



ORIGINAL ARTICLE

Discovery of plasma biomarkers with data-independent acquisition mass spectrometry and antibody microarray for diagnosis and risk stratification of pulmonary embolism

Bingqing Han^{1,2}  | Chuanbao Li³ | Hexin Li⁴ | Ying Li² | Xuanmei Luo² | Ye Liu² | Junhua Zhang² | Zhu Zhang⁵ | Xiaobo Yu⁶ | Zhenguo Zhai⁵  | Xiaomao Xu⁷ | Fei Xiao^{1,2,4}

¹Peking University Fifth School of Clinical Medicine, Beijing, China

²The Key Laboratory of Geriatrics, Beijing Institute of Geriatrics, Beijing Hospital, National Center of Gerontology, National Health Commission, Institute of Geriatric Medicine, Chinese Academy of Medical Sciences, Beijing, China

³Department of Laboratory Medicine, Beijing Hospital, National Center of Gerontology, National Health Commission, Institute of Geriatric Medicine, Chinese Academy of Medical Sciences, Beijing, China

⁴Clinical Biobank, Beijing Hospital, National Center of Gerontology, National Health Commission, Institute of Geriatric Medicine, Chinese Academy of Medical Sciences, Beijing, China

⁵Department of Respiratory and Clinical Care Medicine, China-Japan, Friendship Hospital, Peking University China-Japan Friendship School of Clinical Medicine, Beijing, China

⁶State Key Laboratory of Proteomics, Beijing Proteome Research Center, National Center for Protein Sciences-Beijing (PHOENIX Center), Beijing Institute of Lifeomics, Beijing, China

⁷Department of Respiratory and Critical Care Medicine, Beijing Hospital, National Center of Gerontology, National Health Commission, Institute of Geriatric Medicine, Chinese Academy of Medical Sciences, Beijing, China

Correspondence

Xiaomao Xu, Department of Respiratory and Critical Care Medicine, Beijing Hospital, #1 Dahua Road, DongCheng District, Beijing, China 100730.
Email: xuxiaomao3361@bjhmoh.cn

Fei Xiao, Clinical Biobank, Laboratory of Cell Biology, Beijing Hospital, #1 Dahua Road, DongCheng District, Beijing, China 100730.
Email: xiaofei3965@bjhmoh.cn

Funding information

CAMS Innovation Fund for Medical Sciences, Grant/Award Number: 2018-I2M-1-002; the National Key Research and Development Program of China, Grant/Award Number: 2016YFC0905600, 2016YFC0905602, 2018YFC2000505 and 2020YFC2009003; the National Natural Science Foundation of China, Grant/Award Number: 81870048

Abstract

Background: Pulmonary embolism (PE) is a leading cause of cardiovascular mortality worldwide. Rapid and accurate diagnosis and risk stratification are crucial for timely treatment options, especially in high-risk PE.

Objectives: The study aims to profile the comprehensive changes of plasma proteomes in PE patients and identify the potential biomarkers for both diagnosis and risk stratification.

Patients/Methods: Based on the data-independent acquisition mass spectrometry and antibody array proteomic technology, we screened the plasma samples (13 and 32 proteomes, respectively) in two independent studies consisting of high-risk PE patients, non-high-risk PE patients, and healthy controls. Some significantly differentially expressed proteins were quantified by ELISA in a new study group with 50 PE patients and 26 healthy controls.

Results: We identified 207 and 70 differentially expressed proteins in PE and high-risk PE. These proteins were involved in multiple thrombosis-associated biological processes including blood coagulation, inflammation, injury, repair, and chemokine-mediated cellular

response. It was verified that five proteins including SAA1, S100A8, TNC, GSN, and HRG had significant change in PE and/or in high-risk PE. The receiver operating characteristic curve analysis based on binary logistic regression showed that the area under the curve (AUC) of SAA1, S100A8, and TNC in PE diagnosis were 0.882, 0.788, and 0.795, and AUC of S100A8 and TNC in high-risk PE diagnosis were 0.773 and 0.720.

Conclusion: As predictors of inflammation or injury repair, SAA1, S100A8, and TNC are potential plasma biomarkers for the diagnosis and risk stratification of PE.

KEY WORDS

biomarkers, mass spectrometry, microarray analysis, proteomics, pulmonary embolism

Essentials

- Clinical laboratory tests of pulmonary embolism (PE) are of great significance.
- The protein proteome of PE was analyzed by mass spectrometry and antibody array.
- Serum amyloid A1 is a potential biomarker for PE diagnosis.
- S100A8 and tenascin C help in the diagnosis and risk stratification of PE.

1 | INTRODUCTION

Pulmonary embolism (PE) is one of the leading causes of cardiovascular death.^{1,2} The case fatality of PE was 8.7% in 2008 in a multicenter registration study in China, and there are 60,000 to 100,000 deaths per annum from PE in the United States estimated by the Centers for Disease Control and Prevention in 2009.^{3,4} Acute PE may cause a significant number of deaths within hours after the onset of symptoms, and the condition is difficult to diagnose due to its non-specific causes and clinical presentation.⁵ High-risk PE patients with arterial hypotension or shock have high short- and long-term mortality.⁶⁻⁸ However, once diagnosed accurately in an early stage, patients can effectively recover with immediate targeted treatment.⁹ Therefore, rapid accurate diagnosis and risk stratification of PE patients are the prerequisites for timely and effective treatment.

Up to now, the diagnosis strategy of PE follows the order of assessment of clinical probability, initial risk assessment, to diagnosis determination by selected tests step by step.¹ The biomarker tests can be easily achieved in point-of-care testing (POCT) or in the laboratory to support negative exclusion and treatment choice.¹⁰ Plasma is one of the most easily accessible bodily fluids in biobanks from thousands of clinical studies. It is a predominant source of potential biomarkers used for diagnostic analyses in clinical practice, as it can directly exchange materials with nearly all organs and tissues.^{11,12} In the diagnosis test of PE, D-dimer has been widely used for negative exclusion. As cardiac biomarkers, brain natriuretic peptide (BNP)/N-terminal-proBNP (NT-proBNP), and cardiac troponins (cTn I/T) have been used for prognostic assessment, mainly focusing on hyperfibrinolysis, ventricular dysfunction, myocardial injury, and hemodynamic disorder.¹³⁻¹⁵ Thus, we launched a study on the plasma proteome features to find new efficient biomarkers in PE patients.

With the fast development of biomarker research technology, high-resolution liquid chromatography-tandem mass spectrometry (LC-MS/MS) has become a powerful tool for the high-throughput characterization of plasma proteome in great depth.¹⁶ Data-independent acquisition mass spectrometry (DIA-MS) has become an emerging bottom-up proteomic strategy for biomarker screening in large-scale clinical samples. It acquires both the first MS and the second MS (the coupled tandem mass spectrometry) data in "full scan" mode, which is independent of the composition of precursor ions for their fragmentation avoiding the bias of precursor ion selection, and the limitations of inherent irreproducibility and undersampling of data dependent acquisition (DDA), and retains comparable features of accuracy, reproducibility, and consistency for targeted data analysis.¹⁷⁻¹⁹ Thus, it can deliver high quality and consistent quantification in large-scale projects in areas such as personalized medicine, biomarker research, drug screens, and multiomics studies. Research on urinary protein biomarkers showed that the DIA assay doubled the number of identified peptides and proteins per sample at half the coefficients of variation observed for DDA data.²⁰ Antibody microarray is another proteomic technique based on the antigen-antibody reaction with high sensitivity, high throughput, and concentration independence.^{21,22} The high-density antibody microarray contains hundreds or thousands of different antibodies specially targeted for different proteins in proteome scale. One primary advantage of antibody microarrays over mass spectrometry is that there is no concern about the ultra-high signals of high abundance proteins and ion suppression.²³ Therefore, it is very suitable for detecting low-abundance proteins in serum or plasma samples.

In the past, only a small number of proteomic studies on PE have been performed, and most of them used two-dimensional gel electrophoresis together with matrix-assisted laser desorption/ionization time of flight mass spectrometry (MALDI-TOF-MS), whose resolution and repeatability might not be feasible for complex human samples.²⁴⁻²⁷

