A study of the proteomic expression in patients with complicated parapneumonic pleural effusion

Xin Wu[1], Wei Yang[1], Xiao-Hua Gou[1], Xiao-Yun Xu[1], Na Lu[2], Shi-Ning Jian[1], Yu-Jie Han[1], Tong-Shuai Lv[1], Li-Zhu Luo[1]

[1]The Second Department of Respiration, Baoding No.1 Central Hospital, HeBei, China
[2]Department of Respiration, Lixian County Hospital, HeBei, China

Corresponding author:
Wei Yang
The Second Department
of Respiration
Baoding No. 1
Central Hospital
HeBei, China
E-mail: ywei_112@163.com

Abstract

Introduction: The present study aimed to investigate the differences in the proteomic expression between uncomplicated parapneumonic pleural effusion (UPPE) and complicated parapneumonic pleural effusion (CPPE).

Material and methods: There were 10 patients with UPPE and 10 patients with CPPE. These patients were combined due to the complication of pleural effusion and further divided into group A and group B. An LC-MS analysis was conducted with the extraction of high-abundance proteins, and proteins with 1.5-fold or higher difference multiples were identified as differential proteins. Then, gene ontology (GO) and KEGG analyses were conducted on the differential proteins between the groups.

Results: Compared with the UPPE group, there were 38 upregulated proteins and 29 downregulated proteins in the CPPE group. The GO analysis revealed that the CPPE group had enhanced expressions in monosaccharide biosynthesis, glucose catabolism, fructose-6-phosphate glycolysis, glucose-6-phosphate glycolysis, and NADH regeneration as well as reduced expressions in fibrinogen complexes, protein polymerization, and coagulation. Moreover, the KEGG analysis showed that the CPPE group had enhanced expressions in amino acid synthesis, the HIF-1 signalling pathway, and glycolysis/glyco-isogenesis and decreased expressions in platelet activation and complement activation.

Conclusions: In pleural effusion in patients with CPPE, there are enhanced expressions of proteins concerning glucose and amino acid metabolism, NADH regeneration, and HIF-1 signalling pathways together with decreased expressions of proteins concerning protein polymerization, blood coagulation, platelet activation, and complement activation.

Key words: uncomplicated parapneumonic pleural effusion, complicated parapneumonic pleural effusion, differential protein, GO analysis, KEGG analysis.

Introduction

Parapneumonic pleural effusion (PPE) is associated with a variety of infectious pulmonary diseases [1]. Currently, the prevalence of PPE is increasing in all age groups [2–4]. Approximately 40% of patients hospitalized with pneumonia have varying degrees of pleural effusion. Patients with PPE are more likely to have a longer duration of fever and a more serious condition than those with pneumonia alone. Moreover, such pa-

tients have a higher mortality and worse prognosis than those with only pneumonia [5]. Therefore, we should pay more attention to patients with PPE.

Since the completion of the Human Genome Project, the goal of research has shifted from genomic information to the study of proteins, which has led to a new research direction – proteomics. Proteomics is the study of the mechanisms of disease by analysing protein compositions, protein modification states, quantitative assays, differential expressions, qualitative identifications, and protein interactions, with the core techniques including isolation and identification of proteins, study of protein structure, and analysis and prediction of protein function. In the current study, we questioned whether there are differences in the energy metabolism and signal transduction in the pleural cavity in patients with PPE. Thus, we investigated the expressions of differential proteins in PPE by proteomic techniques and performed GO and KEGG analyses on the differential proteins.

Material and methods

Study subjects

In the present study, a total of 20 patients with PPE, who were hospitalized in the Department of Respiratory Medicine of Baoding First Central Hospital from February 2018 to August 2018, were included, comprising 10 patients with uncomplicated PPE (UPPE) and 10 patients with complicated PPE (CPPE). These patients were divided into Group A and Group B, respectively [5].

Inclusion criteria [6]

(1) New-onset cough, expectoration, and other respiratory symptoms or aggravation of an existing cough, expectoration, and purulent phlegm. (2) Fever (axillary temperature over 37.3°C). (3) Physical examination revealed signs of pulmonary consolidation with or without wet rales. (4) WBC > 10 × 10^9/l revealed by routine blood test. (5) Imaging examination: Chest CT showed flaky, patchy infiltrative opacities or interstitial changes together with pleural effusion (chest CT or thoracic ultrasonography). The included patients had to meet the above conditions, except for pulmonary vasculitis, pulmonary embolism, pulmonary insufficiency, non-infectious interstitial pulmonary disease, pulmonary oedema, pulmonary tumours, tuberculosis, and pulmonary eosinophilic infiltration.

Diagnostic criteria for UPPE and CPPE

Pleural fluid was collected and examined after each patient was admitted to the hospital. According to Light's criteria, those with phases 1–2 were enrolled in the UPPE group, and those with phases 3–5 were enrolled in the CPPE group.

Light's criteria: (1) Phase 1: Small amount of pleural effusion with an area of less than 1 cm on chest CT scan. (2) Phase 2: Chest CT scan of pleural effusion greater than 1 cm; pleural fluid with glucose > 2.24 mmol/l; pH > 7.2; LDH < 3 times the upper limit of the normal serum value; negative pleural fluid culture. (3) Phase 3: 7.00 < pH of the pleural effusion < 7.20; glucose in the pleural effusion > 2.24 mmol/l; LDH > 1000 U/l; negative pleural fluid culture. (4) Phase 4: Pleural effusion without coating and non-purulent in appearance; pH < 7.0, glucose in the pleural effusion < 2.24 mmol/l, or positive pleural fluid culture. (5) Phase 5: Multiple encapsulated effusions; pH < 7.0, glucose in the pleural effusion < 2.24 mmol/l, or positive pleural fluid culture. (6) Phase 6: Purulent pleural effusion with free or solitary package of the pleural fluid. (7) Phase 7: Multiple encapsulated purulent effusions, >5 mm increase in the thickness of the mural pleura.

