

RESEARCH

Open Access



Integrative multi-omics analysis unravels the host response landscape and reveals a serum protein panel for early prognosis prediction for ARDS

Mengna Lin^{1,2†}, Feixiang Xu^{2†}, Jian Sun², Jianfeng Song³, Yao Shen⁴, Su Lu², Hailin Ding², Lulu Lan², Chen Chen², Wen Ma⁵, Xueling Wu^{6*}, Zhenju Song^{1,2,7*} and Weibing Wang^{1,5*}

Abstract

Background The multidimensional biological mechanisms underpinning acute respiratory distress syndrome (ARDS) continue to be elucidated, and early biomarkers for predicting ARDS prognosis are yet to be identified.

Methods We conducted a multicenter observational study, profiling the 4D-DIA proteomics and global metabolomics of serum samples collected from patients at the initial stage of ARDS, alongside samples from both disease control and healthy control groups. We identified 28-day prognosis biomarkers of ARDS in the discovery cohort using the LASSO method, fold change analysis, and the Boruta algorithm. The candidate biomarkers were validated through parallel reaction monitoring (PRM) targeted mass spectrometry in an external validation cohort. Machine learning models were applied to explore the biomarkers of ARDS prognosis.

Results In the discovery cohort, comprising 130 adult ARDS patients (mean age 72.5, 74.6% male), 33 disease controls, and 33 healthy controls, distinct proteomic and metabolic signatures were identified to differentiate ARDS from both control groups. Pathway analysis highlighted the upregulated sphingolipid signaling pathway as a key contributor to the pathological mechanisms underlying ARDS. MAP2K1 emerged as the hub protein, facilitating interactions with various biological functions within this pathway. Additionally, the metabolite sphingosine 1-phosphate (S1P) was closely associated with ARDS and its prognosis. Our research further highlights essential pathways contributing to the deceased ARDS, such as the downregulation of hematopoietic cell lineage and calcium signaling pathways, contrasted with the upregulation of the unfolded protein response and glycolysis. In particular, GAPDH and ENO1, critical enzymes in glycolysis, showed the highest interaction degree in the protein–protein interaction network of ARDS. In the discovery cohort, a panel of 36 proteins was identified as candidate biomarkers, with 8

[†]Mengna Lin, Feixiang Xu have contributed equally.

*Correspondence:

Xueling Wu

wuxueling76@126.com

Zhenju Song

song.zhenju@zs-hospital.sh.cn

Weibing Wang

wwb@fudan.edu.cn

Full list of author information is available at the end of the article



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

proteins (VCAM1, LDHB, MSN, FLG2, TAGLN2, LMNA, MBL2, and LBP) demonstrating significant consistency in an independent validation cohort of 183 patients (mean age 72.6 years, 73.2% male), confirmed by PRM assay. The protein-based model exhibited superior predictive accuracy compared to the clinical model in both the discovery cohort (AUC: 0.893 vs. 0.784; Delong test, $P < 0.001$) and the validation cohort (AUC: 0.802 vs. 0.738; Delong test, $P = 0.008$).

Interpretation Our multi-omics study demonstrated the potential biological mechanism and therapy targets in ARDS. This study unveiled several novel predictive biomarkers and established a validated prediction model for the poor prognosis of ARDS, offering valuable insights into the prognosis of individuals with ARDS.

Keywords Proteomics, Metabolomics, Multi-omics, ARDS, Machine learning, Prognosis model

Introduction

ARDS represents a prevalent manifestation of critical illness, arising from severe infections, major injuries, or the inhalation of harmful irritants [1]. Research indicates that ARDS affects approximately 10% of patients in intensive care units (ICUs), with mortality reaching up to 46% and even 70% during the coronavirus disease 2019 (COVID-19) pandemic [2]. Consequently, there is a need to seek effective treatments and precise prognostic methods to enhance ARDS patient outcomes, a need more intensified by the COVID-19 crisis. However, predicting ARDS prognosis in the early days after ARDS diagnosis is challenging due to the high variability in the underlying mechanisms of ARDS. The systemic response alterations observed in ARDS can be attributed to various factors, including pathogen and injury exposure, genetic susceptibility, and immune responses [3]. These factors may influence protein activity and the downstream metabolites that drive disease progression [4–6]. Therefore, it is crucial to investigate the potential association between host-derived proteins and metabolites circulating in the bloodstream with the pathogenesis and advancement of ARDS.

The integration of multi-omics approaches, particularly the combination of proteomics and metabolomics, elucidates the interactions across different biological system layers, as demonstrated in studies on cardiomyopathy [7], non-small cell lung cancer [8], hepatitis C infection [9], etc. Previous studies have utilized serum proteins or metabolites to investigate infectious diseases. For example, one study identified proteins capable of accurately distinguishing and predicting COVID-19 outcomes [5]. Jacob *et al.* used proteomics and metabolomics to identify early predictive and pathogenic signatures of *Staphylococcus aureus* bacteremia [10]. Yi Wang and his colleagues investigated new diagnosis biomarkers and potential mechanisms in pediatric severe community-acquired pneumonia using proteomics combined with metabolomics [11]. However, the interplay between proteomic and metabolic profiles in the progression of adult ARDS and their collective role from a holistic perspective remains underexplored. Moreover, blood biomarkers

have gained significant interest in ARDS investigations in recent years [12], showing promise in enhancing diagnosis, prognostication, and management strategies for ARDS, either independently or in conjunction with physiological parameters.

