Published in final edited form as:

Electrophoresis. 2015 February ; 36(4): 556–563. doi:10.1002/elps.201400319.

Proteomic profiling of human sera for discovery of potential biomarkers to monitor abstinence from alcohol abuse

Xianyin Lai^{1,5,*}, Suthat Liangpunsakul^{2,3}, Kaigang Li⁴, and Frank A. Witzmann⁵

¹Department of Biochemistry and Molecular Biology, Indiana University School of Medicine, Indianapolis, IN 46202, USA

²Department of Medicine – Division of Gastroenterology/Hepatology, Indiana University School of Medicine, Indianapolis, IN 46202, USA

³Roudebush Veterans Administration Medical Center, Indianapolis, IN 46202, USA

⁴Department of Applied Health Science, Indiana University School of Public Health, Bloomington, IN 47405, USA

⁵Department of Cellular & Integrative Physiology, Indiana University School of Medicine, Indianapolis, IN 46202, USA

Abstract

Although numerous biomarkers or biomarker candidates have been discovered to detect levels of drinking and intervals of time after last drinking episode, only a few biomarkers have been applied to monitor abstinence in a longer interval (6 weeks) from alcohol abuse. Considering sample sources, sensitivity, and specificity, new biomarkers from blood with better accuracy are needed. To address this, serum proteomic profiles were compared between pre- and post- treatment samples from subjects seeking treatment for alcohol abuse and dependence in an intensive 6-week daily outpatient program using high-abundance plasma protein immunodepletion and LC-MS/MS techniques. Protein identification, quantification, candidate biomarker selection, and prioritization analyses were carried out. Among the 246 quantified serum proteins, abundance of 13 and 45 proteins in female and male subjects were significantly changed ($p < 0.05$), respectively. Of these biomarker candidate proteins, 2 (female) and 8 (male) proteins were listed in category 1, with high area under the receiver operating characteristic (ROC) curve (AUC), sensitivity, specificity, and fold change. In summary, several new biomarker candidates have been identified to monitor abstinence from alcohol abuse.

Keywords

Alcohol Abuse; Biomarker; Proteomics; Tandem Mass Spectrometry

*Correspondence: Department of Biochemistry and Molecular Biology, Department of Cellular and Integrative Physiology, Indiana University School of Medicine, 635 Barnhill Drive, Room 0044, Indianapolis, IN 46202 USA xlai@iu.edu, Phone: 317-278-0084, Fax: 317-274-4686 .

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

1 Introduction

The abuse of alcohol is a major public health and social problem. Alcohol is causally related to more than 60 different medical conditions and 4% of the global burden of disease, which accounts for about as much death and disability globally as tobacco and hypertension [1]. Tests that can properly assess the extent of the subject's past drinking activity would be valuable to process a specific clinical setting such as in pre-liver transplant evaluation [2].

Numerous biomarkers or biomarker candidates have been discovered to detect levels of drinking and intervals of time after the last drinking episode. Breath, blood, and urine alcohol can be detected at low level of drinking (< 1 drink/day) within hours after the last drinking. Urine ethyl sulfate (EtS) and ethyl glucuronide (EtG) are detectable at moderate levels (between 1 and 14 drinks/day) within 5 days after the last alcohol consumption. Phosphatidylethanol (PEth) and % carbohydrate-deficient transferrin (CDT) works at moderate level of drinking within 4 weeks. Gamma-glutamyltransferase (GGT) can be detected at risky level of drinking (≥ 14 drinks/day) within 4 weeks. To detect alcohol abuse that occurs after intervals longer than 6 weeks of abstinence, only a few biomarkers exist, such as mean corpuscular volume (MCV), fatty acid ethyl esters (FAEE) in meconium, and FAEE, EtG, and PEth in hair [3]. Considering sample sources, sensitivity, and specificity, we need new biomarkers from blood that enable better accuracy.