Methods

According to Light's criteria, the patients were divided into groups A and B, with the former as the UPPE group and the latter as the CPPE group. There were 10 cases in each group.

Sample collection

In all enrolled patients, pleural puncture or chest tube drainage was performed. A total of 10 ml of the pleural fluid from the first sample was retained and centrifuged at 3000 rpm for 10 min; the supernatant was separated. Then, 40 μl of specimen from each patient in each group was combined to form the samples for group A and group B. The samples were stored at –80°C in a refrigerator.

Protein extraction and quantitative quality control

The proteins were extracted using the ProteoExtract Albumin/IgG Abundance Extraction Kit by Merck. (1) Preparation of the sample solution: 60 μl of sample was mixed with 540 μl of ProteoExtract Albumin/IgG Binding Buffer. (2) The upper cover of the Albumin/IgG Extraction Column was removed, and with the upside-down, the column storage solution was discarded and the lower cover of the Albumin/IgG Extraction Column removed. (3) The column was balanced by 850 μl of the ProteoExtract Albumin/IgG Binding Buffer, with the flowing solution discarded. (4) The sample solution in step 1 was passed through the column connected in step 2, and the outflowing solution was collected. (5) The sample solution prepared in step 1 was placed in the Albumin/

Xin Wu, Wei Yang, Xiao-Hua Gou, Xiao-Yun Xu, Na Lu, Shi-Ning Jian, Yu-Jie Han, Tong-Shuai Lv, Li-Zhu Luo

IgG Extraction Column, and the outflowing solution was collected. (6) Next, 600 μl of the Proteo-Extract Albumin/IgG Binding Buffer was added to clean the column, and the outflowing solution was collected. (7) The procedures in the previous step were repeated, and all outflowing solution was collected. (8) Then, 8 ml of TCA/acetone (1 : 4) was added to the protein extraction solution, which was mixed well and kept at 4°C for 1.5 h. (9) At 4°C, the solution was centrifuged at 6000 rpm for 40 min. (10) Next, 400 μl of ice-cold acetone was added to each sample 4 times to break up the precipitate. The solution was kept at –40°C for 0.5 h, 4°C, and was centrifuged at 15,000 rpm for 15 min. The supernatant was discarded. (11) The procedures in steps 3–4 were repeated, and the solution was precipitated at –40°C for 0.5 h. (12) At 4°C, the solution was centrifuged at 15,000 rpm for 15 min and then dried for 15 min. (13) Finally, the solution was dissolved in 60 μl of 8 M urea and blown evenly with a pipette.

Bradford protein quantification

(1) The standard curve was constructed, and (2) 20 μl of protein solution was prepared: 1 μl of the tested sample was mixed with 19 μl of water, and each sample was tested for 2 replicates. (3) Then, 200 μl of the Bradford working solution was added to each replicate and mixed evenly. (4) At room temperature in the dark, the replicates stood for 15 min. The purple gradient was observed and detected at the 595-nm wavelength by a microplate reader. (5) The total amount of protein in the tested sample was calculated from the standard curve.

SDS-PAGE electrophoresis

To start, (1) 20 μg of the protein sample 5 : 1 (v/v) was mixed with 5× loading buffer, bathed in boiling water for 5 min, and then centrifuged at 14,000 rpm for 10 min. The supernatant was taken to conduct 10% SDS-PAGE electrophoresis. (2) Electrophoresis conditions: constant flow of 14 mA, electrophoresis duration of 90 min. (3) Comas brilliant blue staining.

Trypsin proteolysis

(1) A total of 60 μg of the protein solution was taken and placed in a centrifuge tube with 5 μl of 1 M DTT solution. The solution was shaken,

mixed, and kept at 37°C for 1 h. Then, 20 μl of 1 M IAA solution was added, shaken, and mixed. The tube stood for 1 h in the dark at room temperature. (2) All samples to be tested were added to the ultra-filtration tube and centrifuged, and the collected solution was discarded. (3) Then, 100 μl of UA (pH 8.0, 8 M urea, 100 mM Tris-HCl) was added to the ultra-filtration tube and centrifuged. The collected solution was discarded and repeated twice, after which (4) 100 μl of 50 mM NH_4HCO_3 was added and centrifuged. The collected solution was discarded and repeated 3 times. (5) After changing the collection tube, trypsin was added to the ultra-filtration tube. The amount of trypsin was determined at the ratio of 50 : 1 between the protein and enzyme, and the solution was hydrolysed at 37°C for 12–16 h.

Liquid chromatography–mass spectrometry

A nanolitre-flow-rate, high-performance liquid chromatography (HPLC) system was adopted to separate the samples to be analysed. The chromatography column was balanced with 95% liquid A. Each sample was loaded into the pre-column of an Orbitrap Fusion Lumos mass spectrometer (database: uniprot-human_170221, Software: Proteome Discover) by the auto-sampler and separated by the analytical column. The samples were then separated by capillary HPLC and analysed.

Results

Results of protein quantification

The protein concentration, total protein amount, and electrophoretic amount in the CPPE group were all higher than those in the UPPE group, and the changes in their values were proportional to the OD absorbance values (Table I).

Results of SDS-PAGE electrophoresis

Compared with the UPPE group, the CPPE group had clearer bands around 10, 25, 60, 90, 100, 120, and 220 kDa after the high-abundance proteins were extracted, which was consistent with the expected results (Figure 1).