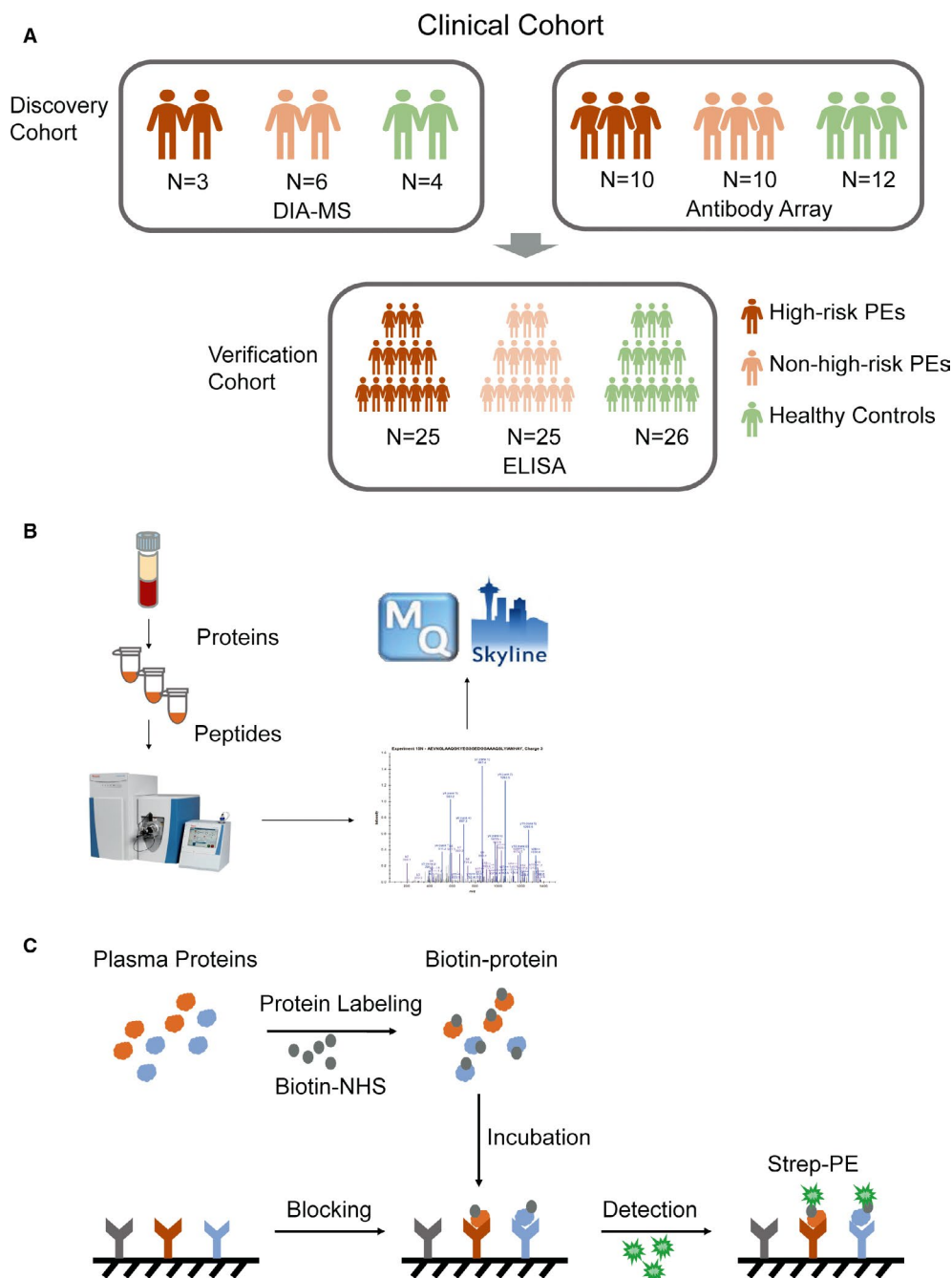


FIGURE 1 Study design. A, Structure of clinical cohort enrolled in this study in different steps including data-independent acquisition mass spectrometry (DIA-MS) method and antibody array method for discovery and ELISA for verification. B, The detection workflow of DIA-MS method. C, The detection principle of the antibody array method. PEs, pulmonary embolism-diagnosed patients

To further understand the plasma proteome in PE patients, we adopt an in-depth proteomics strategy combining DIA-MS and antibody arrays to investigate the plasma protein constructions of 45 proteomes from PE-diagnosed patients and healthy controls in great depth. Differentially expressed proteins revealed the specific changes of biological process in the disease. Among them, candidate biomarkers were verified by ELISA in a new population and assessed for the individual and combined diagnosis performance.

2 | METHODS

2.1 | Clinical samples and study design

The study was a case-control study with the approval of the research ethics committee; the cases were PE-diagnosed patients (PEs) recruited from consecutive respiratory inpatients in Beijing Hospital and China-Japan Friendship Hospital from 2017 to 2018, and the controls

were healthy people or healthy patients with non-PE diseases recruited from regular health check-ups. Written informed consent was provided by all subjects. The design of this study was based on “triangular strategies,” which divides the study into the discovery stage (proteomics) with a small population for screening and the verification stage (classical immunoassays) with a larger individual population.¹⁶ We had recruited more than 580 PE patients and 100 controls for the project. However, only highly paired samples having the same body mass index (BMI), sex, and as few comorbidities as possible were chosen for this study. In the discovery study, six non-high-risk PEs, three high-risk PEs, and four healthy controls were involved in DIA-MS analysis, while 10 non-high-risk PEs, 10 high-risk PEs, and 12 healthy controls were involved in antibody array analysis, which were all semi-quantitative methods. In the verification study, 25 non-high-risk PEs, 25 high-risk PEs, and 26 healthy controls were tested by ELISA, which was an absolute quantitative method with high specificity (Figure 1A). Between the two different stages of study, the PEs were selected randomly except that the patients with fewer comorbidities—such as hypertension, hyperlipemia, coronary artery disease, pulmonary infection, chronic obstructive pulmonary disease (COPD), tuberculosis, dysfunction of liver or kidney, and tumor—were more likely to be selected into the discovery study. The percentage of patients with comorbidities was 27.6% (8/29) in the discovery study and 42% (21/50) in the verification study. Furthermore, in the discovery study, the PE patients had no overlap between DIA-MS and antibody array analysis, while the healthy controls in DIA-MS were pooled from every three healthy controls in antibody array. The verification study was based on completely independent plasma samples.

Sodium citrate-anticoagulated peripheral blood samples were collected from all diagnosed PEs within 24 h of admission and after treatment with subcutaneous low molecular weight heparin (LMWH). Healthy controls were sampled during regular health check-ups. Then the blood was centrifuged at 2050 *g* for 10 min at 4°C. All plasma samples in this study have no jaundice or hemolysis and were immediately aliquoted in sterile tubes, stored at –80°C, and thawed only before the test.

2.2 | Diagnosis and risk stratification

The criteria of diagnostic and risk stratification followed the strategy given by European Society of Cardiology (ESC) both in the discovery and the verification study.¹ All PEs were confirmed with computed tomography pulmonary angiography (CTPA) and ventilation–perfusion scintigraphy (V/Q scan). In the risk stratification study, hemodynamically unstable patients with shock or sustained hypotension were classified as high-risk patients, such as patients with a systolic blood pressure of <90 mmHg or a pressure drop of >40 mmHg for 15 min. Patients who had positive simplified pulmonary embolism severity index (sPESI) score together with either right ventricular dysfunction (by echocardiography or CTPA) or elevated cardiac biomarker levels in the circulation (elevated cardiac troponin I/T or natriuretic peptide concentrations in plasma) were classified as the intermediate-risk

patients. PEs with negative sPESI score were classified as low risk. The latter two categories of patients were grouped as the non-high-risk PEs.

2.3 | Procedure of DIA-MS measurement

The workflow and principle of DIA-MS are integrated in Figure 1B. First, 200 µg of plasma proteins per sample were precipitated by acetone at –20°C for 4 h and redissolved by 0.1 M triethylamine-carbonic acid buffer (TEAB). After being reduced by 5 mM of dithiothreitol for 30 min at 56°C and alkylated by 11 mM of iodoacetamide for 15 min at room temperature in darkness, protein samples were digested overnight by trypsin at the mass ratio of 1:50 (trypsin: protein). Strata X C18 SPE columns (Phenomenex) were used for desalting peptides.

The tryptic peptides were dissolved in 0.1% (v/v) formic acid and separated on EASY-nLC 1200 ultra-performance liquid chromatography (UPLC). The gradient was comprised of an increase from solvent A (0.1% formic acid in 2% acetonitrile) to solvent B (0.1% formic acid in 90% acetonitrile) over 120 min at a constant flow rate of 450 nl/min. The peptides were subjected to nano spray ionization source, followed by tandem mass spectrometry (MS/MS) in Q Exactive™ HF-X (Thermo Fisher Scientific).

In DIA mode, the *m/z* scan range was 385–1200 for full scan in the primary MS, and intact peptides were detected at a resolution of 120,000. Precursor ions in 70 *m/z* windows were fragmented in higher energy collisional dissociation collision (HCD) cell (28% of collision energy). Automatic gain control (AGC) was set at 5E5. Fixed first mass was set as 200 *m/z* in the secondary MS. The spectral library was built by DDA LC-MS/MS spectral data (.dat format) using Skyline (version 4.1.0).²⁸ The corresponding indexed retention time parameters were added into the retention time predictor. The precursor charges were 2,3,4 and ion charges were 1,2 in transition settings. The six most intense product ions were set to be picked from the spectral library for peptide quantification. After adding the decoy peptides, the DIA data were imported and the mProphet method was chosen as the peak scoring model with 1% false discovery rate (FDR). The transformations with log2, normalization, and quantification were carried out in MSstats R package.²⁹

2.4 | Antibody array measurement procedure

The workflow and principle of antibody array are shown in Figure 1C. The biotin label-based antibody array was purchased from Beijing Proteome Research Center, detecting 657 human proteins including cytokines, chemokines, growth and differentiation factors, angiogenic factors, adipokines, adhesion molecules, and matrix metalloproteases, as well as binding proteins, inhibitors, and soluble receptors to these proteins. Bovine serum albumin (BSA, 100 µg/ml; Sigma-Aldrich) was used as negative control and Alexa Fluor 555 goat anti-human IgG (10 µg/ml) was used as positive control.

The detection procedure of the antibody array refers to the previously described procedure with a few modifications.³⁰ In brief, plasma was diluted 10 times with 1×PBS (phosphate buffered saline; pH 7.4) after high speed centrifugation followed by labelling with NHS-PEG4-Biotin (20 g/L in dimethylsulfoxide [DMSO]; Thermo Fisher Scientific) for 1 h at room temperature. Meanwhile, antibody microarrays were blocked with 500 µl of PBS (5% w/v milk) for 1 h at room temperature and then washed by PBST (0.05% w/v Tween-20) and ddH₂O. Bio-Spin® P-6 Gel Columns (Bio-Rad) were used for removing the unbound biotin. The biotinylated proteins were diluted five times with blocking buffer, then incubated with microarrays for 2 h at room temperature. Unbound components were washed out by PBST and water as before. The arrays were incubated with streptavidin-phycoerythrin conjugate (2 µg/ml, Strep-PE, Jackson ImmunoResearch Laboratories Inc.) for 1 h at room temperature in the dark. After being washed, the slide was scanned using the GenePix 4300A microarray scanner at the wavelength of 532 nm.