The primary objective of this study is to utilize proteomics and metabolomics to investigate the host response in ARDS within a discovery cohort comprising ARDS onset, disease, and healthy controls. This approach enabled a comprehensive profiling of biological characteristics and potential mechanisms associated with ARDS. The second objective is to screen the candidate biomarkers of ARDS prognosis based on proteomics and machine learning methods. Then, we employed the Parallel Reaction Monitoring (PRM) method to validate the prognosis biomarkers at ARDS onset in an independent cohort. The multi-omic analysis generated in this study provided a global overview of the molecular changes, which may provide useful insight into the therapy and prognosis of ARDS.

Material and methods

Ethical approval

The study involving human participants was reviewed and approved by the Ethical Committee of Zhongshan Hospital, Fudan University, China (Ethical approval number: B2023-029R), Renji Hospital, Shanghai Jiaotong University School of Medicine, China (No. LY2023-096-B), Minhang Hospital, Fudan University, China (No. 2022-041-01K), and Pudong Hospital, Fudan University, China (No. WZ-22). Written informed consent to participate in this study was provided by the participants or their relatives.

Collections and preparation of clinical specimen

Discovery cohort

Single-center observational study. ARDS patients were recruited from the ICU of Zhongshan Hospital between December 2022 and September 2023. The ARDS cohort included patients aged over 18 years who were admitted to the ICU and clinically diagnosed with ARDS according

to the Berlin definition [13]. Serum samples were collected within 48 hours after diagnosis. Patients whose diagnosis of ARDS was agreed upon by at least two out of three clinical experts were included. A comprehensive analysis of clinical data was performed, utilizing patient medical records to gather information on routine blood tests, liver function tests (including alanine aminotransferase [ALT] and aspartate aminotransferase [AST]), total and conjugated bilirubin (STB and CB), renal function indicators (blood urea nitrogen [BUN] and creatinine [Cr]), albumin [Alb], C-reactive protein [CRP], procalcitonin [PCT], and coagulation profiles (fibrinogen [Fg], prothrombin time [PT], activated partial thromboplastin time [APTT], and D-dimer). The Sepsis-related Organ Failure Assessment (SOFA) score and PaO₂/FiO₂ ratio (P/F ratio) were evaluated concurrently. Comorbidities were also collected. For comparative analysis, disease controls (DC) and healthy controls (HC) were recruited and matched with the ARDS cohort based on age and sex. The inclusion criteria for the DC group were as follows: patients aged > 18 years, patients admitted to the ICU and had risk factors (sepsis, pneumonia, etc.) of ARDS, serum samples will be collected within the first 48 hours of ICU admission. We finally chose who did not progress to ARDS during the hospital stay. Patients who transferred to the other hospital were excluded. The HC group was derived from individuals registered at the Health Check Center of Zhongshan Hospital. HC controls were selected based on the absence of acute diseases such as infection and the lack of significant abnormalities on chest radiography. The control groups were matched for age and sex with the ARDS patients. The details for inclusion and exclusion for ARDS, DC, and HC groups are in the Supporting Information (Fig. S1).

Validation cohort

Multi-center validation. Validation samples were from the ICUs of the other three hospitals. Serum samples from patients diagnosed with ARDS were collected within 48 hours after the onset of ARDS. All venous blood was collected from participants and processed within 8 hours to isolate serum. The serum was separated by centrifugation at 300 × g for 10 min and stored at –80 °C until testing.

Proteomics sequencing and data preprocessing

Proteins were enriched from serum using magnetic nanomaterials [14]. The mass spectrometer was operated in data-independent acquisition (DIA) mode. The DIA raw data were processed using Spectronaut Pulsar 17.5 software (Biognosys) against the Uniprot-Homo sapiens-9606-2023.2.1.fasta database. The details of proteomics

sequencing and data preprocessing were in Supporting Information.

Metabolomics sequencing and data preprocessing

A Dionex Ultimate 3000 RS UHPLC fitted with Q-Exactive plus quadrupole-Orbitrap mass spectrometer equipped with heated electrospray ionization (ESI) source (Thermo Fisher Scientific, Waltham, MA, USA) was used to analyze the metabolic profiling in both ESI positive and ESI negative ion modes. The original liquid chromatography coupled with high-resolution mass spectrometry (LC-MS) data was processed by software Progenesis QI V2.3 (Nonlinear, Dynamics, Newcastle, UK) for baseline filtering, peak identification, integral, retention time correction, peak alignment, and normalization. The main parameters of 5 ppm precursor tolerance, 10 ppm product tolerance, and 5% production threshold were applied. Compound identification was based on the precise mass-to-charge ratio (M/z), secondary fragments, and isotopic distribution using The Human Metabolome Database (HMDB), Lipidmaps (V2.3), Metlin, EMDB, PMDB, and self-built databases to do qualitative analysis. The details of data preprocessing were in Supporting Information.