Proteomic approaches have been used previously to search for biomarkers of alcohol consumption [4]. Two-dimensional differential gel electrophoresis (2D-DIGE) and surface enhanced laser desorption/ionization-time of flight-mass spectrometry (SELDI-TOF-MS) were used to study serum samples from subjects with excessive alcohol use [5, 6]. However, 2D-DIGE suffers from several limitations, such as poor solubilization of hydrophobic proteins and basic proteins in buffers, and poor resolution of very high and low-molecular weight proteins [7]. SELDI-TOF-MS has been limited by the low sensitivity for higher molecular weight proteins [8]. Both techniques are associated with difficulties in the identification of potential biomarkers. In contrast, liquid chromatography - tandem mass spectrometry (LC-MS/MS) can be used for both identification and quantification of proteins in complex sample mixtures. With new techniques to deplete high-abundance plasma proteins, LC-MS/MS can be used to comprehensively investigate the low-abundance proteome for potential biomarkers [9].

In the present study, we applied an LC-MS/MS approach with a high-abundance plasma protein immunodepletion technique to profile the changes of serum proteins to discover potential biomarker candidates to monitor abstinence from alcohol abuse.

2 Materials and methods

2.1 Subjects

The study was approved by the Indiana University Purdue University at Indianapolis institutional review board and by the Fairbanks Addiction Treatment Center. All subjects were recruited from Fairbanks, a private, not-for profit hospital in Indianapolis, Indiana, specializing in treating individuals with alcohol abuse and drug addiction. Subjects were

those seeking treatment for alcohol abuse and dependence. Upon their first visit, subjects were seen by addiction psychiatrists, at which time a full detailed interview, history, and physical examination were obtained. Inclusion criteria screened for subjects who met the DSM-IV Diagnostic Criteria for Alcohol Dependence/abuse and who had been actively drinking up to within 3 days of the first visit. Subjects were excluded if they had a history of any localized or systemic infectious disease within 4 weeks prior to the study and had symptoms and signs of decompensated liver disease (jaundice, ascites, or hepatic encephalopathy). Eligible subjects who met all the enrollment criteria were asked to answer questionnaires which included the following information: a) background information and demographics, b) alcohol consumption, c) alcohol abuse and dependence, d) previous history of alcohol treatment utilization, e) family history of alcoholism, f) tobacco use and dependence, g) medicine use, h) past medical history, i) drug abuse and dependence, and j) family history of drug abuse. Seven mL of blood from each subject were obtained according to the protocol described in the proteomic analysis section. This set of samples was referred to as the “pre-treatment samples”. The alcohol treatment program at Fairbanks was an intensive 6-week daily outpatient treatment program. During this 6-week period, all subjects were followed closely by an addiction specialist and none relapsed. At the end of the 6-week treatment program, another set of blood samples (“post-treatment samples”) were drawn and serum proteomic profiles were compared between pre- and post- treatment samples. During the study period, 75 subjects were recruited, and 16 out of 75 subjects met the established inclusion criteria and provided both pre- and post-treatment samples.

2.2 Materials

Urea, DL-Dithiothreitol (DTT), triethylphosphine (TEP), iodoethanol, and ammonium bicarbonate (NH_4HCO_3) were purchased from Sigma-Aldrich (St. Louis, MO, USA). Acetonitrile (ACN) and MS grade water were purchased from EMD Chemicals (Gibbstown, NJ, USA). Modified sequencing grade porcine trypsin was obtained from Princeton Separations (Freehold, NJ, USA). ProteoPrep® 20 plasma immunodepletion kits were purchased from Sigma-Aldrich.

2.3 Samples

Serum samples were obtained from 16 subjects who met the established inclusion criteria and from whom post-treatment samples were also obtained. Seven mL of blood were obtained for clinical values assessment. Another 7 mL of blood were drawn and centrifuged at $1,500 \times g$ for 10 min at room temperature, within 1-3 hour(s) after acquisition. Aliquots (0.5 mL) were placed in separate 2 mL cryovials and stored at -80°C until analysis. Both pre- and post-treatment groups each contained 16 samples.