The fluorescence signal was calculated in median after subtracting the signal of background surroundings with the corresponding exposure hole. The average of the two replicate dots was adopted. The result was normalized by quantile.

2.5 | Bioinformatics analysis

Data processing was executed in R (version 4.0.2). After normalization, the correlation analysis among all plasma samples and between differentially expressed proteins and clinical index were performed with the *cor* function with the Pearson method in stats package (version 4.0.2), and pictured with the *ggcorrplot* package (version 0.1.3). Principal component analysis (PCA) was performed with the *prcomp* function and pictured with the *ggord* package (version 1.1.5). The hierarchical clustering analysis were performed using the *heatmap.2* function in *gplots* package (version 3.1.0). Venn diagrams were drawn by *VennDiagram* package (version 1.6.20).

The differentially expressed proteins were identified using the two samples *t*-distributed test (*P*-value < .05) and the result was adjusted by multiple testing just for reference, including the Bonferroni, Holm, and Benjamini-Hochberg methods. Classification of differentially expressed proteins employed the PANTHER database (<http://pantherdb.org/>). The proteins were classified according to four categories: biological process, cellular component, molecular function, and protein class. Gene Ontology (GO) enrichment analysis was performed in ClueGO and the interaction network of biological process was visualized in Cytoscape (version 3.8.0) using two-sided hypergeometric test with a *P*-value < .01. Functional enrichment analysis of biological processes based on the foldchange was executed on WebGestalt using minimum category size five (<http://webgestalt.org/>).

2.6 | ELISA verification

The proteins to be tested by ELISA in the verification study were selected from the proteins that were gradually increased or decreased

from healthy controls to non-high-risk PEs and then to high-risk PEs in the discovery study by referring to the previously reported studies. The levels of five proteins were determined by ELISA under the guidance of the manufacturer's instructions in Table S1. The detection panel was designed at an alternate sample sequence of three different groups, so as to minimize the batch effect in manual ELISA assays. The result was plugged into the standard reference curve constructed using the four-parameter algorithm or quadratic regression according to the shape of curves. In each clinical group, the data beyond means $\pm 3 \times$ standard deviations were removed as the outliers. Wilcoxon- test or Student's *t* test were used for difference significance test according to the corresponding result of normality test and homogeneity of variance test.

2.7 | ROC analysis for proteins in verification

The individual and combined diagnostic efficiency of the candidate plasma biomarkers in the verification study were assessed by receiver operating characteristic (ROC) analysis based on the protein concentration detected in ELISA. The Pearson correlation among proteins were calculated before multivariate logistic regression. ROC curve, area under the curve (AUC), and significance tests were executed by the pROC package (version 1.16.2) in R. The cutoff values of diagnosis models were calculated by OptimalCutpoints package (version 1.1-4) in the Youden method, which was the sum of sensitivity and specificity minus 1.

2.8 | Correlation analysis between clinical data and plasma proteome

To further demonstrate the relationship between these biomarkers and existing clinical indicators, we analyzed the correlations between biomarker concentration and the result of existing clinical laboratory tests in PE-diagnosed patients in a verification study. The pairwise Pearson correlation test was used and the circo plot was drawn according to the correlation coefficient and *P*-value using the circlize package (version 0.4.10) in R.

3 | RESULTS

3.1 | Profiling of the PE plasma proteome using DIA-MS and antibody array

The sex, age, BMI, clinical test results, and comorbidities of cases and controls are summarized in Table 1. The DIA-MS analysis acquired 1161 identified proteins and 764 quantifiable proteins and the majority of the proteins (94.98%; 624/657) targeted on the antibody array were detected above the detection limit (Figure S1). Next, the detection stability and sample grouping were assessed. In DIA-MS, the Pearson correlation and the PCA analysis among

samples were analyzed (Figure S2A-B). The Pearson correlation analysis compared the correlation of quantification results of all proteins among all samples one by one to evaluate the characteristics of different groups in general. The Pearson correlations and the PCA analysis results showed that the DIA-MS samples could be divided clearly with good grouping features. In the antibody array test, the correlations of within-array and between-array were detected to be 0.987 and 0.964, respectively (Figure S2C). The concentrations of the identified proteins (Figure 2A) in this research ranged across nearly 10 orders of magnitude (10^0 – 10^{10} pg/ml) referring to the reference concentrations from the human plasma proteome database (<http://www.plasmaproteomedatabase.org/>).

The test results of difference significance were summarized in Table S2. There were only very few proteins that significantly changed (adjusted *P*-value < .05) between the high-risk and non-high-risk PEs, CCL16 and SERPIND1, under the multiple test correction (Bonferroni, Holm, and Benjamini-Hochberg methods). The potential protein markers chosen for verification should be gradually increased or decreased from healthy controls to non-high-risk PEs and then to high-risk PEs. Thus, the hypothesis test *P*-values (*P*-value < .05) as the difference test criteria were applied in this study, as widely used in other proteome studies. We analyzed and obtained 56 and 14 differentially expressed proteins in the high-risk PEs and the non-high-risk PEs (PEH_PEnonH) comparison group (Figure 2B); with 188 and 32 differentially expressed proteins in the PEs and healthy controls (PE_Healthy) comparison

groups by two methods, respectively. We compared the proteins in four comparison groups including PEH_PEnonH and PE_H comparing to Healthy (PEH_Healthy), PE_nonH comparing to Healthy (PEnonH_Healthy), and PE_Healthy in Figure 2C. The results showed that 16 proteins had significant differences in all groups, including immunoglobulin lambda variable 3-19, immunoglobulin kappa variable 1D-39, immunoglobulin kappa variable 4-1, immunoglobulin heavy constant alpha 1, fetuin-B, plasminogen, intercellular adhesion molecule 2, actin, serum amyloid A protein 4, protein AMBP, histidine-rich glycoprotein, coagulation factor XIII B, haptoglobin-related protein, serotransferrin, tenascin C, and alpha-1-acid glycoprotein. Furthermore, we found that almost all samples could be divided into their clinical diagnosis groups separately with 56 proteins in the DIA-MS method (differentially expressed in PEH_PEnonH), or with 36 proteins in the antibody array method (differentially expressed between PEH and other groups) in the non-biased hierarchical cluster analysis (Figure S3).

3.2 | Bioinformatic analysis

First, we dissected the differentially expressed proteins in PEH_PEnonH. Seventy proteins were classified according to the biological process, cell component, molecular function, and protein class (Figure S4A). With the result of the two-sided hypergeometric test (*P*-value < .01), 75 biological processes were identified

TABLE 1 Clinical features of the patients enrolled in the study

Assay	DIA-MS		Antibody array			ELISA		
Group	PE_nonH	PE_H	PE_nonH	PE_H	Healthy	PE_nonH	PE_H	Healthy
Number	6	3	10	10	12	25	25	26
Age (years) ^a	67.9 ± 12.7	60.0 ± 16.8	69 ± 9.4	61.6 ± 19.1	60.2 ± 9.2	68 ± 15.2	62 ± 18.5	54 ± 11.8
Male/female (n/n)	4/2	2/1	5/5	4/6	5/7	15/10	11/14	13/13
BMI (kg/m ²) ^a	23.3 ± 1.0	23.2 ± 4.6	24.6 ± 2.6	25.5 ± 5.5	24.3 ± 3.4	24.2 ± 3.2	26.3 ± 5.2	24.5 ± 4.5
D-dimer (mg/L) ^b	0.3 (0.1, 0.7)	4.9 (3.5, 25.6)	1.8 (0.4, 4.9)	5.2 (2.4, 11.1)	—	1.0 (0.4, 4.1)	5.5 (2.4, 12)	—
cTN I (μg/L) ^b	0 (0, 0)	0.7 (0.4, 1.1)	0 (0, 0)	0.1 (0, 0.9)	—	0 (0, 0)	0.1 (0, 1)	—
cTN T (μg/L) ^b	0 (0, 1)	—	0 (0, 0.1)	0 (0, 0.1)	—	0.1 (0, 0.1)	0 (0, 0.1)	—
BNP (ng/L) ^b	67.4 (28.4, 99.8)	236.5	86.8 (34, 189.5)	121.4 (50.8, 849.4)	—	28.4 (17.1, 458.7)	114.7 (20.0, 458.7)	—
NT-proBNP (μg/L) ^b	—	4.7	1.0 (0.1, 4.0)	2.2 (0.4, 5.5)	—	0.8 (0.3, 1.3)	2.2 (0.7, 4.7)	—
Comorbidities (n)								
Cardiac dysfunction	1	0	3	1	0	0	0	0
CHD	0	1	0	1	0	3	4	0
Pulmonary infection	0	1	1	2	0	2	3	0
Tumor	0	0	2	1	0	4	3	0
Diabetes	0	0	1	1	1	1	1	0

Abbreviations: BMI, body mass index; CHD, coronary heart disease; DIA-MS, data-independent acquisition mass spectrometry; PE_H, high-risk pulmonary embolism patients; PE_nonH, non-high-risk pulmonary embolism patients.

^aData are expressed as mean ± standard deviation.

^bData are expressed as median (1/4 quartile, 3/4 quartile).