Validation of candidate biomarkers using PRM assay

To evaluate the candidate biomarkers of prognosis, a unique PRM assay was generated, incorporating as many candidate proteins as possible. For the proteome profiling samples, the peptide was examined utilizing a Q Exactive HF-X Hybrid Quadrupole-Orbitrap Mass Spectrometer (Thermo Fisher Scientific), integrated with a state-of-the-art high-performance liquid chromatography system (EASY nLC 1200, Thermo Fisher Scientific). The experimental details of the PRM assay were in Supporting Information.

Bioinformatic analysis

Data quality control and pre-processing of proteomics

The proteomics data analysis encompassed the following key procedures mentioned in former studies [15, 16]. Step 1: Screen proteins with unique peptides ≥ 1 for further analysis. Step 2: Proteins with more than 50% null values in three groups were excluded. Step 3: For proteins whose effective value is ≥ 50% in one group, the empty values were filled with the mean value of the group. The remaining missing values were filled using half of the sample's minimum value. Step 4: high-quality proteins were retained by Log₂ transformation and z-normalization for subsequent data analysis.

Specifically, 2669 proteins were collected for downstream statistical and bioinformatics analysis.

Screening of differentially abundant proteins (DAP), differentially abundant metabolites (DAM), and enrichment analysis

Protein abundance changes in different sample groups were conducted through principal component analysis (PCA). For circulating proteomic data from the discovery cohort, the fold change (FC) value was derived from the ratio of ARDS to non-ARDS cases. Statistically significant differentially abundant proteins (DAPs) and differentially abundant metabolites (DAMs) were identified based on the criteria of $FC \geq 2$ or $FC \leq -2$, and $P < 0.05$. P -values were adjusted for false discovery rate (FDR) using Benjamini and Hochberg. We performed Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment to investigate further potential pathological and biological mechanisms associated with ARDS. The functional enrichment analysis of KEGG was performed with the ClusterProfiler package in *R*. KEGG pathway database (<https://www.kegg.jp/kegg/pathway.html>) was applied for metabolites pathway enrichment analysis.

Utilizing the results from KEGG pathways, we identified key interactions between enriched proteins and metabolites in ARDS. These elements were systematically linked to their corresponding pathways to elucidate the molecular mechanisms potentially influencing ARDS development. This linkage is visually represented through network diagrams, created by Cytoscape [17].

Protein-protein interaction network analysis

The protein-protein interactions (PPIs) were obtained from the STRING database [18]. Differentially abundant proteins (P -value < 0.05) were mapped to PPIs to generate the DAP PPI network in ARDS. The Cytoscape software [17] was used to visualize the network. The Cytoscape plugin cytoHubba [19] was utilized to calculate the degree in the PPI network.

Gene set enrichment analysis (GSEA) analysis

For the GSEA enrichment analysis, we utilized the Molecular Signatures Database (MSigDB), specifically focusing on the KEGG gene set. We designated an FDR threshold of 0.05 as the boundary for statistical significance. The process for computing the Normalized Enrichment Score (NES) within the GSEA framework entails the prioritization of proteins based on their statistical relevance, ranging from the most to the least significant, succeeded by the examination of the distribution pattern of the proteins associated with each gene set

throughout the prioritized list. The integrated abundance of proteins was then calculated by utilizing the ClusterProfiler package of *R* [20].

Constructing a prognostic model based on the early serum proteome

We constructed three classification models to predict adverse outcomes of ARDS: first, screened candidate proteins were used to construct a protein-based model. Least absolute shrinkage and selection operator (LASSO) and Boruta methods were used to screen the candidate proteins in the discovery cohort, then combined with DAPs, candidate biomarkers with unique peptides ≥ 3 were proposed for further targeted proteomics analysis. Second, to compare the protein model with current clinical practice, a clinical risk model was constructed and optimized. The clinical parameters were selected by the LASSO method among the lab tests, SOFA, P/F ratio, and comorbidity. A third combined model was formed by stacking the clinical prognostic parameters with the protein parameters. Five machine learning methods, Naïve Bayes (Bayes), Random forest (RF), Generalized Linear Model (Glm), Supporting Vector machine (SVM), and Gradient Boosting Machine (GBM) were used in the three classification models. Discrimination performance was assessed using the receiver operating characteristic (ROC) curve with an area under the curve (AUC), sensitivity, and specificity. The DeLong test was strategically utilized to statistically compare the AUCs across different models, providing a robust assessment of their discriminative capabilities. To prevent overfitting, '10-fold cross-validation' was employed to assess the performance of machine learning methods.

Statistical analysis

Mann–Whitney U tests or Kruskal–Wallis tests were used for comparisons of continuous variables, whereas *Chi-square* tests were used for categorical variables. Correlation analysis was assessed by *Spearman* correlation. A P value of < 0.05 (two-tailed) was considered statistically significant Benjamini–Hochberg correction for multiple testing was applied as appropriate.

