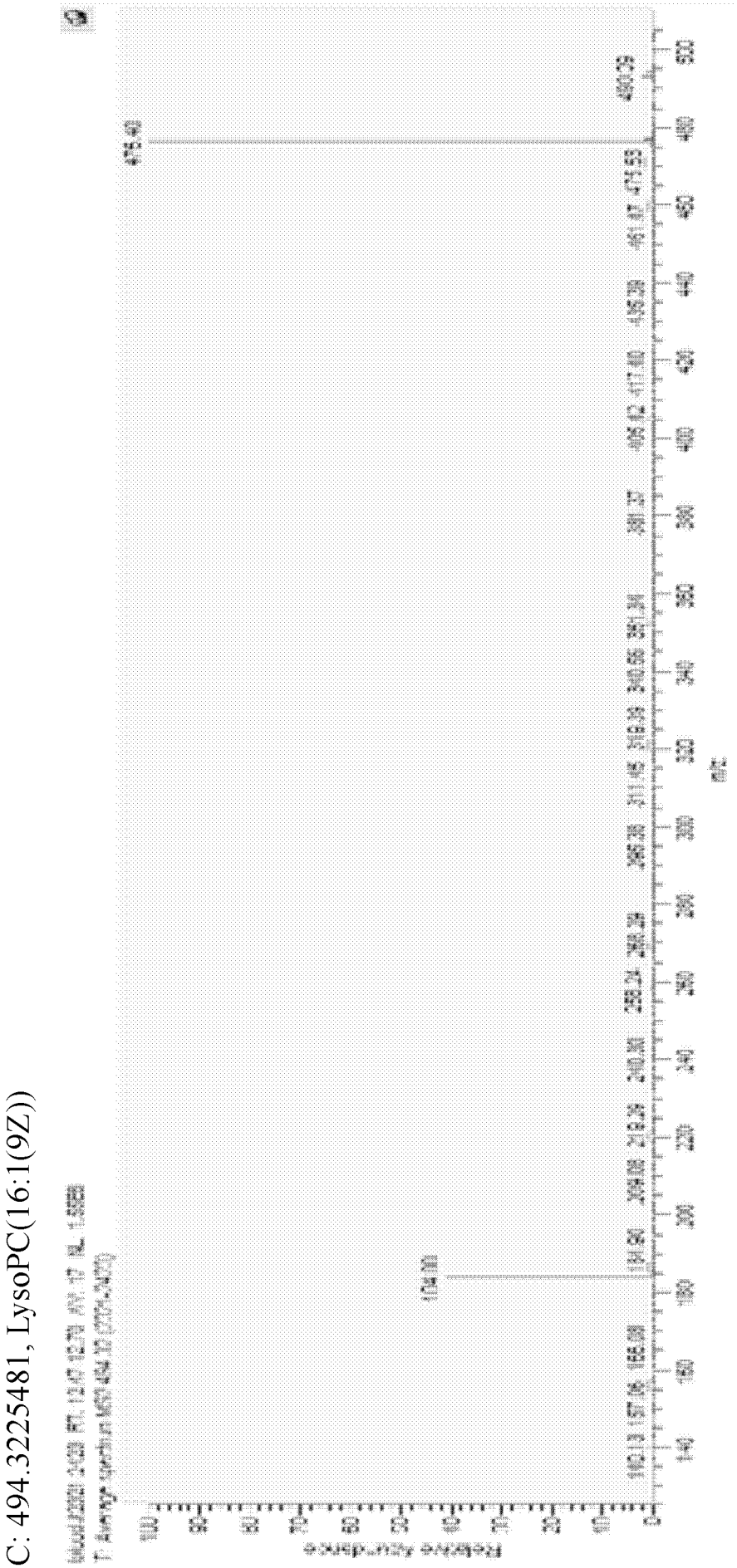
Fig.11



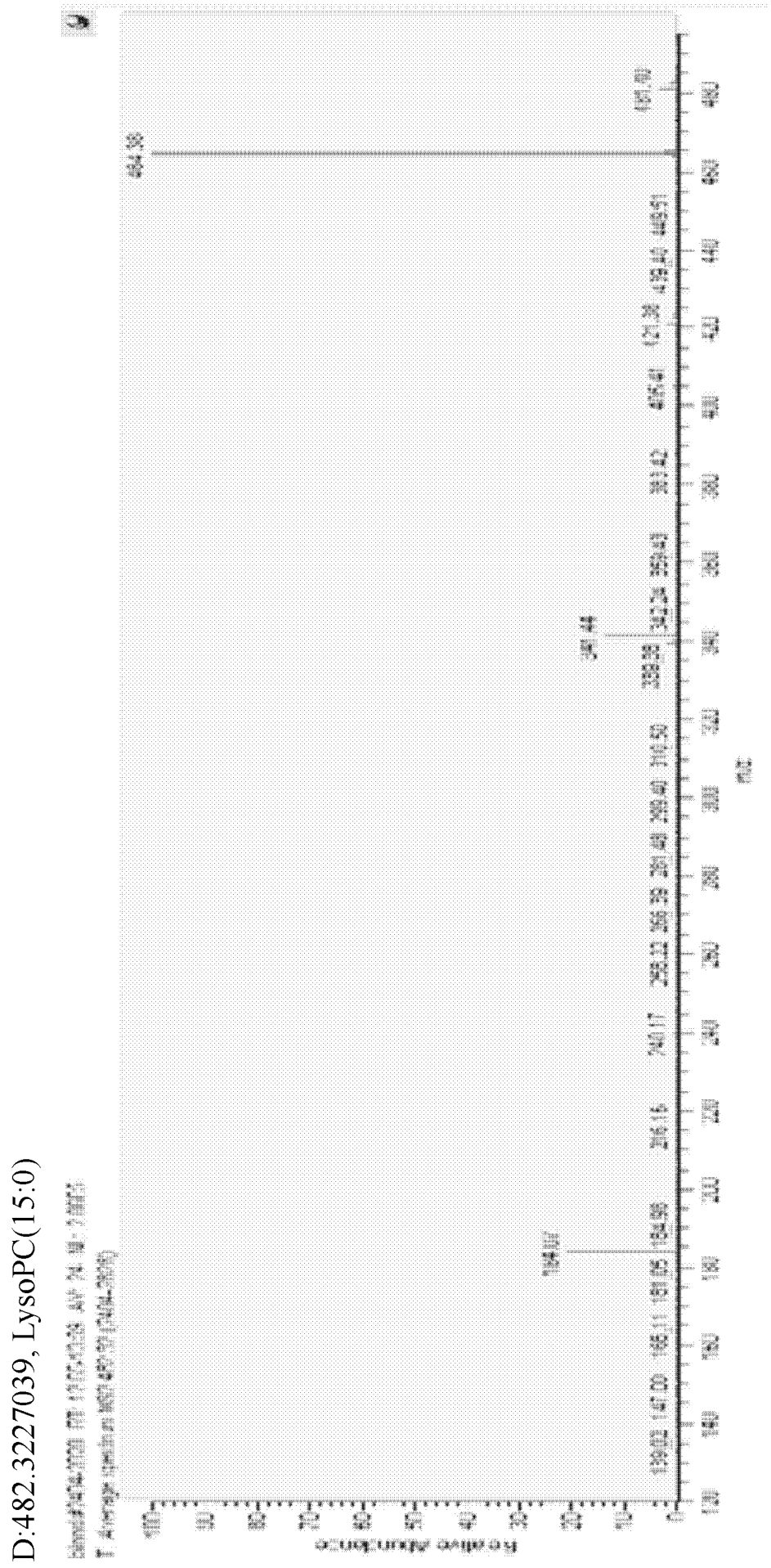
A: 524.3695234, LysoPC(0:0/18:0)

B: 518.3222242, LysoPC(18:3(6Z,9Z,12Z))