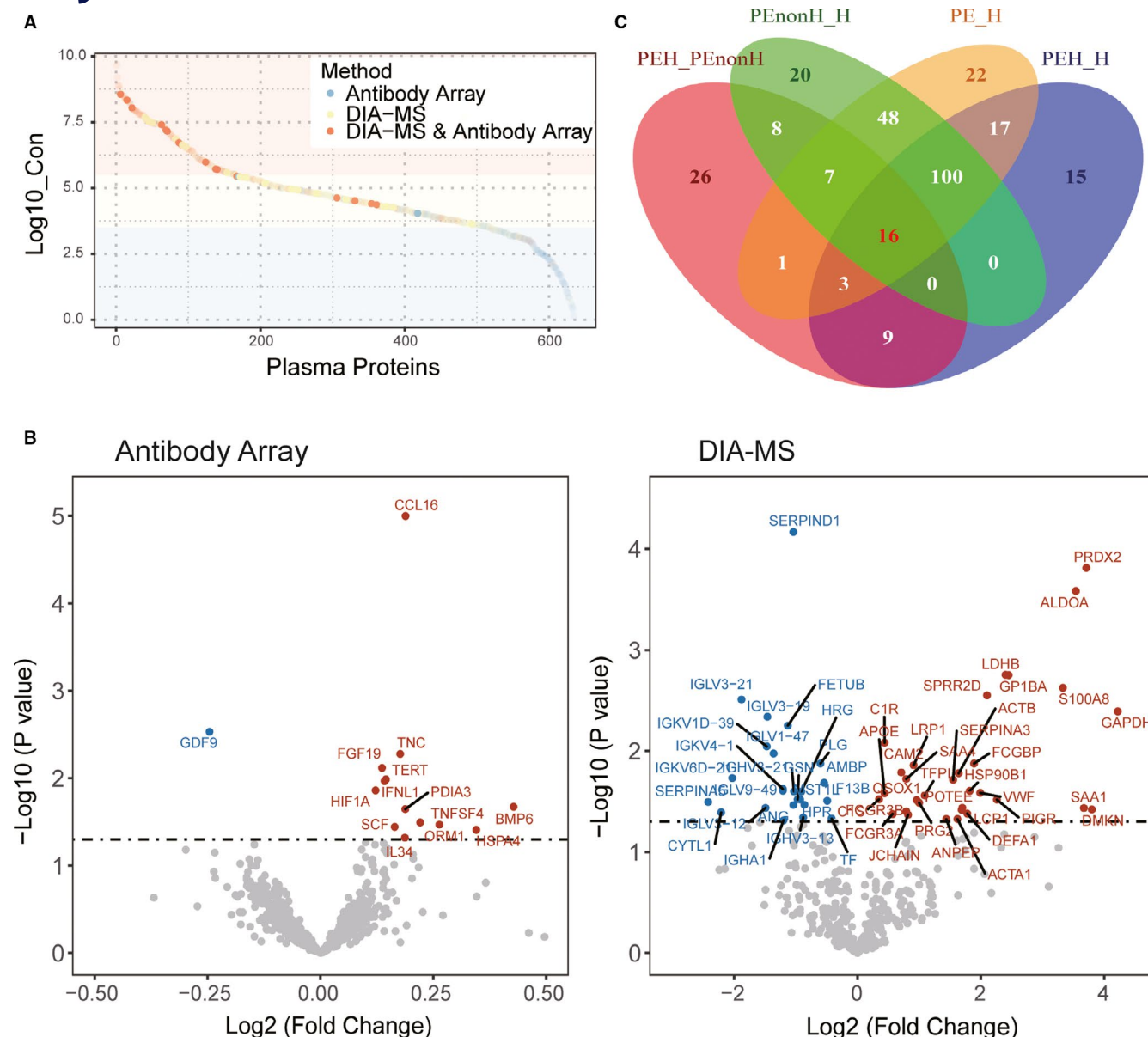


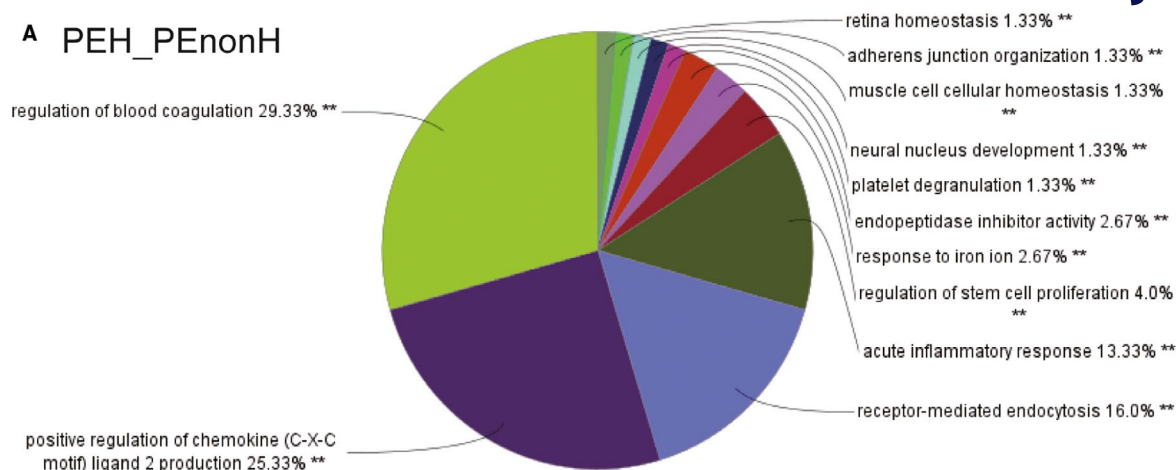
FIGURE 2 General data from the discovery study. A, The concentration of all plasma proteins identified in this in-depth proteome research. B, Volcano plot of proteins identified in antibody array and data-independent acquisition mass spectrometry (DIA-MS), respectively, in the PEH_PEnonH comparison group. Points above the dotted line stand for P -value $< .05$. Red and blue stand for the upregulated and downregulated proteins, respectively. C, Venn diagram of differentially expressed proteins in four comparison groups. PE_H, high-risk pulmonary embolism patients; PE_nonH, non-high-risk pulmonary embolism patients

through GO analysis, among which 4 biological processes had the highest proportion of each group, including regulation of blood coagulation (GO:0030193: APOE, GP1BA, HRG, PLG, PRDX2, TFPI), positive regulation of chemokine (C-X-C motif) ligand 2 production (GO:2000343: LCP1, LRP1, PRG2, S100A8), receptor-mediated endocytosis (GO:0006898: AMBP, APOE, HPR, HSP90B1, LRP1,

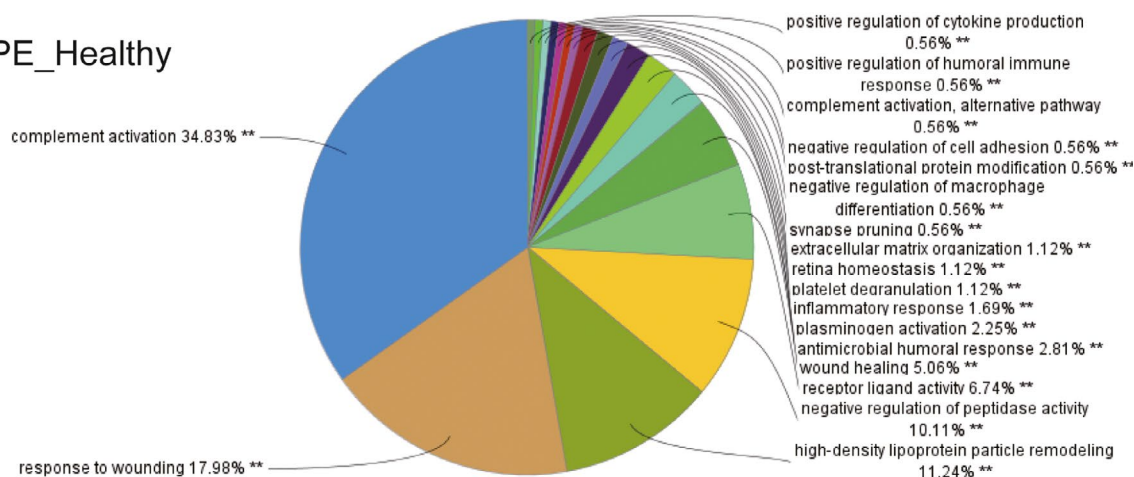
SAA1, TF, and immunoglobulin peptides), and acute inflammatory response (GO:0002526: HPR, ORM1, S100A8, SAA1, SAA4, SERPINA3, TNFSF4) from high to low (Figure 3A). The network of biological processes suggested that there were interactions between the acute inflammatory response and the positive regulation of chemokine production (Figure 3B). The result of functional

FIGURE 3 Biological processes in Gene Ontology (GO) function enrichment analysis. A, Pie plot of enrichment biological processes of differentially expressed proteins in PEH_PEnonH and PE_Healthy. The percentage means the proportion of the category in GO terms. B, The network of biological processes in PEH_PEnonH. The color stands for the enrichment degree and the size stands for the protein number in this term