Results

Participants characteristics

The characteristics of the validation and discovery cohorts are in Table 1. There was no significant difference in the ARDS etiologies between the discovery and validation cohorts. For the discovery cohort, the demographic and clinical characteristics of 130 ARDS cases, including 54 survived ARDS and 76 deceased ARDS are presented in Table 2. At baseline, the mean

Table 1 Baseline characteristics of the participants in the discovery and validation cohorts

| | ARDS in the discovery cohort (N = 130) | ARDS in the validation cohort (N = 183) | Disease control (N = 33) | Healthy control (N = 33) | P1 | P2 |
|--|--|---|--------------------------|--------------------------|---------|-----------|
| Age (mean, SD) | 72.5 (11.4) | 72.6 (14.3) | 72.1 (11.6) | 72.0 (8.2) | 0.889 | 0.883 |
| Male (n, %) | 97 (74.6) | 134 (73.2) | 25 (75.8) | 25 (75.8) | 0.783 | 0.774 |
| SOFA | 6 [4–9] | 4 [3–7] | 3 [2, 3] | – | 0.001** | <0.001*** |
| P/F ratio | 109 [81–151] | 150.0 [90.0–239.0] | 265 [203–354] | – | 0.001** | <0.001*** |
| RBC (*10 ¹² /L) | 3.66 [2.90–4.11] | 3.91 [3.28–4.28] | 4.10 [3.77–4.50] | – | 0.099 | 0.367 |
| Hb (g/L) | 111.0 [90.0–128.0] | 115.0 [96.0–131.0] | 127.5 [113.3–137.3] | – | 0.167 | 0.657 |
| WBC (*10 ⁹ /L) | 10.35 [7.65–14.30] | 10.51 [7.36–14.50] | 7.2 [5.4–11.2] | – | 0.761 | 0.676 |
| NEU (*10 ⁹ /L) | 9.30 [6.60–12.45] | 9.0 [6.2–13.0] | 5.3 [3.7–9.7] | – | 0.968 | 0.352 |
| LYM (*10 ⁹ /L) | 0.4 [0.2–0.6] | 0.6 [0.4–1.0] | 0.8 [0.5–1.1] | – | 0.001** | 0.234 |
| PLT (*10 ⁹ /L) | 153.5 [82.5–229.3] | 155.0 [110.0–222.0] | 191.0 [151.5–254.0] | – | 0.545 | 0.005** |
| ALT (U/L) | 27.0 [18.0–45.3] | 30.0 [18.0–53.0] | 30.0 [22.0–41.0] | – | 0.282 | 0.561 |
| AST (U/L) | 29.0 [20.0–47.0] | 34.0 [21.0–53.0] | 28.0 [21.0–50.8] | – | 0.158 | 0.041* |
| STB (μmol/L) | 11.4 [7.6–19.2] | 11.10 [8.60–20.60] | 9.9 [6.7–11.7] | – | 0.218 | 0.931 |
| CB (μmol/L) | 4.9 [3.1–8.4] | 4.10 [3.00–6.70] | 3.9 [3.0–5.2] | – | 0.479 | 0.938 |
| BUN (mmol/L) | 12.15 [9.18–20.98] | 9.30 [6.10–15.00] | 6.9 [5.5–9.0] | – | 0.001** | 0.049* |
| Cr (μmol/L) | 90.0 [64.8–128.8] | 79.0 [63.0–113.0] | 77.5 [62.5–97.8] | – | 0.131 | 0.097 |
| Alb (g/L) | 31.0 [28.0–34.0] | 32.0 [29.0–35.0] | 34.0 [32.0–37.0] | – | 0.058 | 0.006** |
| CRP (mg/L) | 60.3 [19.6–90.0] | 64.0 [23.4–145.1] | 23.9 [4.4–62.4] | – | 0.020* | 0.109 |
| PCT (ng/mL) | 0.29 [0.11–0.61] | 0.26 [0.11–1.22] | 0.07 [0.03–0.19] | – | 0.177 | 0.568 |
| Fg (mg/dL) | 419.5 [278.3–540.3] | 448.0 [212.0–526.0] | 469.0 [269.0–537.0] | – | 0.655 | 0.921 |
| PT (second) | 13.3 [12.5–14.7] | 13.5 [12.6–15.1] | 12.6 [12.0–13.3] | – | 0.456 | 0.079 |
| APTT (second) | 29.3 [26.4–34.3] | 29.2 [26.8–35.5] | 27.6 [25.9–30.9] | – | 0.108 | 0.944 |
| D-dimer (mg/L) | 4.22 [1.75–10.7] | 2.86 [1.28–8.06] | 1.13 [0.57–3.31] | – | 0.079 | 0.158 |
| Hypertension (n, %) | 87 (66.9) | 90 (49.2) | 23 (69.7) | – | 0.002** | 0.607 |
| Coronary heart disease (n, %) | 30 (23.1) | 36 (19.7) | 7 (21.2) | – | 0.468 | 1.000 |
| Diabetes (n, %) | 50 (38.5) | 57 (31.1) | 14 (42.4) | – | 0.180 | 0.054 |
| Cerebrovascular disease (n, %) | 23 (17.7) | 30 (16.4) | 2 (6.1) | – | 0.763 | 0.659 |
| Immunosuppression (n, %) | 13 (10.0) | 20 (10.9) | 1 (3.0) | – | 0.792 | 0.953 |
| Vasopressor use at enrolment (n, %) | 24 (18.5) | 34 (18.6) | 3 (9.1) | – | 1.000 | 0.303 |
| Neuromuscular blockade use at enrolment (n, %) | 39 (30.0) | 50 (27.3) | 0 (0) | – | 0.696 | <0.001*** |
| Mechanical ventilation (n, %) | 118 (90.8) | 165 (90.2) | 0 (0) | – | 1.000 | <0.001*** |
| Prone positioning at enrolment (n, %) | 70 (53.8) | 96 (52.5) | 0 (0) | – | 0.899 | <0.001*** |
| Risk factors (n, %) | | | | | 0.962 | 0.882 |
| Pneumonia | 96 (73.8) | 137 (74.9) | 26 (75.8) | – | | |
| Sepsis | 26 (20.0) | 36 (19.7) | 6 (18.2) | – | | |
| Other | 8 (6.2) | 10 (5.5) | 1 (3.0) | – | | |
| Source of infection (n, %) | | | | | 0.928 | 0.786 |
| Thorax | 96 (73.8) | 137 (74.9) | 26 (75.8) | – | | |
| Abdomen | 24 (18.5) | 34 (18.6) | 6 (18.2) | – | | |
| Other | 10 (7.7) | 12 (6.5) | 1 (3.0) | – | | |