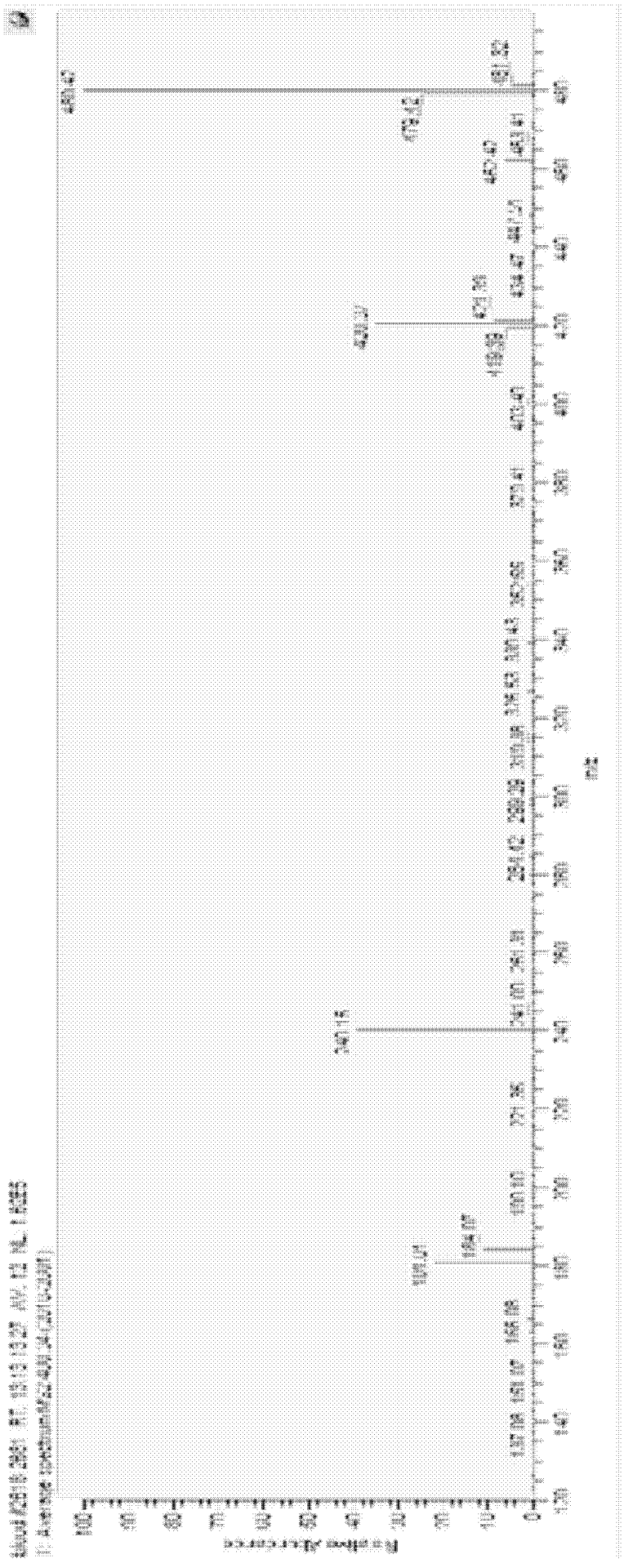




C: 494.3225481, LysoPC(16:1(9Z))


D:482.3227039, LysoPC(15:0)
m/z

E: 480.343497, LysoPC(P-16:0)



F: 468.3070804, LysoPC(14:0)