A PEH_PEnonH



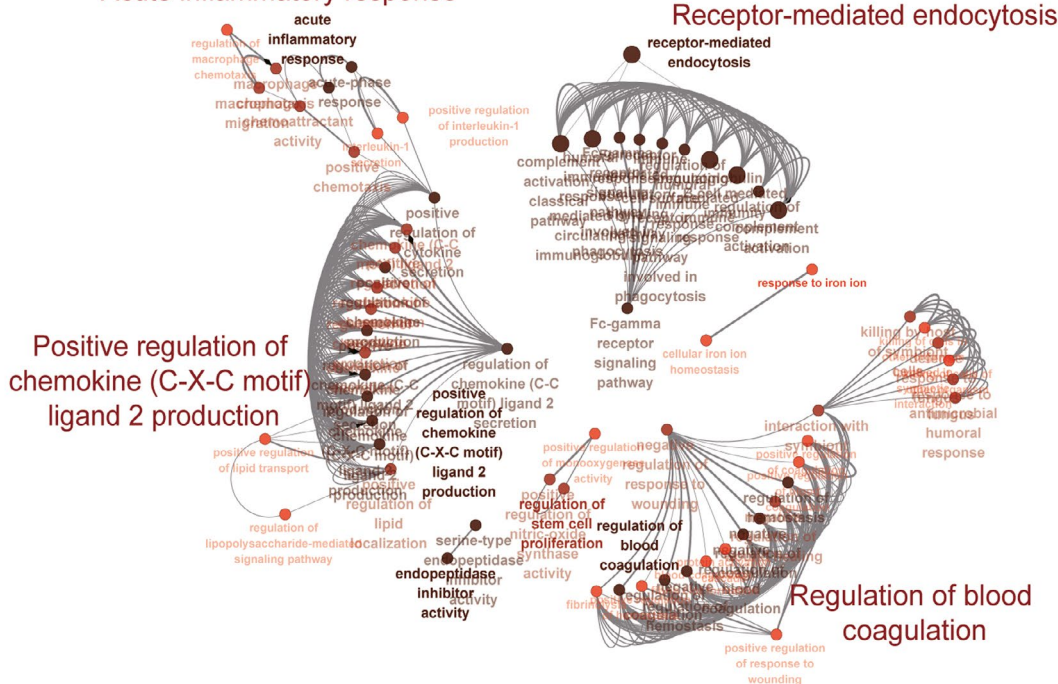
PE_Healthy



B

Acute inflammatory response

Receptor-mediated endocytosis



enrichment analysis based on fold-change rank in WebGestalt (number of IDs in the category >5, top 10 significance level) is shown in Figure S5A. Three upregulated biological processes (FDR < 0.05) were positive regulation of cytokine production (GO:0001819: GAPDH, SAA1), regulation of peptide secretion (GO:0002791: GAPDH, SAA1, S100A8), and positive regulation of secretion (GO:0051047: GAPDH, SAA1, S100A8).

Subsequently, we analyzed the 207 differentially expressed proteins in PE_Healthy. Similarly, proteins were classified in four categories and showed a similar construction (Figure S4B). Then, the most enriched biological processes were derived as complement activation (GO:0006956, 34.83%); response to wounding (GO:0009611, 17.98%); high-density lipoprotein particle remodeling (GO:0034375, 11.24%); and negative regulation of peptidase activity (GO:0010466, 10.11%), including proteins such as C3, C5, plasminogen, and apolipoproteins that were widely known as key players in the initiation and therapy of thrombus embolism disease³¹⁻³³ (Figure 3A). The functional enrichment analysis (top 15 significance level) received a supporting result that multiple cellular responses changed significantly (Figure S5B), and most of the enriched proteins were polypeptide chains of fibrinogen, platelet-associated protein, and serum amyloid A1 in upregulated biological processes like ERK1 and ERK2 cascade (GO:0070371), regulation of cell–cell adhesion (GO:0022407), positive regulation of cell adhesion (GO:0045785), and downregulated processes like response to chemokine (GO:1990868).

3.3 | ELISA verification for five potential plasma biomarkers

ELISA was used for the verification of candidate biomarkers. Five differentially expressed proteins including SAA1 (serum amyloid A-1), S100A8/A9 (calprotectin), TNC (tenascin-C), GSN (gelsolin), and HRG (histidine-rich glycoprotein) were chosen and detected within a few days to reduce repeated freezing and thawing. In the regression method of standard curve, GSN was calculated by quadratic regression and the others were calculated by four-parameter algorithm. There was one outlier in each of the PE_nonH and healthy groups in S100A8, one outlier in each of three groups in TNC and HRG, one outlier in each of PE_H and PE_nonH groups in GSN, and no outlier in SAA1. The result of normality test showed that only GSN was normally distributed so that Student's *t* test was used for GSN comparison and the Wilcoxon test was used for the others. Differences in concentrations of five proteins among three groups are shown in Figure 4 and Table S3, which demonstrates that almost all of the proteins in all comparison groups have significant differences (*P*-value < .05 or less) except for SAA1 and HRG in PEH_PEnonH and S100A8 in PEnonH_Healthy. In addition, no significant change was observed before and after the outliers were removed. Furthermore, all proteins showed significant differences (*P*-value < .01 or less) in PE_Healthy.

3.4 | Diagnosis performance of candidate plasma biomarkers

First, we inspected the individual diagnostic performance of five proteins between the PEs and healthy controls (Figure S6A) and three significantly changed proteins between the high-risk PEs and non-high-risk PEs (Figure S6B). In Figure S6, the AUC of SAA1 was 0.882 and those of other proteins were in the range of 0.7 to 0.8 in PE_Healthy diagnosis, meanwhile the AUC of S100A8 and TNC were higher than 0.7 in PEH_PEnonH diagnosis. In PE_Healthy diagnosis, the sensitivity and the specificity of SAA1 were 0.837 and 0.818, respectively, at the cutoff value of 1.26 µg/ml; the sensitivity and the specificity of S100A8 were 0.714 and 0.840, respectively, at the cutoff value of 1.19 µg/ml; the sensitivity and the specificity of TNC were 0.625 and 0.920, respectively, at the cutoff value of 12.62 ng/ml (Figure 5A). In PEH_PEnonH diagnosis, S100A8 had 0.760 sensitivity and 0.708 specificity at the cutoff value of 1.70 µg/ml; TNC had 0.667 sensitivity and 0.750 specificity at the cutoff value of 17 ng/ml (Figure 5B).

Second, the combined diagnosis performances for PE were analyzed. In Pearson correlation analysis, there were significant correlations between SAA1 and S100A8/TNC/GSN, TNC and S100A8, GSN and HRG (*P*-value < .5). However, the significance test of AUC of combined proteins showed that there was no significant improvement in all combined groups both in PE_Healthy diagnosis and PEH_PEnonH diagnosis (Figure S7). In consequence, SAA1, S100A8, and TNC show a good performance in PE diagnosis, and the latter two proteins could further help with the high-risk PE diagnosis.

3.5 | Correlation analysis between clinical data and biomarkers

The results are shown in Figure 6. Homocysteine (HCY) was positively correlated with SAA1 and S100A8, blood glucose was positively correlated with S100A8 and TNC, and NT-proBNP was negatively correlated with GSN and HRG, consistent with the reported elevated plasma levels of HCY and NT-proBNP in PE and elevated blood glucose level in PEs with high short-term mortality.³⁴⁻³⁶

4 | DISCUSSION

In our study, there were 70 differentially expressed proteins between the high-risk and the non-high-risk PEs and 207 proteins between the PEs and the healthy controls. Our study reveals that the regulation of chemokine (especially the chemokine ligand 2), acute inflammatory response, and some other cellular responses may play important roles in thrombosis and embolism and might be correlated with the risk level of the disease. Previous studies have suggested that the imbalance between prothrombotic and antithrombotic cytokines/chemokines might be involved in the pathophysiology of