2.4 Depletion of high-abundance serum proteins

High-abundance proteins were depleted from each serum sample aliquot as described in the ProteoPrep® 20 User Guide. Briefly, 100 μL of diluted sample (400 μg) was placed on the equilibrated spin column and incubated at room temperature for 20 min. The spin column and collection tube were then centrifuged at $2,000 \times g$ for 30 s at room temperature. The flow-through volume was saved in the collection tube. The spin column was then washed

twice with 100 μ L of equilibration buffer. Total fluid collection was 300 μ L. The proteins bound to the column were eluted with 2 mL of elution solution. Each sample was processed as nine technical replicates. The nine replicates were pooled and concentrated with the filter, and its final volume was 300 μ L. A final depletion was carried out on the concentrated depleted serum. These protein samples were subsequently referred to as “depleted samples”. Protein concentration was determined by the Bradford assay [10]. A 100 μ g aliquot of each sample was placed into a new tube and dried via SpeedVac.

2.5 Protein reduction, alkylation, and digestion for LC-MS/MS

The depleted serum samples were reconstituted in 200 μ L of 4 M urea and then reduced and alkylated using TEP and iodoethanol as described by Lai et al. [9]. Briefly, 200 μ L of the reduction/alkylation cocktail was added to the protein solution. The sample was incubated at 37°C for 120 min, dried by SpeedVac, and reconstituted with 100 μ L of 100 mM NH_4HCO_3 at pH 8.0. A 150 μ L aliquot of a 20 μ g/mL trypsin solution was added to the sample and incubated at 37°C for 3 h, after which another 150 μ L of trypsin was added, and the solution incubated at 37°C overnight.

2.6 LC-MS/MS

The digested samples were analyzed using a Thermo-Finnigan linear ion-trap (LTQ) mass spectrometer coupled with a Surveyor autosampler and MS HPLC system (Thermo-Finnigan). Tryptic peptides were injected onto the C18 microbore RP column (Zorbax SB-C18, 1.0 mm x 150 mm) at a flow rate of 50 μ L/min. The mobile phases A, B, and C were 0.1% formic acid in water, 50% ACN with 0.1% formic acid in water, and 80% ACN with 0.1% formic acid in water, respectively. The gradient elution profile was as follows: 10% B (90% A) for 5 min, 10-95% B (90-5% A) for 120 min, 100% C for 5 min, and 10% B (90% A) for 12 min. The data were collected in the “Triple-Play” (MS scan, Zoom scan, and MS/MS scan) mode with the ESI interface using normalized collision energy of 35%. Dynamic exclusion settings were set to repeat count 1, repeat duration 30 s, exclusion duration 120 s, and exclusion mass width 0.75 m/z (low) and 2.0 m/z (high). Each sample was injected twice.

2.7 Protein identification and quantification

The acquired data were searched against the UniProt protein sequence database of HUMAN (released on 10/31/2013) using SEQUEST (v. 28 rev. 12) algorithms in Bioworks (v. 3.3). General parameters were set as follows: peptide tolerance 2.0 amu, fragment ion tolerance 1.0 amu, enzyme limits set as “fully enzymatic - cleaves at both ends”, and missed cleavage sites set at 2. The searched peptides and proteins were validated by PeptideProphet [11] and ProteinProphet [12] in the Trans-Proteomic Pipeline (TPP, v. 3.3.0) (<http://tools.proteomecenter.org/software.php>). Only proteins and peptides with protein probability 0.9000 and peptide probability 0.8000 were reported. Protein quantification was performed using a label-free quantification software package, IdentiQuantXLTM. [13] If the mean of the pre-treatment was greater than the mean of the post-treatment, the fold change was positive and calculated using the formula pre-mean/post-mean, indicating a protein was up-regulated because of the drinking. If the mean of the pre-treatment was less than the mean of the post-treatment, the fold change was negative and calculated using the formula

post-mean/pre-mean, indicating a protein was down-regulated because of the drinking. Since only two groups were compared, the p value was calculated using Student's t-test. The area under the receiver operating characteristic curve (AUC), sensitivity, and specificity analyses for each biomarker candidate were performed using XLSTAT (Addinsoft, New York, NY. <http://www.xlstat.com/en/>).