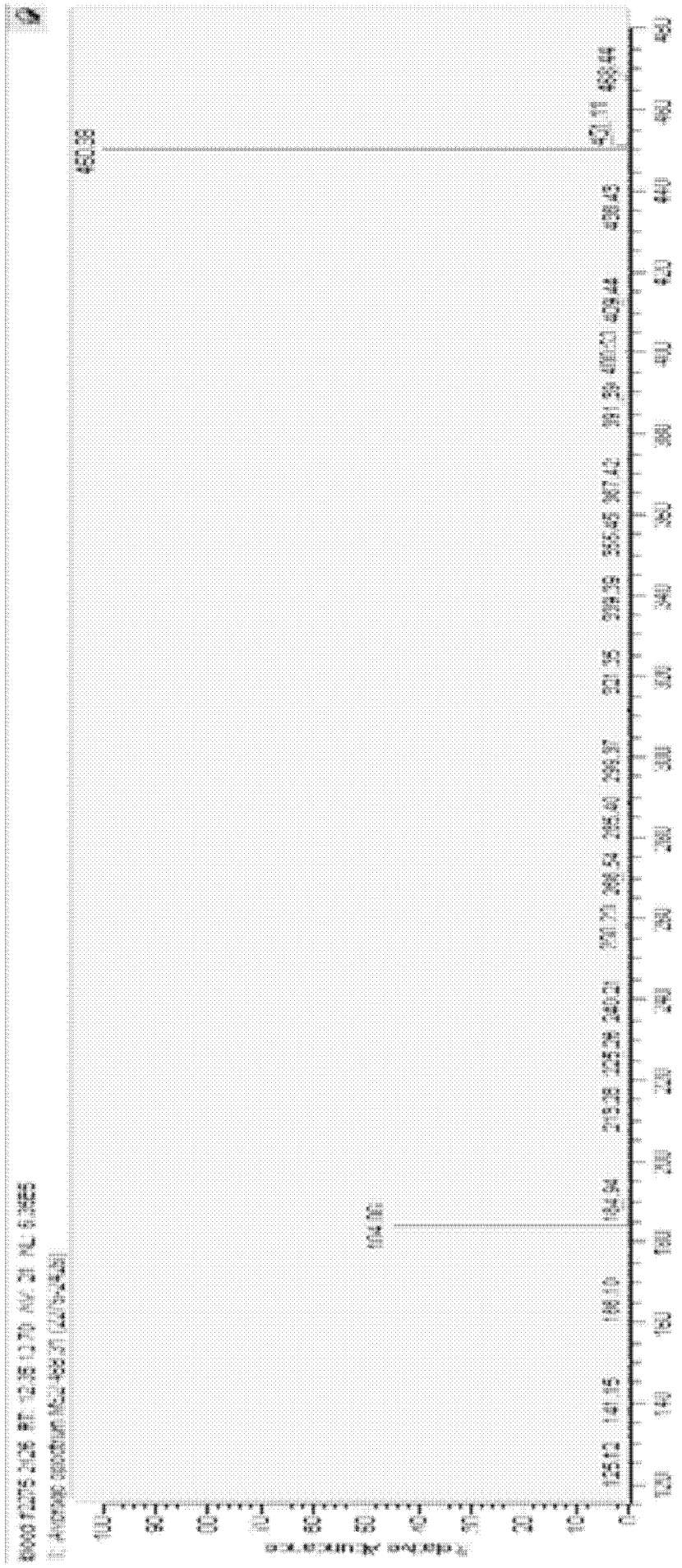


Fig. 12

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2015/087664

A. CLASSIFICATION OF SUBJECT MATTER

G01N 30/88(2006.01)i; G01N 30/00(2006.01)i

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G01N

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

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

WPI, EPODOC, CNPAT, CNTXT, EPTXT, USTXT, WOTXT, CNKI, WANFANG, EBI, NCBI, GOOGLE, BAIDU, ISI Web of Knowledge, PUBMED:BGI SHENZHEN, FENG Qiang, LIU Zhipeng, WANG Jun, CORONARY HEART DISEASE, BIOMARKERS, coronary artery heart disease, CHD, coronary atherosclerosis, ischemic heart, LysoPC, LPC, lysophosphatidylcholine, lysophospholipids, Palmitoylglycerophosphocholine, 1-Palmitoylglycerophosphocholine, 1-Oleoylglycerophosphocholine, Oleoylglycerophosphocholine, Paraxanthine, L-Arginine, N-Acetyl-D-glucosamine 6-phosphate, gas chromatography mass spectrometry, GC-MS, liquid chromatography mass spectrometry, LC-MS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	GANNA, Andrea et al. "Large-scale Metabolomic Profiling Identifies Novel Biomarkers for Incident Coronary Heart Disease" *PLOS Genetics*, Vol. 10, No. 12, 11 December 2014 (2014-12-11), e1004801	1-53(partially)
X	WO 2014043421 A1 (BERG LLC.) 20 March 2014 (2014-03-20) claims 1-11	1-53(partially)
X	WO 2011063470 A1 (BAKER IDI HEART AND DIABETES INSTITUTE HOLDINGS LTD.) 03 June 2011 (2011-06-03) claims 1-22, description, paragraph 【0012】	1-53(partially)
X	WO 2009049189 A2 (BG MEDICINE, INC.) 16 April 2009 (2009-04-16) claims 1-29, description, paragraph 【0011】, the embodiment 1	1-53(partially)
X	WO 2011161062 A2 (ZORA BIOSCIENCES OY) 29 December 2011 (2011-12-29) claims 1-26	1-53(partially)

☑ Further documents are listed in the continuation of Box C. ☑ See patent family annex.