Results of liquid chromatography–mass spectrometry

The differential PPE proteins were investigated (Figure 2) and statistically analysed (Figures 3 and 4)

Table I. Results of protein quantification

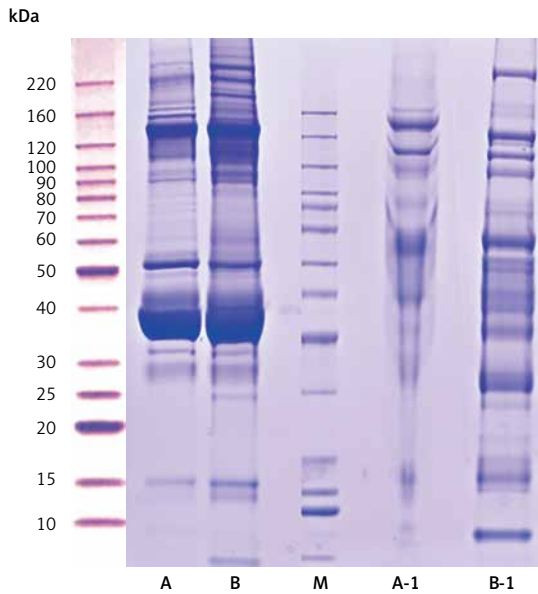| Group	OD absorbance	Protein concentration [μg/μl]	Total protein [μg]	Volume of electrophoresis [μl]
A	0.988	2.11	532	6.33
B	1.162	4.36	870	4.51

Figure 1. Results of protein electrophoresis. Note: A and B were the raw samples. A-1 and B-1 were the samples from which high-abundance proteins were extracted, by which the bands of high-abundance proteins were effectively removed. The total proteins in the 2 samples were effectively separated within the molecular weight range of 15–220 kDa, with no degradation of the proteins

Figure 2. Expressions of the differential proteins

Gene ontology (GO) analysis

A GO analysis is a series of classification systems based on gene function and is an important tool in the field of bioinformatics. This method has a complete set of control characters, the main purpose being to explain the key functions that eukaryotic genes and proteins perform in cells.

Gene ontology can be divided into 3 ontologies: (1) Cell components – A cell component is a structural object, mainly used to describe the sub-cellular structures (e.g. the nucleus) or protein production components (proteasome, ribosome, etc.). (2) Molecular functions – These are mainly used to describe the biological activities, such as catalytic reactions, that occur in cells. They can also represent the operations performed by specific objects without designating specific conditions such as when and where the operations occur. (3) Biological processes – These are events wherein many molecular functions are combined and coordinated. It may be difficult to strictly distinguish between biological processes and molecular functions, but, in general, biological processes require multiple different steps.

after being identified in the 2 uploading repetitions. The differential proteins were determined under the condition of 1.5 difference multiples. An FC value ≥ 1.5 and p-value < 0.05 signified upregulation; FC ≤ 0.67 and p < 0.05 signified downregulation; and 0.67 < FC < 1.5 or p > 0.05 indicated no significant change in the expression level. The proteins with upregulated expression are illustrated in Table II, and the proteins with downregulated expression are shown in Table III.
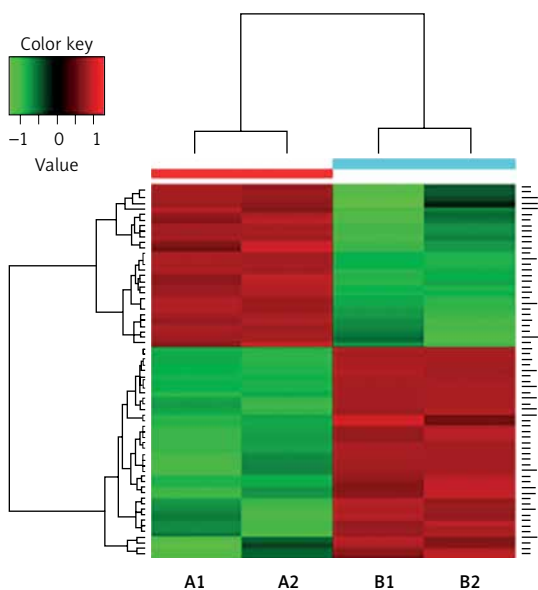

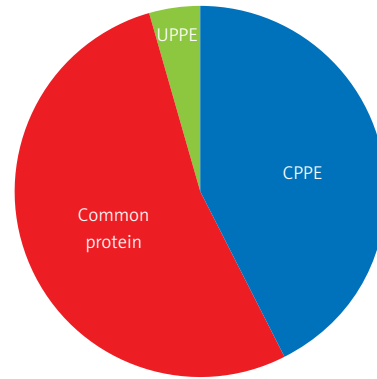**Figure 3.** Plots of the protein cluster analysis

Figure 4. Volcano diagram of the differential proteins between the groups

Xin Wu, Wei Yang, Xiao-Hua Gou, Xiao-Yun Xu, Na Lu, Shi-Ning Jian, Yu-Jie Han, Tong-Shuai Lv, Li-Zhu Luo