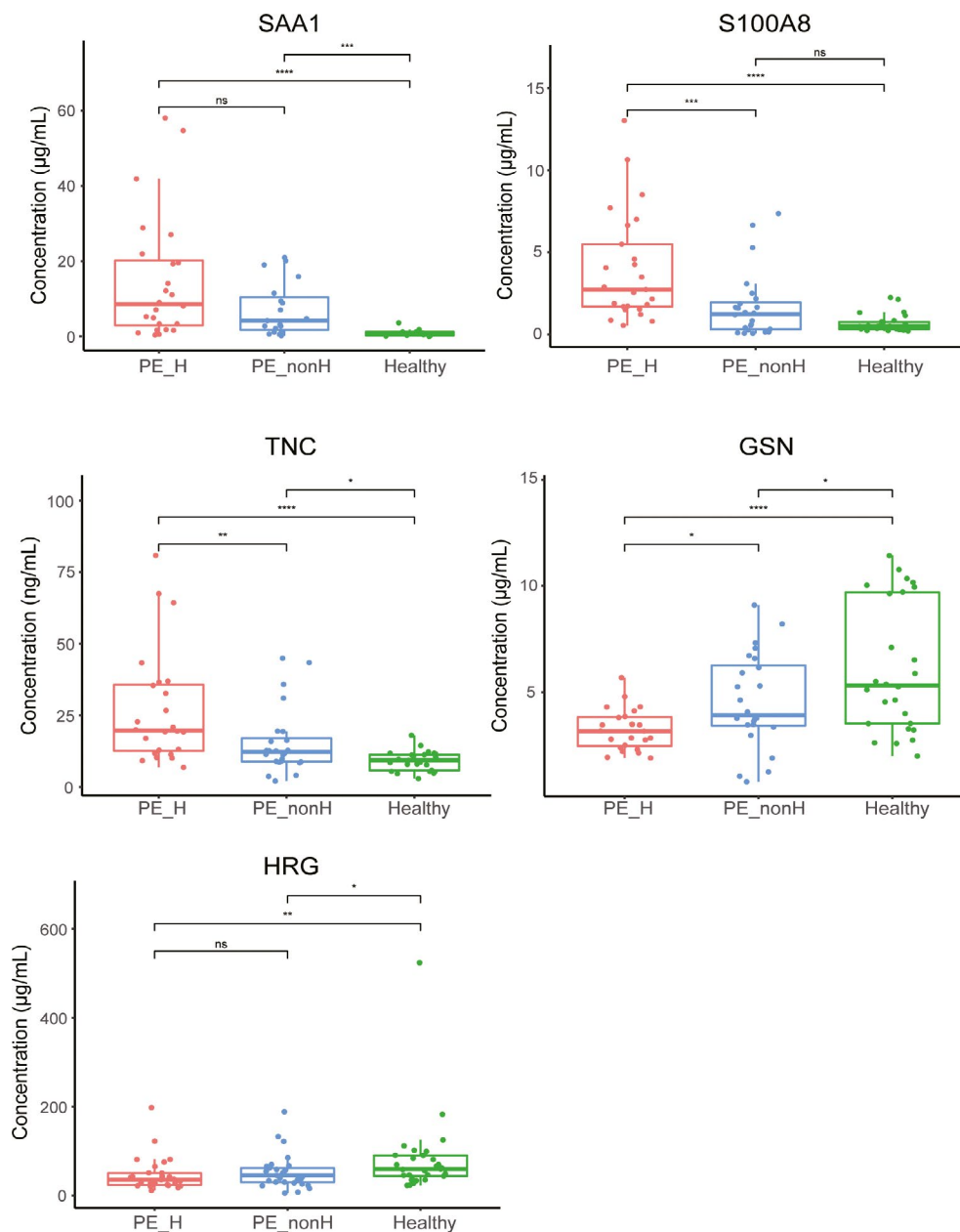


FIGURE 4 ELISA results of SAA1, S100A8, TNC, GSN, and HRG in high-risk PEs, non-high-risk PEs, and healthy controls without outliers. As the result of significance test, * means P -value $< .05$; ** means P -value $< .01$; *** means P -value $< .001$; **** means P -value $< .0001$; ns means P -value $> .05$

venous thromboembolism,³⁷ but it remains controversial whether inflammation is a cause or a result of thrombosis.³⁸ Nevertheless, inflammatory regulation may be a valuable target in PE.

Five plasma proteins were verified for the concentration difference in different PE risk groups and healthy controls. Interestingly, two inflammation-associated proteins, SAA1 and S100A8, were overwhelmingly expressed in PE. SAAs have four members in human, among which SAA1/2 are the most prominent members of the acute phase response in inflammation, known as "A-SAA."^{39,40} In addition, SAA4 was found to be significantly upregulated in PEs in a lower fold change in DIA-MS detection, known as "C-SAA." SAAs have cytokine-like activities to stimulate multiple receptors in

inflammation, as demonstrated at the level of proteomics (Figure 3B). The acute inflammatory response is linked to the positive regulation of chemokine (C-X-C motif) ligand 2 production in the biological process network of differentially expressed proteins between high-risk and non-high-risk PEs. In addition, SAAs are reported to be lipophilic and contribute to high density lipoproteins (HDL) and cholesterol transport,^{41,42} which is consistent with the results of the enriched process of HDL particle remodeling in our study. Recently, SAAs have become a hotspot in the search for biomarkers of inflammation compared to c-reactive protein (CRP) and procalcitonin (PCT) and are gradually being promoted to the clinical laboratory.^{43,44} Interestingly, we did not see significant change in CRP, one of the

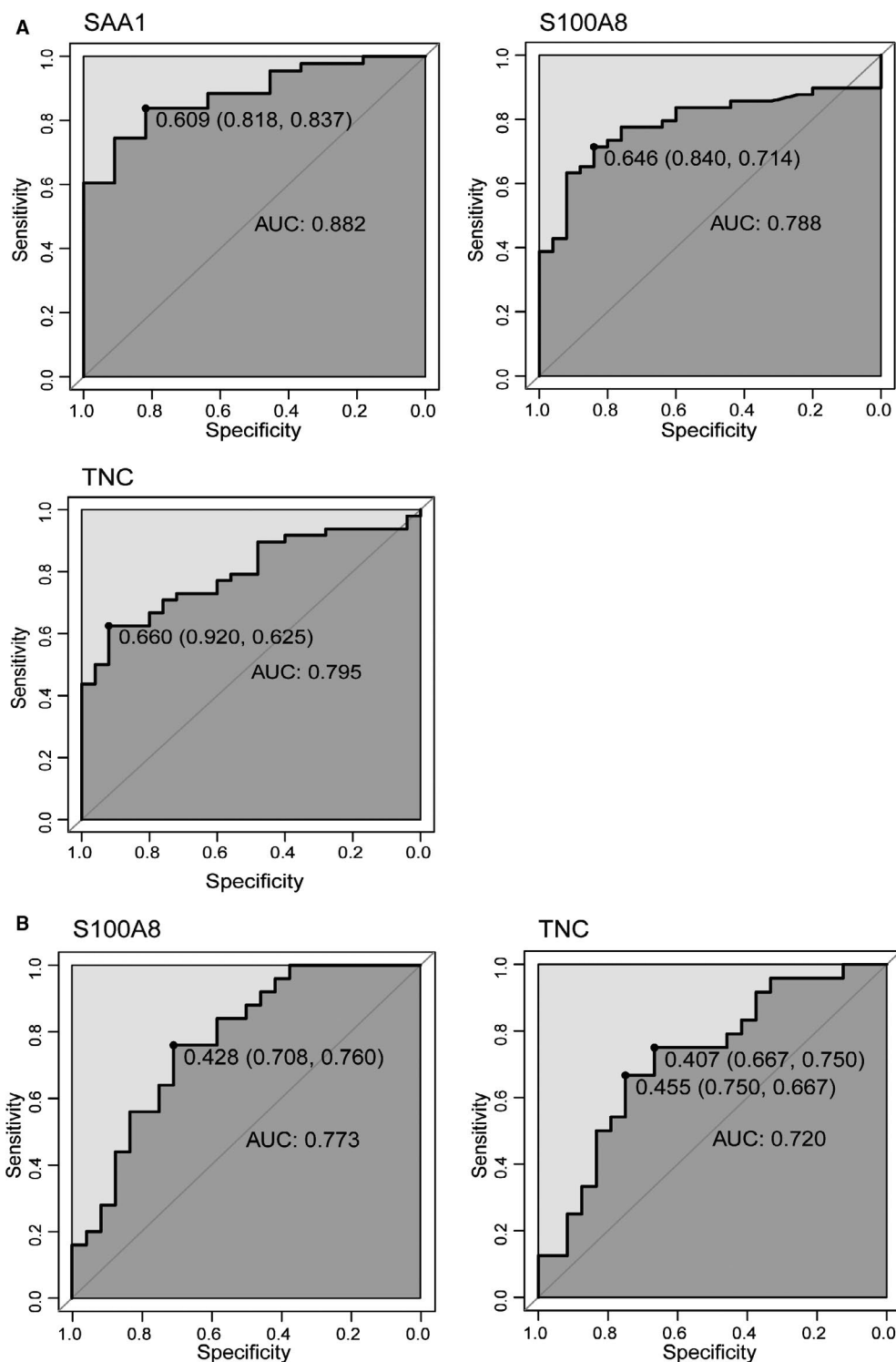


FIGURE 5 Receiver operating characteristic (ROC) results of different diagnosis test models in PEH_PENonH and PE_Healthy. A, SAA1, S100A8, and TNC testing model for PE diagnosis. B, S100A8 and TNC testing model for risk stratification. The point on the curve with label is the performance at the best cutoff-value

most widely used inflammatory indicators and cardiovascular biomarkers, in the discovery study both by DIA-MS and antibody array method. Therefore, the application of SAA can provide additional information for the underlying inflammatory process and lipid metabolism state. S100A8 is a Ca^{2+} -binding protein belonging to the

S100 family and exists in the form of heterodimer with S100A9 in plasma, also well known as MRP8. It also exerts a critical role in modulating the inflammatory response by stimulating leukocyte recruitment and inducing cytokine secretion, which are crucial in the occurrence and development of PE.⁴⁵ It has been recognized as

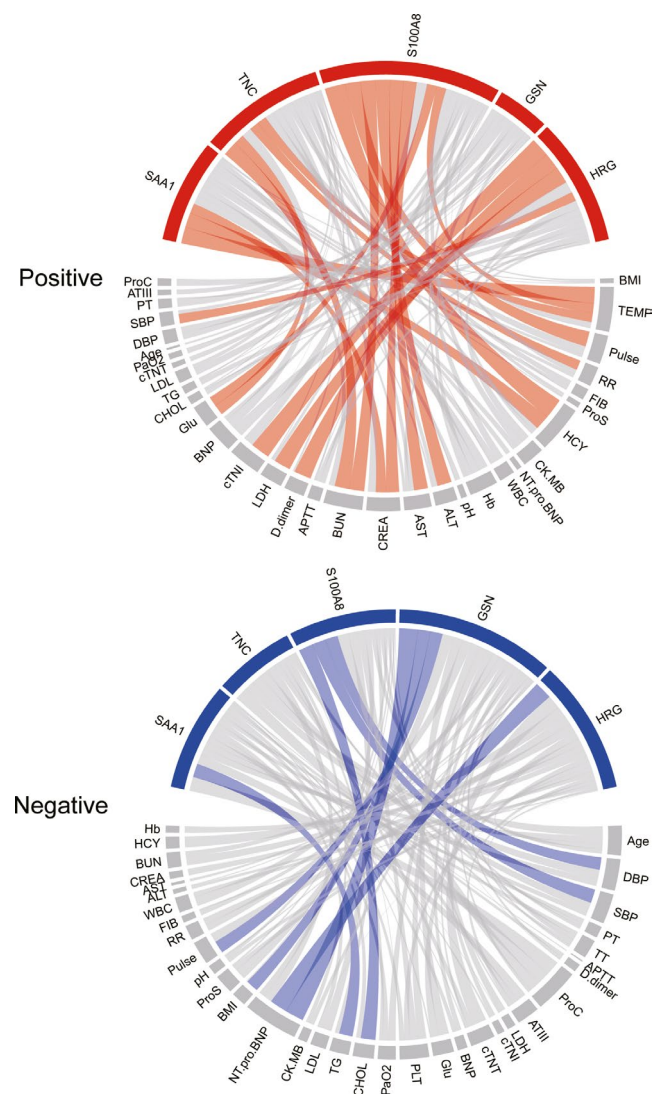


FIGURE 6 Correlation network of five biomarkers and clinical indicators in pulmonary embolism patients using circos. Correlations with statistical significance (P -value $< .05$) are indicated in red for positive correlations and in blue for negative correlations. Non-significant correlations (P -value $\geq .05$) are indicated in gray