3 Results and discussion

3.1 Subject data

Pre-treatment clinical data for the 16 subjects who met inclusion criteria (8 males and 8 females) are shown in Supporting Information Table 1. Self-reported pre-treatment drinking levels for all 16 pre/post subjects averaged 17 ± 7 drinks/day and 103 ± 37 drinks/week. We define a "drink" as the following equivalents: 1.5 oz liquor (40% EtOH), 5 oz wine, or 12 oz beer. Many subjects' self-reported drinking reflected a combination of these. Drinking onset was reported as 18 ± 3 yrs of age, and 7 ± 7 drinks were consumed on the day of last use, which were reported to be from 0-144 hours prior to pre-treatment sampling.

3.2 Protein identification and quantification

Peptides and proteins were searched against the Human database using SEQUEST and protein identifications were evaluated using PeptideProphet and ProteinProphet in TPP to further validate their identification probabilities. In total, 246 protein groups (unique proteins) with a probability ≥ 0.9000 from 1,816 peptides with a probability ≥ 0.8000 were identified and quantified. The complete list of identified and quantified proteins is available in the Supplementary Tables 2 (Female) and 3 (Male), where proteins identified by identical peptides are placed into a single protein group. While only 13 proteins in the female sera were significantly different ($p \leq 0.05$) after 6 weeks of abstinence, 45 unique proteins were significantly different ($p \leq 0.05$) in sera obtained from male subjects. Excluding protein isoforms and those proteins that we attempted to deplete, 7 unique proteins in female sera and 30 in male sera are listed in Tables 1 and 2, respectively.

Recent evidence suggests that the benefits and risks of similar levels of alcohol use are not equivalent for women and men. Biological (e.g. sex hormones, body water, etc.) and behavioral (e.g. pace of drinking, combining alcohol with other drugs or with food, etc.) differences related to alcohol consumption affect the rate of alcohol absorption, metabolism, and elimination. However, these effects are complicated and many inconsistent results fail to reveal gender differences in alcohol metabolism [14]. These findings imply that some biomarkers are not reliable for use in both males and females, while others are.

Among the 16 subjects, males drank more than females, while females had higher ALT, AST, and MCV levels with non-significant trends. As illustrated in Tables 1 and 2, five proteins, APOC1, APOD, KLKB1, LGALS3BP, and S100A9, were significantly elevated before treatment both in male and female subjects. Proteins CA1 and GP5 were significantly different only in female subjects and 25 other proteins were significantly different only in male subjects.

These results indicate that gender is a critical consideration in the development of biomarkers of alcohol abuse. Though the exact mechanism remains uncertain, the level of sex hormones may play a role. Alcohol consumption alters female hormone levels by affecting the hypothalamic-pituitary-gonadal (HPG) axis, which includes the hypothalamus, pituitary gland, and ovaries [15].

3.3 Prioritization of candidate proteins

To prioritize the candidate proteins in Tables 1 and 2, multiple criteria are applied. The first one is the area under the receiver operating characteristic (ROC) curve (AUC). The AUC provides a measure of the overall performance of a diagnostic test with a value between 0 and 1. A perfect diagnostic performance is a test with an AUC value of 1 and the practical lower limit for the AUC of a diagnostic test is 0.5 [16]. For the current study, a cut-off of 0.750 was applied. Proteins with an AUC less than 0.750 in the candidate list were labeled as category 4, where proteins have the lowest priority in the candidate list for further validation.

The second criterion is fold change. For future validation of candidates discovered by proteomics, different approaches, such as ELISA and multiple reaction monitoring (MRM) [17], are usually applied. Because these analytical approaches have different sensitivity and reproducibility, selection of greater fold change values is more likely to insure successful biomarker detection across different approaches. The cut-off for fold change selected in the current study is 1.4; thus proteins with a fold-change less than 1.4 were placed into category 3, where proteins have a lower priority for further validation.