Table II. Proteins with upregulated expressions

Protein sequence	Description	Protein existence	FC value	*P*-value
D9YZU5	Beta-globin	3	138.97	< 0.01
P06733	Alpha-enolase	1	102.13	< 0.01
P04406	Glyceraldehyde-3-phosphate dehydrogenase	1	102.13	< 0.01
P29401	Transketolase	1	91.19	< 0.01
P00558	Phosphoglycerate kinase 1	1	80.84	< 0.01
A0A1K0GXZ1	Globin C1	4	62.36	< 0.01
P04075	Fructose-bisphosphate aldolase A	1	59.33	< 0.01
V9HWB8	Pyruvate kinase	1	57.42	< 0.01
A0A140VK27	Leukotriene A(4) hydrolase	2	35.63	< 0.01
P07737	Profilin-1	1	35.29	< 0.01
P63261	Actin, cytoplasmic 2	1	21.03	< 0.01
Q8NF20	FLJ00382 protein (Fragment)	2	17.54	< 0.01
S6BGD4	IgG H chain	2	12.31	< 0.01
Q0ZCH1	Immunoglobulin heavy chain variable region	4	12.29	< 0.01
A0A0K0K1H8	Epididymis secretory sperm binding protein Li 71p	2	10.06	< 0.01
A0N7I9	F5-20 (Fragment)	2	9.85	< 0.01
A8K3K1	cDNA FLJ78096, highly similar to Homo sapiens actin, alpha, cardiac muscle (ACTC), mRNA	2	8.63	< 0.01
A0A0X9TD88	GCT-A3 heavy chain variable region (Fragment)	2	8.57	< 0.01
P98160	Basement membrane-specific heparan sulphate proteoglycan core protein	1	7.87	0.03
Q6N030	Uncharacterized protein	1	7.27	< 0.01
Q0ZCF6	Immunoglobulin heavy chain variable region	4	5.39	< 0.01
A2J1N6	Rheumatoid factor RF-ET9	2	4.99	0.03
A0A120HG44	MS-E1 heavy chain variable region	2	4.70	0.02
O75636	Ficolin-3	1	4.57	< 0.01
A0A0X9T7V9	GCT-A4 light chain variable region	2	3.84	0.03
A0A068LKQ2	Ig heavy chain variable region	2	3.71	< 0.01
P02741	C-reactive protein	1	3.68	< 0.01
B2RBZ5	cDNA, FLJ95778	1	3.41	< 0.01
P18428	Lipopolysaccharide-binding protein	1	3.32	< 0.01
P23381	Tryptophan--tRNA ligase, cytoplasmic	1	3.01	0.04
B4E1B2	cDNA FLJ53691	2	2.89	< 0.01
Q4LE33	TNC variant protein	2	2.56	0.04
P02760	Protein AMBP	1	2.51	0.01
A0A109PP82	MS-C2 light chain variable region	2	2.50	0.03
A0A0X9UWM4	IBM-A3 heavy chain variable region	2	2.26	< 0.01
E9KL23	Epididymis secretory sperm binding protein Li 44a	2	2.21	< 0.01
C9JF17	Apolipoprotein D	1	2.16	0.02
A0A075B6L1	Ig lambda-7 chain C region	1	2.09	0.02

In the present study, a GO analysis was conducted on the identified proteins. The results showed that the upregulated differential proteins were mostly involved in the biosynthesis of monosaccharides, glucose catabolism, glycolysis of fructose-6-phosphate, glycolysis of glucose-6-phosphate, and NADH regeneration (Figure 5), while the downregulated differential proteins mostly participated in the processes of fibrinogen complexes, protein polymerization, and coagulation (Figure 6).

Table III. Proteins with downregulated expressions

Protein sequence	Description	Protein existence	FC value	*P*-value
A0A140VKF3	Testis tissue sperm-binding protein Li 70n	2	0.50	0.02
A0A0X9V9B3	MS-F1 light chain variable region	2	0.50	0.02
B2R6V9	cDNA, FLJ93141, highly similar to Homo sapiens coagulation factor XIII	2	0.50	0.02
A6XGL1	Transthyretin	2	0.49	0.03
B0YIW2	Apolipoprotein C-III	1	0.49	< 0.01
D3DNU8	Kininogen 1, isoform CRA_a	4	0.48	< 0.01
A2JA19	Anti-mucin1 light chain variable region	2	0.42	0.02
A0A0S2Z3Y1	Lectin galactoside-binding soluble 3 binding protein isoform 1	2	0.42	<0.01
P04430	Immunoglobulin kappa variable 1-16	1	0.39	< 0.01
Q6ZW64	cDNA FLJ41552 fis, clone COLON2004478, highly similar to Protein Tro alpha1 H	2	0.38	< 0.01
A0A096LPE2	Protein SAA2-SAA4	4	0.38	< 0.01
Q9UNU2	Complement protein C4B frameshift mutant	4	0.38	< 0.01
A2NWW3	VH-3 family (VH26)D/J protein	4	0.35	< 0.01
B2R773	cDNA, FLJ93312, highly similar to Homo sapiens adipose most abundant gene transcript 1	2	0.35	< 0.01
Q96PD5	N-acetylmuramoyl-L-alanine amidase	1	0.35	0.01
A0A024R930	Proteoglycan 4, isoform CRA_a OS=Homo sapiens	4	0.29	0.03
V9H1C1	Gelsolin exon 4	4	0.27	< 0.01
A0A109PS32	MS-A1 heavy chain variable region	2	0.27	0.02
B2R815	cDNA, FLJ93695, highly similar to Homo sapiens serpin peptidase inhibitor	2	0.26	0.03
B1AHL2	Fibulin-1	1	0.24	< 0.01
P01706	Immunoglobulin lambda variable 2-11	1	0.23	< 0.01
P02671	Fibrinogen alpha chain	1	0.21	< 0.01
D3DP16	Fibrinogen gamma chain	4	0.21	0.01
V9HVY1	Epididymis secretory sperm binding protein Li 78p	2	0.20	< 0.01
P23142	Fibulin-1	1	0.20	< 0.01
H7C0L5	Inter-alpha-trypsin inhibitor heavy chain H4	1	0.19	< 0.01
P02679	Fibrinogen gamma chain	1	0.16	< 0.01
B2R950	cDNA, FLJ94213, highly similar to Homo sapiens pregnancy-zone protein	2	0.10	< 0.01
A8KAJ3	cDNA FLJ77823, highly similar to Homo sapiens EGF-containing fibulin-like extracellular matrix protein 1	2	0.10	< 0.01

KEGG analysis

Organisms are composed of different kinds of proteins that perform their functions in a coordinated and orderly manner. The purpose of a pathway analysis is to better investigate the roles of proteins. One of the public databases about pathways is KEGG. Through pathway analysis, the functions of proteins in the process of biological activities are determined with the main participation in the metabolic pathways and signal transductions.