a novel inflammatory biomarker and screening tool of autoinflammatory diseases.⁴⁶ Tenascin-C was another upregulated protein in our study, acting on the PEs in a different way. TNC is a hexameric multimodular extracellular matrix protein that is often used as a signal for successful tissue repair and has been used in tissue-repair and stem-cell-based tissue replacement.⁴⁷ TNC expression is sparse in the normal healthy adult organs but can be rapidly induced in many tissues in response to pathological stress. It keeps pace with infection and inflammation and appears after mechanical or chemical injury, including mechanical overload of muscle and tendons,⁴⁸ which is consistent with trauma, the common risk factor of PE. It corresponds to the enriched wound and wound-healing processes in Figure 3A in our study. On the contrary, GSN and HRG were downregulated in PEs. GSN is a negatively regulated inflammation factor

and its plasma form may depolymerize and sequester actin released into the vasculature after cell damage and death so as to rapidly clear actin from the circulation.^{49,50} Similarly, the upregulated cytoplasmic actin (ACTB) and downregulated GSN in plasma were all observed in our study (Figure 2B). Because of the extracellular actin scavenging function, gelsolin is also a potential therapeutic target in addition to inflammatory predictor. HRG is one of the major plasma proteins and thought to function in blood coagulation, fibrinolysis, and innate immune systems by interacting with heparin/heparan sulfate, fibrinogen, and plasminogen.⁵¹ In addition, HRG in plasma is functionally regulated by Zn^{2+} by providing protection against proteolytic disassembly by serine proteases in settings of tissue injury.⁵² Interestingly, our research proved that serine protease inhibitors (Serpins) had complex diverse performance in PEs (Figure 2B).

Finally, we evaluated the diagnosis performance of these five plasma biomarkers in PEs. SAA1 presented good performance for the PE diagnosis with 0.818 specificity and 0.837 sensitivity at the optimal cut-off point (1.26 $\mu\text{g/ml}$) but is not obviously useful for high-risk PE. S100A8 and TNC have certain diagnosis values at the optimal cut-off point in PEs (1.19 $\mu\text{g/ml}$ and 12.62 ng/ml) and high-risk PEs (1.7 $\mu\text{g/ml}$ and 17.0 ng/ml). However, none of them showed confirmed synergistic effect with each other. To further demonstrate the clinical utility in practice, we compared the novel biomarkers in this study with the existing biomarkers, including the temperature, pulse, respiration, blood pressure, and PE-associated clinical tests. It further indicated the relationship between biomarkers and clinical features that might give additional advice on the clinical application. Certainly, the possibility of clinical application and potential diagnostic panel need further validation.

However, there were some limitations in our study. First, the widely used laboratory tests in PE serve for negative prediction (D-dimer) and risk stratification (cardiac markers). In our study, some of the significantly changed proteins in PE were related to the diagnosis and severity assessment, and might be helpful for treatment by indicating severe inflammation or injury. Nevertheless, clinically the diagnosis and risk stratification are performed mainly based on clinical parameters, and whether these protein biomarker tests (which may take more than an hour) work in the real-world settings remain doubtful. Second, treatment with subcutaneous LMWH in PE patients may affect the levels of circulating plasma proteins. Therefore, the differences reported in our study were inevitably affected by heparin treatment, which limited their application and interpretation. It has been reported that heparin can cause the formation of various amyloid deposits aroused by amyloid A (AA) amyloidosis,^{53,54} suppress the induction of S100A8, block the deposition of S100A8/9 heterodimer onto the endothelium of venules in inflamed tissues,^{55,56} and enhance the production of tenascins.⁵⁷ Thus, the effect of LMWH is a complex, multidimensional, and comprehensive process to be further investigated. Moreover, according to the European Respiratory Society guidelines on the diagnosis and management of acute pulmonary embolism,¹ parenteral anticoagulation (e.g., intravenous unfractionated heparin, subcutaneous LMWH, or

subcutaneous fondaparinux) should be initiated while awaiting the results of diagnostic tests in patients with high or intermediate clinical probability for PE. Therefore, it is impractical and unethical to ask if a patient would like to participate in this study before receiving emergency treatments including anticoagulation under an urgent situation. Finally, our current study was not a prospective cohort study so all participants were not followed-up. Because all the patients were sampled within 24 h after admission, the samples reflected more about the initial stage of the disease.

In addition, there were some other technical limitations in this study. First, the number of plasma samples was relatively small, and the results need to be further validated in a larger population both for quantification and for ROC analysis. Second, the quality of antibody, array printing, and manual operation can deeply influence the quality of plasma proteomic measurement. Third, many differentially expressed proteins including serpins, CCL16, ALDOA, and FETUB have not been checked for further usage in the verification stage under the strict screening criteria.

In conclusion, we comprehensively profiled the plasma proteomic feature of PE in the depth of almost 10 orders of concentration and discovered significantly changed proteins and biological processes.

ACKNOWLEDGMENTS

This work was supported by the National Key Research and Development Program of China (2016YFC0905600, 2016YFC0905602, 2018YFC2000505, and 2020YFC2009003), CAMS Innovation Fund for Medical Sciences (2018-I2 M-1-002), and the National Natural Science Foundation of China (81870048). Thanks go to Ms. Meng Xu for help with antibody array detection. Thanks also go to Ms. Chang Zheng for help with the data analysis in the antibody array. The DIA-MS measurement was performed by PTM Biolabs, Inc.

CONFLICTS OF INTEREST

The authors declare no competing financial interests.

AUTHOR CONTRIBUTIONS

F.X. conceived and designed the experiments. B.Q.H. conducted the experiments and data analysis. C.B.L., H.X.L., J.H.Z., Z.Z., and Z.G.Z. performed the collection of clinical samples and case information. Ying.L., X.M.L., and Ye.L. performed the sample pretreatment. X.B.Y. provided technical support of experiments. B.Q.H. and F.X. wrote the manuscript. F.X. and X.M.X. played advisory roles.

ORCID

Bingqing Han  <https://orcid.org/0000-0001-6920-4480>

Zhenguo Zhai  <https://orcid.org/0000-0002-7096-8792>

REFERENCES

- Konstantinides SV, Meyer G, Becattini C, et al. 2019 ESC Guidelines for the diagnosis and management of acute pulmonary embolism developed in collaboration with the European Respiratory Society (ERS): The Task Force for the diagnosis and management of acute pulmonary embolism of the European Society of Cardiology (ESC). *The European respiratory journal*. 2019;54(3):1901647.
- Cohen AT, Agnelli G, Anderson FA, et al. Venous thromboembolism (VTE) in Europe: the number of VTE events and associated morbidity and mortality. *Thromb Haemost*. 2007;98(4):756-764.
- Yang Y, Liang L, Zhai Z, et al. Pulmonary embolism incidence and fatality trends in Chinese hospitals from 1997 to 2008: a multicenter registration study. *PLoS ONE*. 2011;6(11):e26861.
- Beckman MG, Hooper WC, Critchley SE, Ortel TL. Venous thromboembolism: a public health concern. *Am J Prev Med*. 2010;38(4 Suppl):S495-501.
- Girard P, Sanchez O, Leroyer C, et al. Deep venous thrombosis in patients with acute pulmonary embolism: prevalence, risk factors, and clinical significance. *Chest*. 2005;128(3):1593-1600.
- Di Nisio M, van Es N, Büller HR. Deep vein thrombosis and pulmonary embolism. *Lancet (London, England)*. 2016;388(10063):3060-3073.
- Raskob GE, Angchaisuksiri P, Blanco AN, et al. Thrombosis: a major contributor to global disease burden. *Arterioscler Thromb Vasc Biol*. 2014;34(11):2363-2371.
- Goldhaber SZ, Visani L, De Rosa M. Acute pulmonary embolism: clinical outcomes in the International Cooperative Pulmonary Embolism Registry (ICOPER). *Lancet (London, England)*. 1999;353(9162):1386-1389.
- Torbicki A, Perrier A, Konstantinides S, et al. Guidelines on the diagnosis and management of acute pulmonary embolism: the Task Force for the Diagnosis and Management of Acute Pulmonary Embolism of the European Society of Cardiology (ESC). *Eur Heart J*. 2008;29(18):2276-2315.
- Borohovitz A, Weinberg MD, Weinberg I. Pulmonary embolism: Care standards in 2018. *Prog Cardiovasc Dis*. 2018;60(6):613-621.
- Végvári A, Welinder C, Lindberg H, Fehniger TE, Marko-Varga G. Biobank resources for future patient care: developments, principles and concepts. *Journal of clinical bioinformatics*. 2011;1(1):24.
- Keshishian H, Burgess MW, Specht H, et al. Quantitative, multiplexed workflow for deep analysis of human blood plasma and biomarker discovery by mass spectrometry. *Nat Protoc*. 2017;12(8):1683-1701.
- Kearon C, de Wit K, Parpia S, et al. Diagnosis of pulmonary embolism with d-dimer adjusted to clinical probability. *The New England journal of medicine*. 2019;381(22):2125-2134.
- du Fay de Lavallaz J, Badertscher P, Nestelberger T, et al. B-type natriuretic peptides and cardiac troponins for diagnosis and risk-stratification of syncope. *Circulation*. 2019;139(21):2403-2418.
- Klok FA, Mos IC, Huisman MV. Brain-type natriuretic peptide levels in the prediction of adverse outcome in patients with pulmonary embolism: a systematic review and meta-analysis. *Am J Respir Crit Care Med*. 2008;178(4):425-430.
- Geyer PE, Holdt LM, Teupser D, Mann M. Revisiting biomarker discovery by plasma proteomics. *Mol Syst Biol*. 2017;13(9):942.
- Zhang F, Ge W, Ruan G, Cai X, Guo T. Data-independent acquisition mass spectrometry-based proteomics and software tools: a glimpse in 2020. *Proteomics*. 2020;20(17-18):e1900276.
- Gillet LC, Navarro P, Tate S, et al. Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. *Mol Cell Proteomics*. 2012;11(6):O111.016717.
- Ludwig C, Gillet L, Rosenberger G, Amon S, Collins BC, Aebersold R. Data-independent acquisition-based SWATH-MS for quantitative proteomics: a tutorial. *Mol Syst Biol*. 2018;14(8):e8126.
- Muntel J, Xuan Y, Berger ST, et al. Advancing urinary protein biomarker discovery by data-independent acquisition on a quadrupole-orbitrap mass spectrometer. *J Proteome Res*. 2015;14(11):4752-4762.
- Yu Z, Liu L, Yu X, et al. High-throughput antibody generation using multiplexed immunization and immunogen array analysis. *J Biomol Screen*. 2010;15(10):1260-1267.