A third concern is the number of peptides used for protein quantification in the proteomic approach. The more peptides that are used, the more confident one can be in the quantification result. A protein quantified by multiple peptides has higher priority than a protein quantified by a single peptide. Based on this criterion, proteins with a single peptide are put into category 2, in which proteins have the low priority for further validation. Other proteins are put into category 1, in which proteins has the highest priority among the entire list.

To prioritize proteins inside each category, protein connectivity in a network is applied. To find potential links between the quantified proteins, QIAGEN's Ingenuity® Pathway Analysis (IPA®, QIAGEN Redwood City, www.qiagen.com/ingenuity) was used. IPA Networks are generated based on protein connectivity with other proteins. The more connected a protein is, the more important it is and the more influence it has. Therefore, the higher priority a protein is assigned, if a protein has more interconnection with other proteins from Tables 1 and 2 in a network. When seven proteins from female subjects were submitted to IPA, two networks were generated, but there is no connectivity between any of the seven proteins. When 30 proteins from male subjects were submitted to IPA, two networks (Figure 1) were generated, indicating direct and indirect connectivity among some of the submitted proteins. Proteins in the two networks were assigned with higher priority than other proteins in the candidate list.

Finally, according to the AUC of ROC, fold change, and number of peptides, proteins in the entire candidate list are put into four categories with different priorities. Inside each category, proteins are prioritized according to their connectivity, AUC of ROC, fold change, and number of peptides. The prioritized proteins are shown in Tables 1 and 2, based on their rank. For the female subjects, 2 proteins are listed in category 1. For the male subjects, 8 proteins are listed in category 1.

3.4 Biomaker candidates

3.4.1 Apolipoproteins—The proteomic results indicate that changes in serum apolipoprotein levels are related to abusive alcohol consumption. Apolipoproteins have important functions in health, serving as plasma lipid transfer carriers, enzyme co-factors, and receptor ligands by forming lipoproteins [18]. In the UniprotKB database, human apolipoproteins are sorted into 11 classes and multiple subclasses: Apolipoprotein As (APOA1, APOA2, APOA4, and APOA5), B (APOB), Cs (APOC1, APOC2, APC2, APOC3, and APOC4), D (APOD), E (APOE), F (APOF), H (APOH), J (APOJ), Ls (APOL1, APOL2, APOL3, APOL4, APOL5, and APOL6), M (APOM), and O (APOO). Apolipoprotein A-I (APOA1), A-II (APOA2), and B (APOB) are high-abundance proteins in serum and were depleted by the ProteoPrep® 20 Plasma Immunodepletion column. Therefore, they are not considered as biomarker candidates. In the female subjects, APOC1 and APOD are the only proteins in the category 1 of the prioritization list. In the male subjects, three out of eight proteins are apolipoproteins (APOC2, APOD, and APOM) in category 1 of the prioritization list. The ROC curve, sensitivity and specificity, and box plot of APOC1 and APOD in female subjects are illustrated in Figures 2 and 3, respectively. The ROC curve, sensitivity and specificity, and box plot of APOC2, APOD, and APOM in male subjects are illustrated in Supplementary Figure 1. The cut-off and its corresponding sensitivity and specificity are included in Table 3. The results show that all of them are good biomarker candidates.

3.4.2. Other proteins in category 1—In the male subjects, there are 5 other proteins (KLKB1, GPX3 CD14, VTN, and SERPINA5) listed in category 1. The ROC curve, sensitivity and specificity, and box plots of the 5 proteins are illustrated in Supplementary Figure 1. The minimum and maximum of AUC of the ROC curve are 0.797 and 0.875. Their cut-off and the corresponding sensitivity and specificity are included in Table 3. The ranges of sensitivity and specificity are 0.625 - 0.875 and 0.750 - 1.000. The results indicate that the 5 proteins are good candidates for further biomarker validation.