One method of pathway significant enrichment analysis is to use KEGG Pathway as the research unit and apply hyper-geometric testing to determine the pathways significantly enriched by differential proteins. Using this research method, the most important signal transduction and biochemical metabolic pathways involved with differential proteins can be determined. Herein, the KEGG function analysis was adopted to investigate the most prominent biochemical metabolic pathways and related signal transductions of the differential proteins in UPPE and CPPE. The results showed that CPPE had increased protein expressions in the HIF-1 signalling pathway and glycolysis/gluconeogenesis (Figure 7) but decreased expressions
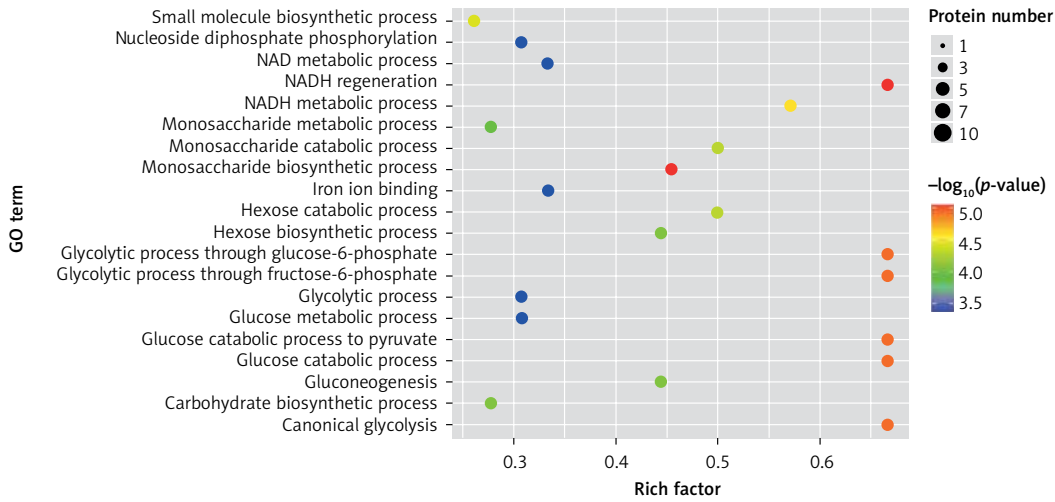
Xin Wu, Wei Yang, Xiao-Hua Gou, Xiao-Yun Xu, Na Lu, Shi-Ning Jian, Yu-Jie Han, Tong-Shuai Lv, Li-Zhu Luo

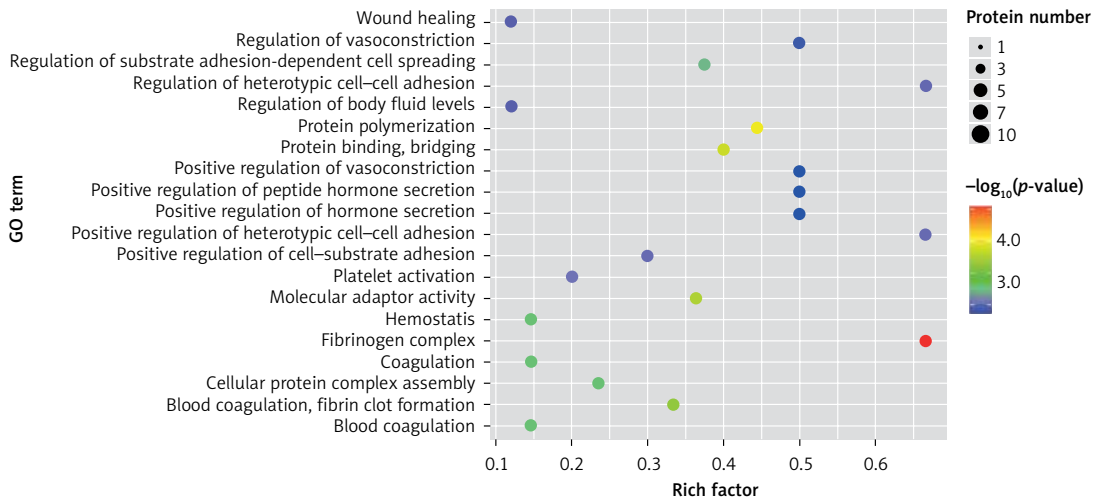**Figure 5.** Plots of proteins with upregulations

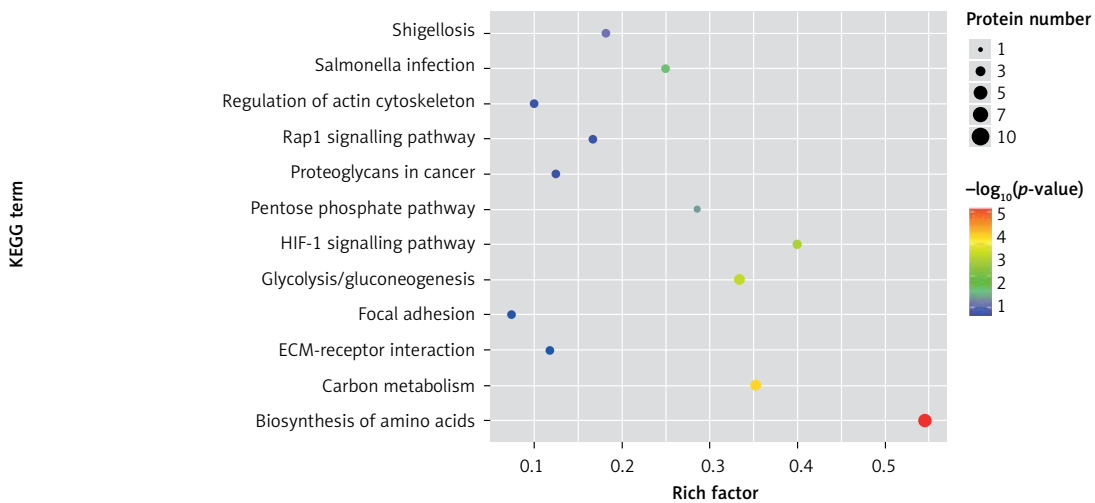**Figure 6.** Plots of proteins with downregulations