P1: Comparison between ARDS patients in the discovery cohort and validation cohort; P2: Comparison between ARDS patients in the discovery cohort and Disease control

*P<0.05, **P<0.01, ***P<0.001

age of the ARDS group was 72.5 (SD, 11.4) years, and 74.6% were male. After the diagnosis of ARDS, non-survivors showed decreased P/F ratio, PLT, and Alb

levels, alongside increased SOFA scores, BUN, and AST, compared to the survivors. Gender distribution and age were matched in DC and HC with ARDS

Table 2 Baseline characteristics of the discovery cohort

| | Survived ARDS (n = 54) | Deceased ARDS (n = 76) | P |
|--|------------------------|------------------------|-----------|
| Age (mean, SD) | 72.2 (11.8) | 73.7 (11.2) | 0.883 |
| Male (n, %) | 41 (75.9) | 56 (73.7) | 0.774 |
| SOFA | 5 [4–7] | 6 [5–9] | <0.001*** |
| P/F ratio | 122 [95–200] | 98 [73–138] | <0.001*** |
| RBC (*10 ¹² /L) | 3.55 [2.88–4.00] | 3.71 [2.95–4.16] | 0.367 |
| Hb (g/L) | 108.5 [85.8–128.0] | 111.5 [91.3–127.8] | 0.657 |
| WBC (*10 ⁹ /L) | 9.83 [7.56–12.32] | 10.41 [7.66–15.04] | 0.676 |
| NEU (*10 ⁹ /L) | 8.80 [6.38–10.65] | 9.55 [6.70–13.28] | 0.352 |
| LYM (*10 ⁹ /L) | 0.4 [0.3–0.8] | 0.4 [0.2–0.6] | 0.234 |
| PLT (*10 ⁹ /L) | 189.5 [126.8–249.5] | 129.0 [68.0–205.0] | 0.005** |
| ALT (U/L) | 29.5 [19.8–54.8] | 25.5 [17.0–44.0] | 0.561 |
| AST (U/L) | 27.0 [20.0–34.0] | 34.0 [21.0–54.0] | 0.041* |
| STB (μmol/L) | 10.8 [7.7–17.6] | 12.3 [7.4–20.1] | 0.931 |
| CB (μmol/L) | 4.5 [3.2–6.5] | 5.5 [3.1–8.6] | 0.938 |
| BUN (mmol/L) | 11.55 [9.45–15.85] | 12.8 [8.6–24.38] | 0.049* |
| Cr (μmol/L) | 81.0 [63.5–111.8] | 96.0 [67.5–169.8] | 0.097 |
| Alb (g/L) | 32.0 [29.0–35.3] | 30.0 [28.0–32.0] | 0.006** |
| CRP (mg/L) | 55.1 [8.6–68.3] | 68.9 [26.3–98.3] | 0.109 |
| PCT (ng/mL) | 0.23 [0.08–0.76] | 0.30 [0.11–0.56] | 0.568 |
| Fg (mg/dL) | 410.5 [249.8–560.3] | 427.0 [301.0–538.0] | 0.921 |
| PT (second) | 13.1 [12.2–14.0] | 13.5 [12.7–15.1] | 0.079 |
| APTT (second) | 29.4 [25.6–34.2] | 29.3 [27.0–34.3] | 0.944 |
| D-dimer (mg/L) | 3.27 [1.42–7.13] | 4.41 [2.02–11.50] | 0.158 |
| Hypertension (n, %) | 38 (70.4) | 49 (64.5) | 0.607 |
| Coronary heart disease (n, %) | 12 (22.2) | 18 (23.7) | 1.000 |
| Diabetes (n, %) | 15 (27.8) | 35 (46.1) | 0.054 |
| Cerebrovascular disease (n, %) | 11 (20.4) | 12 (15.8) | 0.659 |
| Immunosuppression (n, %) | 6 (11.1) | 7 (9.2) | 0.953 |
| Vasopressor use at enrolment (n, %) | 8 (14.8) | 16 (21.1) | 0.500 |
| Neuromuscular blockade use at enrolment (n, %) | 14 (25.9) | 25 (32.9) | 0.509 |
| Mechanical ventilation (n, %) | 48 (88.9) | 70 (92.1) | 0.553 |
| Prone positioning at enrolment (n, %) | 30 (55.6) | 40 (52.6) | 0.880 |
| ARDS risk factors (n, %) | | | 0.397 |
| Pneumonia | 42 (77.8) | 54 (71.0) | |
| Sepsis | 8 (14.8) | 18 (23.7) | |
| Other | 4 (7.4) | 4 (5.3) | |
| Source of infection (n, %) | | | 0.369 |
| Thorax | 42 (77.8) | 54 (71.0) | |
| Abdomen | 7 (13.0) | 17 (22.4) | |
| Other | 5 (9.2) | 5 (6.6) | |