3.4.3 Proteins not in category 1—There are 5 and 32 additional proteins not in category 1 of biomarker candidates in female and male subjects, respectively. Some of them have a reasonable AUC of the ROC curve. Some of their sensitivities and specificities are high and others have been reported to be affected by alcohol intake. For example, the level of Apolipoprotein E (APOE) has been reported to increase significantly in alcohol-dependent individuals and return to normal level after completion of alcohol detoxification. A significant correlation has been observed between alcohol consumption during the previous year of alcohol abuse and the APOE values both upon admission to and on discharge from the detoxification program. This suggests that APOE is dependent on alcohol

consumption and may serve as a marker of severe alcohol abuse [19]. Although these proteins are not in category 1, they potentially are valuable biomarker candidates. There is no unlimited source for validation, we have to start with the most promising candidate. The categorization and prioritization approaches are intended to uncover potential protein candidates in a quicker and less expensive way.

4 Concluding remarks

The objective of this research was to detect a potential biomarker (or confined set of biomarkers) to monitor abstinence from alcohol abuse. As a result of the comprehensive mass spectrometry-based analysis used in this study, we have identified potential biomarker candidates whose presence in serum is altered in subjects who drink alcohol abusively compared to subsequent 6 weeks of abstinence. A follow-up study is needed to validate our results.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by NIH-NIAAA R21 AA016217-01 (F.W.) and K08 AA016570 from the NIH/NIAAA, 1101CX000361-01 from the Veterans Affairs Research and Administration, Indiana University Research Support Fund Grant, and W81XWH-12-1-0497 from United States Department of Defense (S.L.).

Abbreviations

AUC	area under the receiver operating characteristic curve
CDT	carbohydrate-deficient transferrin
EtG	ethyl glucuronide
EtS	ethyl sulfate
FAEE	fatty acid ethyl esters
GGT	gamma-glutamyltransferase
MCV	mean corpuscular volume
PEth	Phosphatidylethanol
ROC	receiver operating characteristic
TEP	triethylphosphine

REFERENCES

- [1]. Room R, Babor T, Rehm J. Lancet. 2005; 365:519–530. [PubMed: 15705462]
- [2]. Mathurin P. Liver Transpl. 2005; 11:S21–S24. [PubMed: 16237730]
- [3]. Szabo G, Bakhireva LN, Savage DD. Alcohol Res. Health. 2011; 34:56–63. [PubMed: 23580042]
- [4]. Kasinathan C, Vrana K, Beretta L, Thomas P, Gooch R, Worst T, Walker S, Xu A, Pierre P, Green H, Grant K, Manowitz P. Alcohol. Clin. Exp. Res. 2004; 28:228–232. [PubMed: 15112930]

- [5]. Wu D, Tomonaga T, Sogawa K, Satoh M, Sunaga M, Nezu M, Oh-Ishi M, Kodera Y, Maeda T, Ochiai T, Nomura F. Alcohol. Clin. Exp. Res. 2007; 31:S67–71. [PubMed: 17331169]
- [6]. Nomura F, Tomonaga T, Sogawa K, Ohashi T, Nezu M, Sunaga M, Kondo N, Iyo M, Shimada H, Ochiai T. Proteomics. 2004; 4:1187–1194. [PubMed: 15048998]
- [7]. Faber MJ, Agnetti G, Bezstarosti K, Lankhuizen IM, Dalinghaus M, Guarnieri C, Caldarera CM, Helbing WA, Lamers JM. Cell Biochem. Biophys. 2006; 44:11–29. [PubMed: 16456231]
- [8]. Seibert V, Wiesner A, Buschmann T, Meuer J. Pathol. Res. Pract. 2004; 200:83–94. [PubMed: 15237917]
- [9]. Lai XY, Bacalla RL, Blazer-Yost BL, Hong D, Mason SB, Witzmann FA. Proteomics Clin. Appl. 2008; 2:1140–1152. [PubMed: 20411046]
- [10]. Bradford MM. Anal. Biochem. 1976; 72:248–254. [PubMed: 942051]
- [11]. Keller A, Nesvizhskii AI, Kolker E, Aebersold R. Anal. Chem. 2002; 74:5383–5392. [PubMed: 12403597]
- [12]. Nesvizhskii AI, Keller A, Kolker E, Aebersold R. Anal. Chem. 2003; 75:4646–4658. [PubMed: 14632076]
- [13]. Lai XY, Wang LS, Tang HX, Witzmann FA. J. Proteome Res. 2011; 10:4799–4812. [PubMed: 21888428]
- [14]. Graham K, Wilsnack R, Dawson D, Vogeltanz N. Addiction. 1998; 93:1137–1147. [PubMed: 9813895]
- [15]. Emanuele MA, Wezeman F, Emanuele NV. Alcohol Res. Health. 2002; 26:274–281. [PubMed: 12875037]
- [16]. Park SH, Goo JM, Jo CH. Korean J. Radiol. 2004; 5:11–18. [PubMed: 15064554]
- [17]. Schmidt C, Urlaub H. Methods Mol. Biol. 2012; 893:249–265. [PubMed: 22665306]
- [18]. Franceschini G. Eur. J. Clin. Invest. 1996; 26:733–746. [PubMed: 8889435]
- [19]. Liappas IA, Nikolaou C, Michalopoulou M, Paparrigopoulos T, Tzavellas EO, Piperi C, Cambouri C, Soldatos CR. In Vivo. 2007; 21:699–702. [PubMed: 17708369]