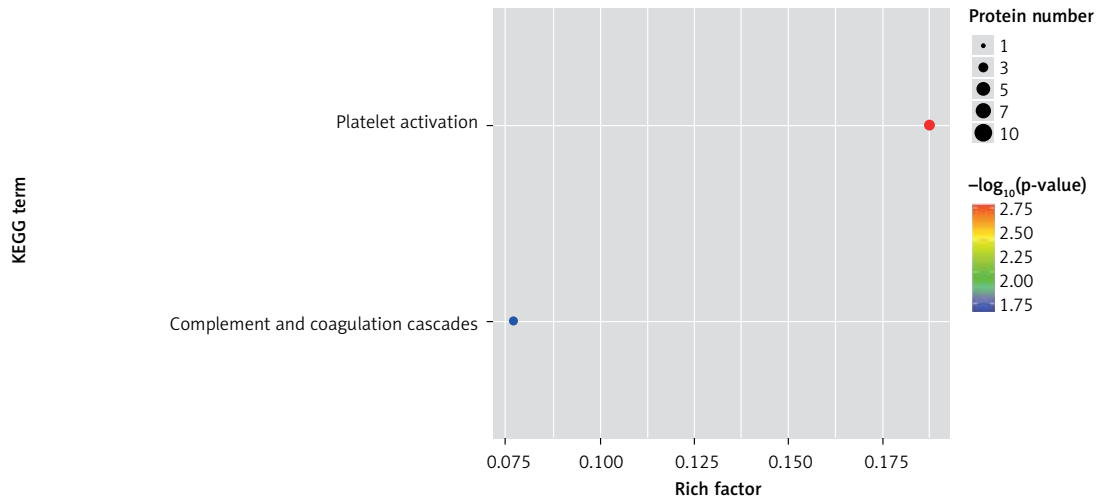**Figure 7.** Plots of differential proteins with upregulations

Figure 8. Plots of differential proteins with downregulations

in platelet activation and complement activation (Figure 8).

Discussion

Human genome sequencing has been completed, and we have used genomics technology to conduct a large number of related studies on genes and diseases, bringing about more opportunities for the development of modern medicine. Proteomics is an emerging research field in the post-gene era. It is well known that although proteins are encoded by genes, they are affected by the internal and external environments and types of cells and are constantly changing. Moreover, due to RNA splicing, editing, or post-transcriptional modification, a single gene may be related to multiple proteins. Therefore, the relationship between genes and proteins is not completely linear; rather, they are 2 fields that complement each other and are closely correlated.

Thus far, the research fields of proteomics have mainly focused on the following aspects: (1) It is proteins, not genes, that determine the functions of cells. Proteomics can investigate the micro-features of proteins. While certain body fluids and plasma may not express much DNA or RNA, proteins may be present in large quantities, so proteomics may have advantages over genomics. (2) Protein isoforms and post-transcriptional modifications cannot be identified through genomic studies, but important protein information can be obtained through proteomic studies. Thus, changes in protein composition can be compared between pathological and normal conditions to find markers for the diagnosis of disease. (3) Regarding the study of protein–protein interactions, with the invention of biomass spectrometry, advances in medical informatics, and the establishment of relevant databases, we are now able to complete high-throughput proteomics studies [7, 8].

In recent years, rapid advances in molecular biology and bioinformatics have made it possible to identify and quantify any kind of trace protein, thus enabling efficient analysis of low-abundance proteins and post-translational modifications [9, 10]. Malignant tumours comprise the most extensive field of proteomics research. Researchers have begun to apply this technology to investigate tumour markers and have successfully discovered new protein markers for the diagnosis of malignant tumours [11]. In a previous study, leaky pleural effusions were used instead of normal pleural effusions as a control to study the expression of proteins expressed in malignant pleural effusions, and 7 differential proteins were found [12]. In respiratory diseases, proteomics is also the most widely used research method concerning malignant lung cancer, and several differentially expressed proteins have been found [13–16]. Currently, the research on CPPE is mostly focused on inflammatory cytokines such as tumour necrosis factor-α, IL-8, and IL-1 as well as neutrophil elastase, myeloperoxidase, metalloproteinase, and other enzymes [17–22], and there are a few studies on proteomics concerning pleural effusion. At present, the development of high-throughput proteomic technology provides technical support for the proteomic analysis of body fluids, which may help in the discovery of new biomarkers [23–25]. Moreover, emerging computer and bioinformatics technologies can analyse and interpret a large number of experimental proteomic data. These technological advances make it possible to investigate the proteomics of pleural effusion.

There are several high-abundance proteins in exudative pleural effusion, such as albumin, immunoglobulin, and α-antitrypsin. The presence of these high-abundance proteins may affect the expression of low-abundance proteins; thus, the presence of the former is not conducive to the de-

Xin Wu, Wei Yang, Xiao-Hua Gou, Xiao-Yun Xu, Na Lu, Shi-Ning Jian, Yu-Jie Han, Tong-Shuai Lv, Li-Zhu Luo

tection of the latter, especially with the obvious interference of albumin [12]. Therefore, to ensure the reliability of an experiment, it is necessary to extract the high-abundance proteins [26]. However, this process also has a certain degree of negative impact because a small number of low-abundance proteins are combined with the high-abundance proteins, causing the loss of both protein types.