22. Huang R, Jiang W, Yang J, et al. A biotin label-based antibody array for high-content profiling of protein expression. *Cancer Genomics Proteomics*. 2010;7(3):129-141.
23. Hou X, Zhang X, Wu X, et al. Serum protein profiling reveals a landscape of inflammation and immune signaling in early-stage COVID-19 infection. *Mol Cell Proteomics*. 2020;19(11):1749-1759.
24. Li SQ, Qi HW, Wu CG, et al. Comparative proteomic study of acute pulmonary embolism in a rat model. *Proteomics*. 2007;7(13):2287-2299.
25. Insenser M, Montes-Nieto R, Martínez-García M, et al. Identification of reduced circulating haptoglobin concentration as a biomarker of the severity of pulmonary embolism: a nontargeted proteomic study. *PLoS ONE*. 2014;9(6):e100902.
26. Zhang YX, Li JF, Yang YH, et al. Identification of haptoglobin as a potential diagnostic biomarker of acute pulmonary embolism. *Blood Coagul Fibrinolysis*. 2018;29(3):275-281.
27. Watts JA, Lee YY, Gellar MA, Fulkerson MB, Hwang SI, Kline JA. Proteomics of microparticles after experimental pulmonary embolism. *Thromb Res*. 2012;130(1):122-128.
28. MacLean B, Tomazela DM, Shulman N, et al. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics*. 2010;26(7):966-968.
29. Choi M, Chang CY, Clough T, et al. MSstats: an R package for statistical analysis of quantitative mass spectrometry-based proteomic experiments. *Bioinformatics*. 2014;30(17):2524-2526.
30. Xu M, Deng J, Xu K, et al. In-depth serum proteomics reveals biomarkers of psoriasis severity and response to traditional Chinese medicine. *Theranostics*. 2019;9(9):2475-2488.
31. Ellis DA, Neville E, Hall RJ. Subacute massive pulmonary embolism treated with plasminogen and streptokinase. *Thorax*. 1983;38(12):903-907.
32. McFadyen JD, Stevens H, Peter K. The emerging threat of (Micro) Thrombosis in COVID-19 and its therapeutic implications. *Circ Res*. 2020;127(4):571-587.
33. Obermayer G, Afonyushkin T, Binder CJ. Oxidized low-density lipoprotein in inflammation-driven thrombosis. *J Thromb Haemost*. 2018;16(3):418-428.
34. Cellai AP, Lami D, Antonucci E, et al. Hyperhomocysteinemia in patients with pulmonary embolism is associated with impaired plasma fibrinolytic capacity. *J Thromb Thrombolysis*. 2014;38(1):45-49.
35. Pruszczyk P, Kostrubiec M, Bochowicz A, et al. N-terminal pro-brain natriuretic peptide in patients with acute pulmonary embolism. *The European respiratory journal*. 2003;22(4):649-653.
36. Scherz N, Labarère J, Aujesky D, Méan M. Elevated admission glucose and mortality in patients with acute pulmonary embolism. *Diabetes Care*. 2012;35(1):25-31.
37. Najem MY, Couturaud F, Lemarié CA. Cytokine and chemokine regulation of venous thromboembolism. *J Thromb Haemost*. 2020;18(5):1009-1019.
38. Christiansen SC, Naess IA, Cannegieter SC, Hammerstrøm J, Rosendaal FR, Reitsma PH. Inflammatory cytokines as risk factors for a first venous thrombosis: a prospective population-based study. *PLoS Med*. 2006;3(8):e334.
39. Ye RD, Sun L. Emerging functions of serum amyloid A in inflammation. *J Leukoc Biol*. 2015;98(6):923-929.
40. Sun L, Ye RD. Serum amyloid A1: structure, function and gene polymorphism. *Gene*. 2016;583(1):48-57.
41. Sack GH Jr. Serum amyloid A (SAA) proteins. *Subcell Biochem*. 2020;94:421-436.
42. Tannock LR, De Beer MC, Ji A, et al. Serum amyloid A3 is a high density lipoprotein-associated acute-phase protein. *J Lipid Res*. 2018;59(2):339-347.
43. Voudris KV, Chanin J, Feldman DN, Charitakis K. Novel inflammatory biomarkers in coronary artery disease: potential therapeutic approaches. *Curr Med Chem*. 2015;22(22):2680-2689.
44. Zhou J, Lu Y, Wang S, Chen K. Association between serum amyloid A levels and coronary heart disease: a systematic review and meta-analysis of 26 studies. *Inflamm Res*. 2020;69(4):331-345.
45. Wang S, Song R, Wang Z, Jing Z, Wang S, Ma J. S100A8/A9 in Inflammation. *Front Immunol*. 2018;9:1298.
46. Holzinger D, Roth J. Alarming consequences - autoinflammatory disease spectrum due to mutations in proline-serine-threonine phosphatase-interacting protein 1. *Curr Opin Rheumatol*. 2016;28(5):550-559.
47. Midwood KS, Chiquet M, Tucker RP, Orend G. Tenascin-C at a glance. *J Cell Sci*. 2016;129(23):4321-4327.
48. Flück M, Tunc-Civelek V, Chiquet M. Rapid and reciprocal regulation of tenascin-C and tenascin-Y expression by loading of skeletal muscle. *J Cell Sci*. 2000;113(Pt 20):3583-3591.
49. Li GH, Arora PD, Chen Y, McCulloch CA, Liu P. Multifunctional roles of gelsolin in health and diseases. *Med Res Rev*. 2012;32(5):999-1025.
50. Lee WM, Galbraith RM. The extracellular actin-scavenger system and actin toxicity. *N Engl J Med*. 1992;326(20):1335-1341.
51. Wakabayashi S, Koide T. Histidine-rich glycoprotein: a possible modulator of coagulation and fibrinolysis. *Semin Thromb Hemost*. 2011;37(4):389-394.
52. Priebsch KM, Kvensakul M, Poon IK, Hulett MD. Functional Regulation of the plasma protein histidine-rich glycoprotein by Zn(2+) in settings of tissue injury. *Biomolecules*. 2017;7(1):22.
53. Takase H, Tanaka M, Miyagawa S, Yamada T, Mukai T. Effect of amino acid variations in the central region of human serum amyloid A on the amyloidogenic properties. *Biochem Biophys Res Comm*. 2014;444(1):92-97.
54. Aguilera JJ, Zhang F, Beaudet JM, Linhardt RJ, Colón W. Divergent effect of glycosaminoglycans on the in vitro aggregation of serum amyloid A. *Biochimie*. 2014;104:70-80.
55. Erbel C, Tyka M, Helmes CM, et al. CXCL4-induced plaque macrophages can be specifically identified by co-expression of MMP7+S100A8+ in vitro and in vivo. *Innate Immun*. 2015;21(3):255-265.
56. Robinson MJ, Tessier P, Poulson R, Hogg N. The S100 family heterodimer, MRP-8/14, binds with high affinity to heparin and heparan sulfate glycosaminoglycans on endothelial cells. *J Biol Chem*. 2002;277(5):3658-3665.
57. Meiners S, Marone M, Rittenhouse JL, Geller HM. Regulation of astrocytic tenascin by basic fibroblast growth factor. *Dev Biol*. 1993;160(2):480-493.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Han B, Li C, Li H, et al. Discovery of plasma biomarkers with data-independent acquisition mass spectrometry and antibody microarray for diagnosis and risk stratification of pulmonary embolism. *J Thromb Haemost*. 2021;19:1738-1751. <https://doi.org/10.1111/jth.15324>