P* < 0.05, *P* < 0.01, ****P* < 0.001

(Table 1). We acquired serum proteome profiles of all participants (*n* = 196) using a DIA strategy and global metabolome by the LC-MS method. An overview of proteomics and metabolomics workflow is showed in Fig. 1A. The independent prospective validation cohort incorporating 183 early ARDS patients with 85

deceased and 98 survived (Fig. 1A and Table S1). The characteristics of the validation and discovery cohorts were nearly consistent, except for Alb and Cr (Table 2 and Table S1).

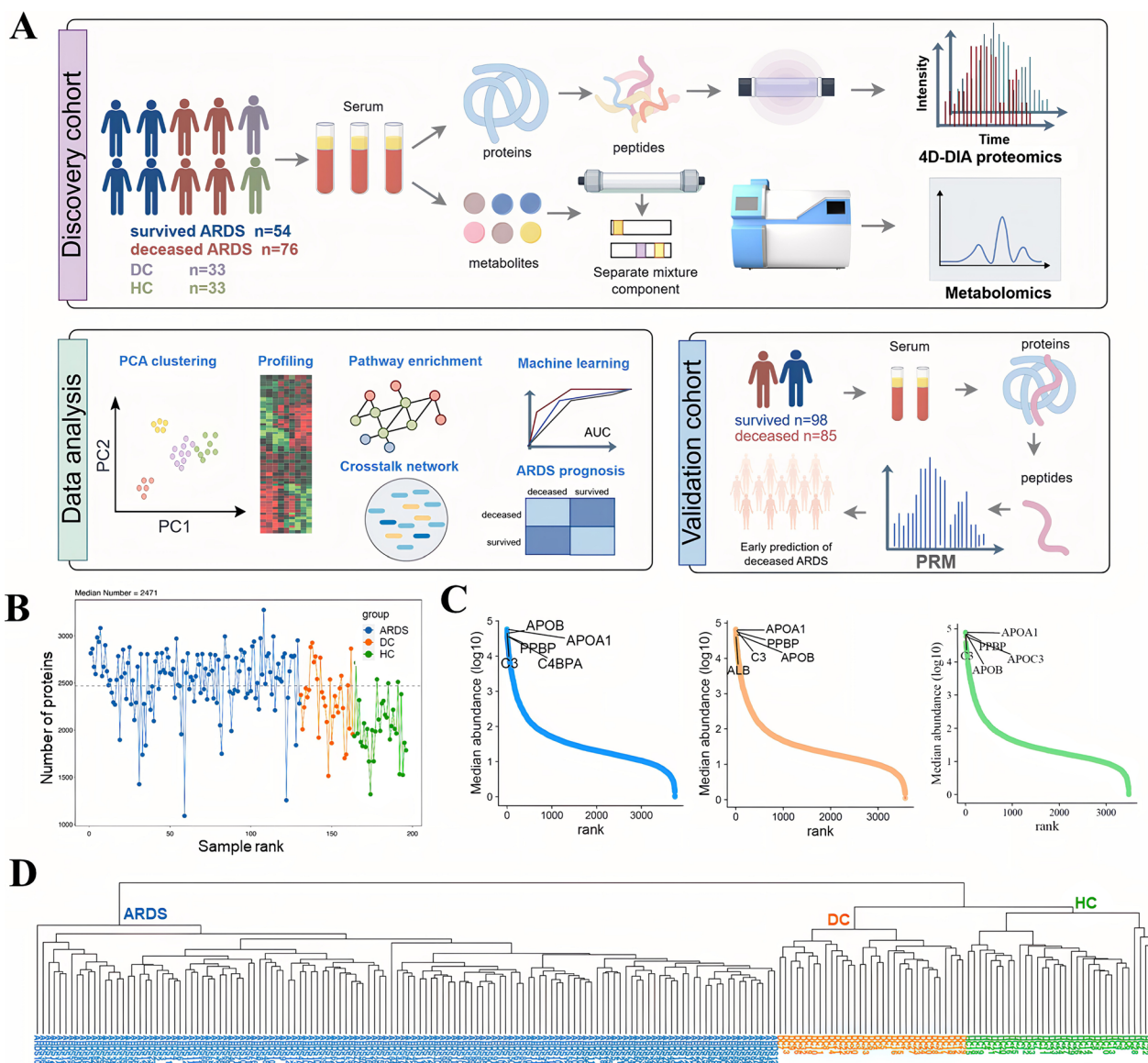


Fig. 1 Study design overview and summary of the serum proteomic analysis of ARDS, disease control, and health control. **A** Overview of the serum proteomic and metabolic workflow, including cohort construction (survived ARDS: N = 54, deceased ARDS: N = 76, disease control: N = 33, and healthy control: N = 33), serum validation cohort (survived ARDS: N = 98 and deceased ARDS: N = 85) and data analysis (proteomic and metabolic data). **B** A number of proteins were identified and quantified in three groups. **C** The top five most abundant proteins are labeled. **D** The clustering tree of each sample

Altered serum proteomic profiling in the ARDS group and the biological pathways

After filtering the low abundant proteins, 2669 high-quality proteins were collected for the data analysis (Additional file 1). The median protein number for 196 samples was 2471 (Fig. 1B–C). The clustering tree indicated the appropriate pre-processing method in this study (Fig. 1D). Our preliminary analysis employed PCA method to explore the clustering patterns among the groups (Fig. 2A and Figs. S2A–C), which demonstrated

a distinct separation between ARDS samples and non-ARDS control groups. A comprehensive differential expression analysis revealed that 1069, 319, and 511 proteins were uniquely altered in ARDS vs. HC, ARDS vs. DC, and DC vs. HC groups, respectively (Fig. 2B–C and Fig. S2D, as detailed in Additional file 2). These findings demonstrated the progressive features of serum protein alterations correlating with the severity of the disease. A critical subset of 214 differentially abundant proteins emerged as consistently regulated across ARDS, with