The application of proteomic technology in the field of PPE pathogenesis is still in the exploratory stage. The accumulated experience is not yet rich, and many research areas need to be explored. Currently, several researchers have begun to apply proteomic technology in the research of interstitial pulmonary diseases and pulmonary infectious diseases, and related research on proteomics has also begun in the study of ARDS [27, 28]. Therefore, we hope that through the study of proteomics, we will gain better understanding of the protein expression pattern, protein biological activity, metabolic pattern, and transmission pathway of PPE to understand the disease in depth. At present, the proteomics of pleural effusion has not been studied in depth, but more and more scholars have paid attention to it. Chiu *et al.* found that the level of glycolytic glucose in pleural effusion of patients with complex pleural effusion was significantly reduced, which was similar to the results of our study, and they also found that glucose and 3-hydroxybutyric acid could be used as markers to distinguish different metabolic pathways [29]. Based on proteomic analysis of pleural effusion near pneumonia, Hsueh found that dipeptides were closely related to the expression of related inflammatory factors in the pleural cavity, suggesting that dipeptides may be involved in the inflammatory response in the pleural cavity through immune regulation and affect the course of the disease [30]. By proteomic method, Chiu *et al.* found that in patients with pleural effusion near pneumonia, anaerobic glycolysis metabolism was enhanced, lactic acid and butyric acid were increased, and the increased expression of hypoxanthine may be related to the formation of fibrin in the pleural cavity [31]. In addition, the pleural effusion around simple pneumonia and complex pneumonia can be distinguished by protein spectrum analysis, and the people who need active treatment or even surgical treatment can be predicted [32, 33]. Through proteomics, Wu *et al.* found that the four markers of BPI, NGAL, AZU1, and calcium-guard protein are most significantly increased, among which BPI has the highest value in diagnosing CPPE, if combining with BPI and LDH, then the sensitivity and specificity to the diagnosis of CPPE will be increased to 100% and 91.4%, so the study of thoracic fluid protein

histology of pneumonia may find more novel biomarkers [5].

The expressions of differential proteins were investigated in the present study. To increase the study's representativeness, the pleural effusion samples of 10 patients in the study group and 10 in the control group were combined to form group A and group B, respectively. A similar method has been reported in related studies. Using label-free proteomic technology and bioinformatics analysis, it was determined that there were 257 proteins in total from the 2 groups, including 200 specific to the CPPE group, 21 specific to the UPPE group, and 38 upregulated and 29 downregulated in the CPPE group. A GO analysis showed that the upregulated differential proteins participated in the biosynthesis of monosaccharides, glucose catabolism, glycolysis of fructose-6-phosphate, glycolysis of glucose-6-phosphate, and NADH regeneration. The downregulated differential proteins were concentrated in the processes of fibrinogen complexes, protein polymerization, and coagulation. These results suggest that the glucose metabolism in the chest cavity of the CPPE patients was active and that NADH was also produced in glycolysis and the citric acid cycle in cell respiration. It is an important coenzyme for electron transfer in the body as well as an important part of the energy production chain in mitochondria control markers. Increased levels of the two often indicate metabolic imbalance. The downregulated differential proteins mostly participated in the processes of fibrinogen complexes, protein polymerization, and blood coagulation, suggesting that the coagulation process in the pleural cavity was enhanced and the formation of fibrin increased, which might be correlated with the infection and inflammation in the cavity. It was revealed in the KEGG analysis that the CPPE group had increased protein expressions in amino acid synthesis, the HIF-1 signalling pathway, and glycolysis/gluconeogenesis as well as weakened protein expressions in platelet activation and complement activation. These results were consistent with the results of the GO analysis. We considered that it might be correlated with intrapleural infection, enhanced bacterial metabolism, hypoxic stimulation, and imbalances in the blood coagulation system and immune system. Overall, the results suggested that proteomics could be used to investigate the pathogenesis of PPE and that new biomarkers that can distinguish between UPPE and CPPE might be discovered through subsequent verification experiments.

This study also has some limitations, such as a small sample size and insufficient representativeness. And into the group of patients with standard need to be further refinement, such as pneumonia side pleural effusion patients of different ages there may be some difference between pro-

tein content [33], so in a follow-up study, we need to further increase the sample size, at the same time reduce the influence of some confounding factors, and look forward to the future we can have better results.

In conclusion, proteomics is a new technology of the post-gene era, with high sensitivity and accuracy. However, it is still in its infancy in the research of pulmonary infectious diseases, and related research is not yet mature. In this study, 38 upregulated proteins and 29 downregulated proteins were found in the pleural effusions of patients with PPE. Moreover, it was found that glucose and amino acid metabolism, NADH regeneration, and the HIF-1 signalling pathway in the effusions of patients with CPPE were enhanced, while protein polymerization, blood coagulation, platelet activation, and complement activation were weakened. These results suggest that there were obvious signs of infection in the pleural cavity in the CPPE group. The process of infection was accompanied by changes such as hypoxia and abnormal energy metabolism in addition to obvious inflammation and inflammatory disorders in the pleural cavity.

Conflict of interest

The authors declare no conflict of interest.