Figure 1. Among the 30 proteins from male subjects that were submitted to IPA, 13 and 6 proteins are included in networks 1 and 2, respectively. Top diseases and functions involved in network 1 are cellular movement, hematological system development and function, and immune cell trafficking. Top diseases and functions involved in network 2 are lipid metabolism, molecular transport, and small molecule biochemistry.

Figure 2.
The ROC curve, sensitivity and specificity, and box plot of APOC1 in female subjects. Its AUC is 0.906 with a sensitivity of 0.875 and specificity of 1.000, indicating that it is a good biomarker candidate.

Figure 3.

The ROC curve, sensitivity and specificity, and box plot of APOD in female subjects. The AUC of APOD is 0.891 with a sensitivity of 0.750 and specificity of 0.875, establishing it as a good biomarker candidate.

Table 1

Proteins whose abundance significantly differed ($p < 0.05$) in female subjects and their prioritization.

R	C	ID	GN	Protein Name	AUC	FC	#P	NC	p
1	1	P02654	APOC1	Apolipoprotein C-I	0.906	1.6	5	0	0.01
2	1	P05090	APOD	Apolipoprotein D	0.891	1.8	7	0	0.007
3	2	E5RII2	CA1	Carbonic anhydrase 1	0.875	1.6	1	0	0.04
4	2	P06702	S100A9	Protein S100-A9	0.875	1.5	1	0	0.004
5	2	K7EJD3	LGALS3BP	Galectin-3-binding protein	0.781	1.7	1	0	0.04
6	3	P03952	KLKB1	Plasma kallikrein	0.797	1.3	7	0	0.04
7	4	P40197	GP5	Platelet glycoprotein V	0.203	-1.4	1	0	0.03

Note: R, Rank; C, Category; ID, Protein ID; GN, Gene Name; AUC, Area Under the Receiver Operating Characteristic (ROC) Curve; FC (FC_PR_PO), fold change; PR, pre-treatment; PO, post-treatment; #P, Number of Peptides; NC, Network Connectivity; and p (p_PR_PO), p value of PR vs. PO.

Table 2

Proteins whose abundance significantly differed ($p < 0.05$) in male subjects and their prioritization.