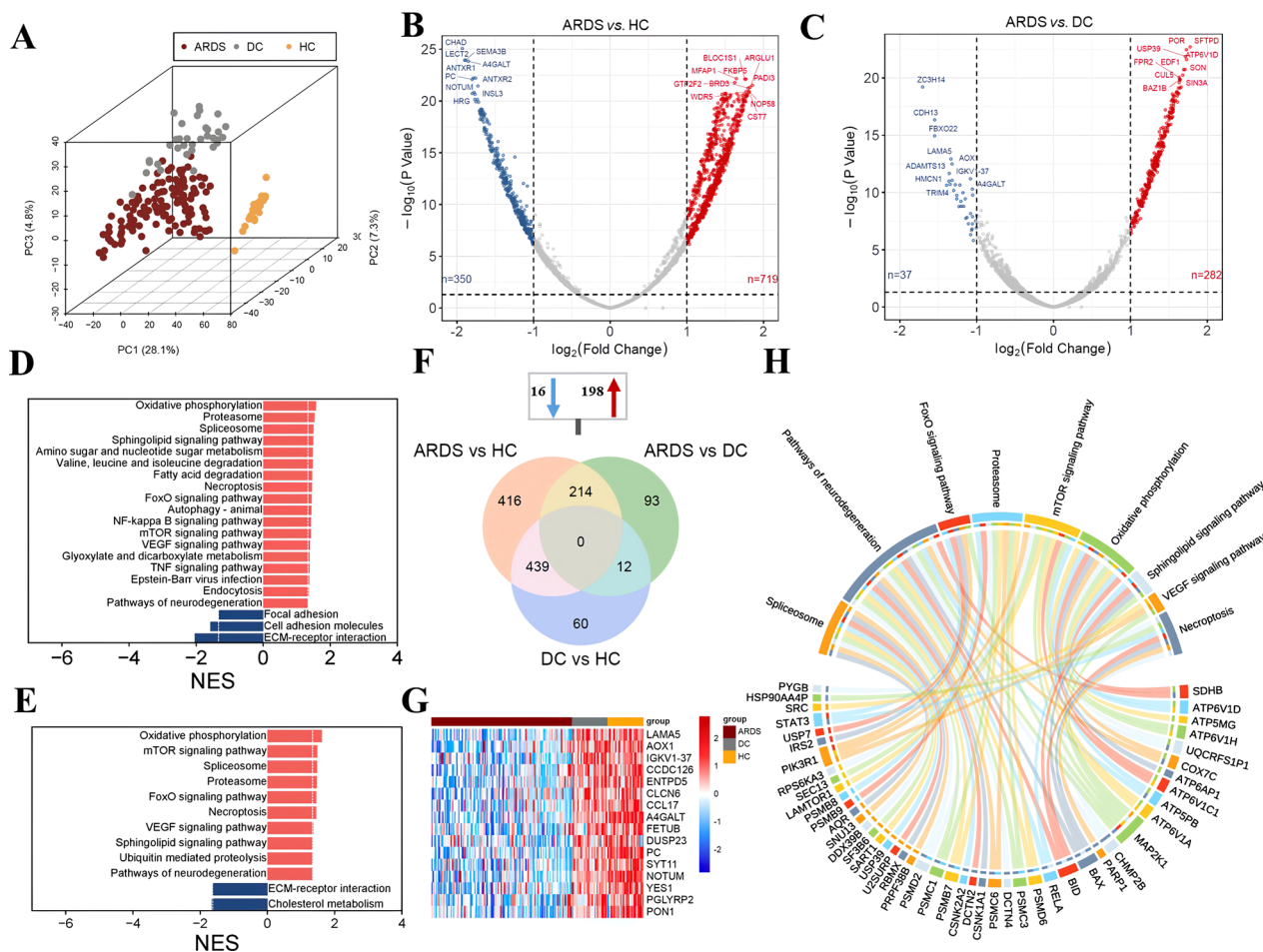


Fig. 2 Differentially abundant proteins and functional alterations in ARDS. **A** PCA discriminates ARDS patients from non-ARDS subjects. The % value indicates the explained variance. Volcano plot displaying differentially abundant proteins between ARDS vs. HC **B** and ARDS vs. DC **C**. Each dot represents a protein, with red dots for proteins significantly upregulated in ARDS and blue dots for proteins significantly downregulated in ARDS. *P* values were calculated using the *R* package ‘limma’ and adjusted using the Benjamini–Hochberg method. Two-sided *P* values were calculated. Fold change ≥ 2.0 and FDR-adjusted *P* < 0.05 is considered statistically significant. **D–E** KEGG pathway enrichments of ARDS compared with HC and DC groups, respectively. **F** A Venn diagram elucidates the distribution of DAP across the ARDS, DC, and HC groups. **G** The heatmap showed the 16 overlapped down-regulated proteins in ARDS. **H** A chordal graph maps the interaction between 9 common pathways and their associated proteins

16 proteins showing persistent downregulation and 198 showing upregulation in comparisons of both ARDS vs. DC and ARDS vs. HC (Fig. 2E, Additional file 3). Among the downregulated proteins, Fetuin B (FETUB) can be considered a reliable biomarker for predicting mortality in *Staphylococcus aureus* bacteremia (SaB) patients because decreased FETUB levels correlate strongly with poor clinical outcomes in patients with SaB [10]. Paraoxonase 1 (PON1), negatively associated with higher mortality of sepsis [21], underlines significance in ARDS pathology compared to DC and HC groups (Fig. 2G). Conversely, certain proteins, including Surfactant Protein D (SFTPD) and Signal Transducers and Activators of Transcription 3 (STAT3) et al., were elevated (Additional file 3). Moreover, SFTPD, a circulating epithelial marker

[22], and STAT3, an activator of macrophages and neutrophils [23], were implicated as potential key contributors to the pathogenesis of ARDS.