* Special categories of cited documents:

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

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

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

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

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

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
07 May 2016	**12 May 2016**
Name and mailing address of the ISA/CN **STATE INTELLECTUAL PROPERTY OFFICE OF THE P.R.CHINA** **6, Xitucheng Rd., Jimen Bridge, Haidian District, Beijing 100088, China** Facsimile No. **(86-10)62019451**	Authorized officer **LI,Juanjuan** Telephone No. **(86-10)82246978**

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2015/087664

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2011138419 A1 (ZORA BIOSCIENCES OY) 10 November 2011 (2011-11-10) claims 1-25	1-53(partially)
A	WO 0023612 A1 (ATAIRGIN BIOTECHNOLOGIES, INC.) 27 April 2000 (2000-04-27) the whole document	1-53(partially)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2015/087664

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] 1.Claims1-53(partially): direct to biomarker 1 ;

[2] 2. Claims1-53(partially): direct to biomarker 2 ;

[3] 3. Claims1-53(partially): direct to biomarker 3 ;

[4] ……

[5] 18. Claims1-53(partially): direct to biomarker 18.

[6] The same or corresponding technical features among the inventions above are as follows: a biomarker for predicting the risk of CHD. However, the same or corresponding technical features above are well known in the art.It follows that the same or corresponding technical features of claims above do not make a contribution over the prior art and can not be considered as special technical features within the meaning of Rule 13.2 PCT. The application, hence does not meet the requirement of unity of invention as defined in Rule 13.1 PCT.

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.: **claims1-53(partially) directing to biomarker 1**

Remark on Protest

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

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

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2015/087664

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
WO	2014043421	A1	20 March 2014	IL	237557	D0	30 April 2015
				KR	20150043566	A	22 April 2015
				SG	11201501684U	A	29 April 2015
				AU	2013315371	A1	09 April 2015
				IL	237557	A	30 April 2015
				CN	104781670	A	15 July 2015
				CA	2884637	A1	20 March 2014
				JP	2015529335	A	05 October 2015
				EP	2895861	A1	22 July 2015
				US	2014100128	A1	10 April 2014
WO	2011063470	A1	03 June 2011	US	2013023054	A1	24 January 2013
				EP	2504708	A4	01 May 2013
				EP	2944964	A3	25 November 2015
				ES	2536712	T3	27 May 2015
				EP	2944964	A2	18 November 2015
				EP	2504708	B1	11 March 2015
				US	9255935	B2	09 February 2016
				AU	2010324544	A1	21 June 2012
				CN	102884435	A	16 January 2013
				EP	2504708	A1	03 October 2012
				US	2015301070	A1	22 October 2015
				US	9110086	B2	18 August 2015
WO	2009049189	A2	16 April 2009	ES	2374762	T3	21 February 2012
				EP	2210108	B1	07 December 2011
				WO	2009049189	A3	18 June 2009
				CN	101939651	A	05 January 2011
				AU	2008310669	A1	16 April 2009
				AT	536553	T	15 December 2011
				JP	2011501133	A	06 January 2011
				IL	204922	D0	30 November 2010
				US	2011045514	A1	24 February 2011
				EP	2210108	A2	28 July 2010
				CA	2702268	A1	16 April 2009
WO	2011161062	A2	29 December 2011	CA	2801459	A1	29 December 2011
				CN	103154742	A	12 June 2013
				EP	2583107	A2	24 April 2013
				JP	2013531238	A	01 August 2013
				WO	2011161062	A3	28 June 2012
				US	9201080	B2	01 December 2015
				US	2013216560	A1	22 August 2013
				US	2016025751	A1	28 January 2016
WO	2011138419	A1	10 November 2011	JP	5871244	B2	01 March 2016
				CA	2798238	A1	10 November 2011
				US	2015247873	A1	03 September 2015
				DK	2385374	T3	17 March 2014
				EP	2385374	B1	08 January 2014
				JP	2013527449	A	27 June 2013
				ES	2455124	T3	14 April 2014
				CN	102971633	A	13 March 2013
				EP	2385374	A1	09 November 2011

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2015/087664

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
		US 2013045217 A1	21 February 2013
		EP 2567240 A1	13 March 2013
		US 9046538 B2	02 June 2015
WO 0023612 A1	27 April 2000	US 6248553 B1	19 June 2001
		JP 4472872 B2	02 June 2010
		CN 1354799 A	19 June 2002
		JP 2003524762 A	19 August 2003
		US 2002039757 A1	04 April 2002
		IL 142510 D0	10 March 2002
		EP 1123411 A1	16 August 2001
		CA 2347103 A1	27 April 2000
		AU 1321800 A	08 May 2000
		EP 1123411 A4	02 May 2002
		MX PA01004010 A	10 March 2003
		US 6255063 B1	03 July 2001