R	C	ID	GN	Protein Name	AUC	FC	#P	NC	p
1	1	P03952	KLKB1	Plasma kallikrein	0.875	1.4	7	2	0.009
2	1	O95445	APOM	Apolipoprotein M	0.828	1.4	3	2	0.03
3	1	P22352	GPX3	Glutathione peroxidase 3	0.813	1.4	2	2	0.05
4	1	P08571	CD14	Monocyte differentiation antigen CD14	0.797	1.5	5	1	0.03
5	1	P05090	APOD	Apolipoprotein D	0.938	1.9	7	0	0.001
6	1	Q6P163	APOC2	APOC2 protein	0.922	1.5	3	0	0.008
7	1	P04004	VTN	Vitronectin	0.844	1.4	14	0	0.02
8	1	P05154	SERPINA5	Plasma serine protease inhibitor	0.828	1.4	5	0	0.02
9	2	P59665	DEFA1	Neutrophil defensin 1	0.781	2.5	1	2	0.03
10	2	E9PQ57	RAE1	mRNA export factor	0.891	1.8	1	1	0.007
11	2	Q5VV67	PPRC1	Peroxisome proliferator-activated receptor gamma coactivator-related protein 1	0.859	2.2	1	1	0.03
12	2	P02533	KRT14	Keratin, type I cytoskeletal 14	0.781	2.5	1	1	0.04
13	2	K7EJI9	APOC1	Truncated apolipoprotein C-I	0.844	2.6	1	0	0.05
14	2	P06702	S100A9	Protein S100-A9	0.844	1.4	1	0	0.02
15	3	B4E1F0	SERPING1	Plasma protease C1 inhibitor	0.906	1.3	22	4	0.003
16	3	P02649	APOE	Apolipoprotein E	0.875	1.3	17	4	0.005
17	3	P02776	PF4	Platelet factor 4	0.813	1.3	1	2	0.03
18	3	B0YIW2	APOC3	Apolipoprotein C-III	0.766	1.2	5	2	0.05
19	3	P07996	THBS1	Thrombospondin-1	0.750	1.3	5	2	0.05
20	3	R4GMN6	C1R	Complement C1r subcomponent (Fragment)	0.750	1.3	1	2	0.05
21	3	P08603	CFH	Complement factor H	0.859	1.3	71	1	0.02
22	3	Q08380	LGALS3BP	Galectin-3-binding protein	0.859	1.3	6	1	0.02
23	3	P35858	IGFALS	Insulin-like growth factor-binding protein complex acid labile subunit	0.844	1.2	11	1	0.02
24	3	F5GY80	C8B	Complement component C8 beta chain	0.828	1.3	12	1	0.02
25	3	F5H1A8	GSN	Gelsolin	0.781	1.3	28	1	0.05
26	3	P01042	KNG1	Kininogen-1	0.828	1.3	19	0	0.02
27	3	Q96PD5	PGLYRP2	N-acetylmuramoyl-L-alanine amidase	0.828	1.3	18	0	0.03
28	3	C9JV77	AHSG	Alpha-2-HS-glycoprotein	0.813	1.3	14	0	0.03
29	3	D6RF35	GC	Vitamin D-binding protein	0.766	1.3	41	0	0.04
30	3	P07357	C8A	Complement component C8 alpha chain	0.766	1.2	10	0	0.05

Note: R, Rank; C, Category; ID, Protein ID; GN, Gene Name; AUC, Area Under the Receiver Operating Characteristic (ROC) Curve; FC (FC_PR_PO), fold change; PR, pre-treatment; PO, post-treatment; #P, Number of Peptides; NC, Network Connectivity; and p (p_PR_PO), p value of PR vs. PO.

Table 3

The intensity cut-off and its corresponding sensitivity and specificity of proteins in category 1 of the prioritization list.

Sex	Protein ID	Gene Name	Protein Name	Cut-off	Sensitivity	Specificity
Female	P02654	APOC1	Apolipoprotein C-I	13,189,032	0.875	1.000
	P05090	APOD	Apolipoprotein D	4,489,123	0.750	0.875
	P03952	KLKB1	Plasma kallikrein	11,009,035	0.875	0.750
Male	O95445	APOM	Apolipoprotein M	8,474,595	1.000	0.625
	P22352	GPX3	Glutathione peroxidase 3	5,228,642	0.875	0.750
	P08571	CD14	Monocyte differentiation antigen CD14	40,060,652	0.750	0.750
	P05090	APOD	Apolipoprotein D	4,853,342	0.875	1.000
	Q6P163	APOC2	APOC2 protein	576,240	0.875	1.000
	P04004	VTN	Vitronectin	106,023,294	0.625	1.000
	P05154	SERPINA5	Plasma serine protease inhibitor	5,939,552	0.750	0.750