References

1. Light RW. Parapneumonic effusions and empyema. Proc Am Thorac Soc 2006; 3: 75-80.
2. Sahn SA. Diagnosis and management of parapneumonic effusions and empyema. Clin Infect Dis 2007; 45: 1480-6.
3. Finley C, Clifton J, Fitzgerald JM, Yee J. Empyema: an increasing concern in Canada. Can Respir J 2008; 15: 85-9.
4. Farjah F, Symons RG, Krishnadasan B, Wood DE, Flum DR. Management of pleural space infections: a population-based analysis. J Thorac Cardiovasc Surg 2007; 133: 346-51.
5. Wu KA, Wu CC, Chen CD, et al. Proteome profiling reveals novel biomarkers to identify complicated parapneumonic effusions. Sci Rep 2017; 7: 4026.
6. Cao B, Huang Y, She DY, et al. Diagnosis and treatment of community-acquired pneumonia in adults: 2016 clinical practice guidelines by the Chinese Thoracic Society, Chinese Medical Association. Clin Respir J 2018; 12: 1320-60.
7. Nyman TA. The role of mass spectrometry in proteome studies. Biomol Eng 2001; 18: 221-7.
8. Bakhtiar R, Nelson RW. Electrospray ionization and matrix-assisted laser desorption ionization mass spectrometry. Emerging technologies in biomedical sciences. Biochem Pharmacol 2000; 59: 891-905.
9. Starita-Geribaldi M, Roux F, Garin J, Chevallier D, Fénichel P, Pointis G. Development of narrow immobilized pH gradients covering one pH unit for human seminal plasma proteomic analysis. Proteomics 2003; 3: 1611-9.
10. Patton WF. Detection technologies in proteome analysis. J Chromatogr B Analyt Technol Biomed Life Sci 2002; 771: 3-31.
11. Rodland KD. Mass spectrometry and biomarker development. Dis Markers 2004; 20: 129-30.
12. Hsieh WY, Chen MW, Ho HT, You TM, Lu YT. Identification of differentially expressed proteins in human malignant pleural effusions. Eur Respir J 2006; 28: 1178-85.
13. Yanagisawa K, Shyr Y, Xu BJ, et al. Proteomic patterns of tumour subsets in non-small-cell lung cancer. Lancet 2003; 362: 433-9.
14. Li C, Chen Z, Xiao Z, et al. Comparative proteomics analysis of human lung squamous carcinoma. Biochem Biophys Res Commun 2003; 309: 253-60.
15. Gharib TG, Chen G, Wang H, et al. Proteomic analysis of cytokeratin isoforms uncovers association with survival in lung adenocarcinoma. Neoplasia 2002; 4: 440-8.
16. Chen G, Gharib TG, Huang CC, et al. Proteomic analysis of lung adenocarcinoma: identification of a highly expressed set of proteins in tumors. Clin Cancer Res 2002; 8: 2298-305.
17. Alegre J, Jufresa J, Segura R, et al. Pleural-fluid myeloperoxidase in complicated and noncomplicated parapneumonic pleural effusions. Eur Respir J 2002; 19: 320-5.
18. Alemán C, Alegre J, Segura RM, et al. Polymorphonuclear elastase in the early diagnosis of complicated pyogenic pleural effusions. Respiration 2003; 70: 462-7.
19. Porcel JM, Vives M, Esquerda A. Tumor necrosis factor-alpha in pleural fluid: a marker of complicated parapneumonic effusions. Chest 2004; 125: 160-4.
20. Iglesias D, Alegre J, Alemán C, et al. Metalloproteinases and tissue inhibitors of metalloproteinases in exudative pleural effusions. Eur Respir J 2005; 25: 104-9.
21. Porcel JM, Vives M, Cao G, et al. Biomarkers of infection for the differential diagnosis of pleural effusions. Eur Respir J 2009; 34: 1383-9.
22. Marchi E, Vargas FS, Acencio MM, et al. Proinflammatory and antiinflammatory cytokine levels in complicated and noncomplicated parapneumonic pleural effusions. Chest 2012; 141: 183-9.
23. Yu CJ, Wang CL, Wang CI, et al. Comprehensive proteome analysis of malignant pleural effusion for lung cancer biomarker discovery by using multidimensional protein identification technology. J Proteome Res 2011; 10: 4671-82.
24. Wu CC, Chu HW, Hsu CW, Chang KP, Liu HP. Saliva proteome profiling reveals potential salivary biomarkers for detection of oral cavity squamous cell carcinoma. Proteomics 2015; 15: 3394-404.
25. Liu PJ, Chen CD, Wang CL, et al. In-depth proteomic analysis of six types of exudative pleural effusions for nonsmall cell lung cancer biomarker discovery. Mol Cell Proteomics 2015; 14: 917-32.
26. Bhargava M, Higgins L, Wendt CH, Ingbar DH. Application of clinical proteomics in acute respiratory distress syndrome. Clin Transl Med 2014; 3: 34.
27. Kan M, Shumyatcher M, Himes BE. Using omics approaches to understand pulmonary diseases. Respir Res 2017; 18: 149.
28. Bhargava M, Becker TL, Viken KJ, et al. Proteomic profiles in acute respiratory distress syndrome differentiates survivors from non-survivors. PLoS One 2014; 9: e109713.
29. Chiu CY, Lin G, Cheng ML, et al. Metabolomic profiling of infectious parapneumonic effusions reveals biomarkers for guiding management of children with Streptococcus pneumoniae pneumonia. Sci Rep 2016; 6: 24930.
30. Hsueh PC, Wu KA, Yang CY, et al. Metabolomic profiling of parapneumonic effusion reveals a regulatory role of

Xin Wu, Wei Yang, Xiao-Hua Gou, Xiao-Yun Xu, Na Lu, Shi-Ning Jian, Yu-Jie Han, Tong-Shuai Lv, Li-Zhu Luo

dipeptides in interleukin-8 production in neutrophil-like cells. Anal Chim Acta 2020; 1128: 238-50.

31. Chiu CY, Hsieh SY, Wong KS, Lai SH, Chen JK, Huang JL. The value of total protein in guiding management of infectious parapneumonic effusion by using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. J Microbiol Immunol Infect 2015; 48: 483-9.

32. Chiu CY, Cheng ML, Wong KS, et al. Metabolomics reveals anaerobic bacterial fermentation and hypoxanthine accumulation for fibrinous pleural effusions in children with pneumonia. J Proteome Res 2019; 18: 1248-54.

33. Zarogiannis SG, Tsilioni I, Hatzoglou C, Molyvdas PA, Gourgoulianis KI. Pleural fluid protein is inversely correlated with age in uncomplicated parapneumonic pleural effusions. Clin Biochem 2013; 46: 378-80.