Pathway analysis identified nine distinct pathways that were specifically modulated in ARDS (Fig. 2D–E and Fig. S3). Notably, the oxidative phosphorylation pathway emerged as significantly upregulated in comparisons of ARDS with both DC (NES = 1.615, *P* = 0.001) and HC (NES = 1.582, *P* < 0.001), revealing its pivotal role in the energetic metabolism associated with ARDS (comprehensive pathway listings are available in Additional file 4 and Supporting Information). Fig. 2H illustrated the enriched proteins within these nine overlapping pathways, all exhibiting increased serum levels in ARDS cases. Among the highlighted proteins, Succinate

Dehydrogenase Complex Iron-Sulfur Subunit B (SDHB), involved in mitochondrial function, and various components of the vacuolar ATPase family (ATP6V1D, ATP5MG, ATP6V1H) were implicated in oxidative phosphorylation. Moreover, proteins such as Mitogen-Activated Protein Kinase Kinase 1 (MAP2K1), Phosphoinositide-3-Kinase Regulatory Subunit 1 (PIK3R1), and Non-Receptor Tyrosine Kinase (SRC), associated with the VEGF signaling pathway, were significantly upregulated in ARDS, illustrating a comprehensive network of molecular interactions contributing to the pathogenesis and progression of ARDS.

Cross-talk between proteomics and metabolomics implicates the sphingolipid signaling pathway as a mediator in ARDS

In our metabolic investigation, we cataloged a comprehensive array of 3331 metabolites, encompassing amino acids, lipids, and other critical serum metabolites, as detailed in Additional file 5. The PCA plots revealed a pronounced metabolic differentiation of the ARDS in comparison to both DC and HC groups (Fig. 3A–C). We identified 214, 113, and 206 differentially abundant metabolites (DAMs) in the ARDS vs. HC, ARDS vs. DC, and DC vs. HC, respectively (Additional file 6). Among these, lysophosphatidylcholine (LysoPC) emerged as the most markedly altered metabolites in ARDS, exemplified

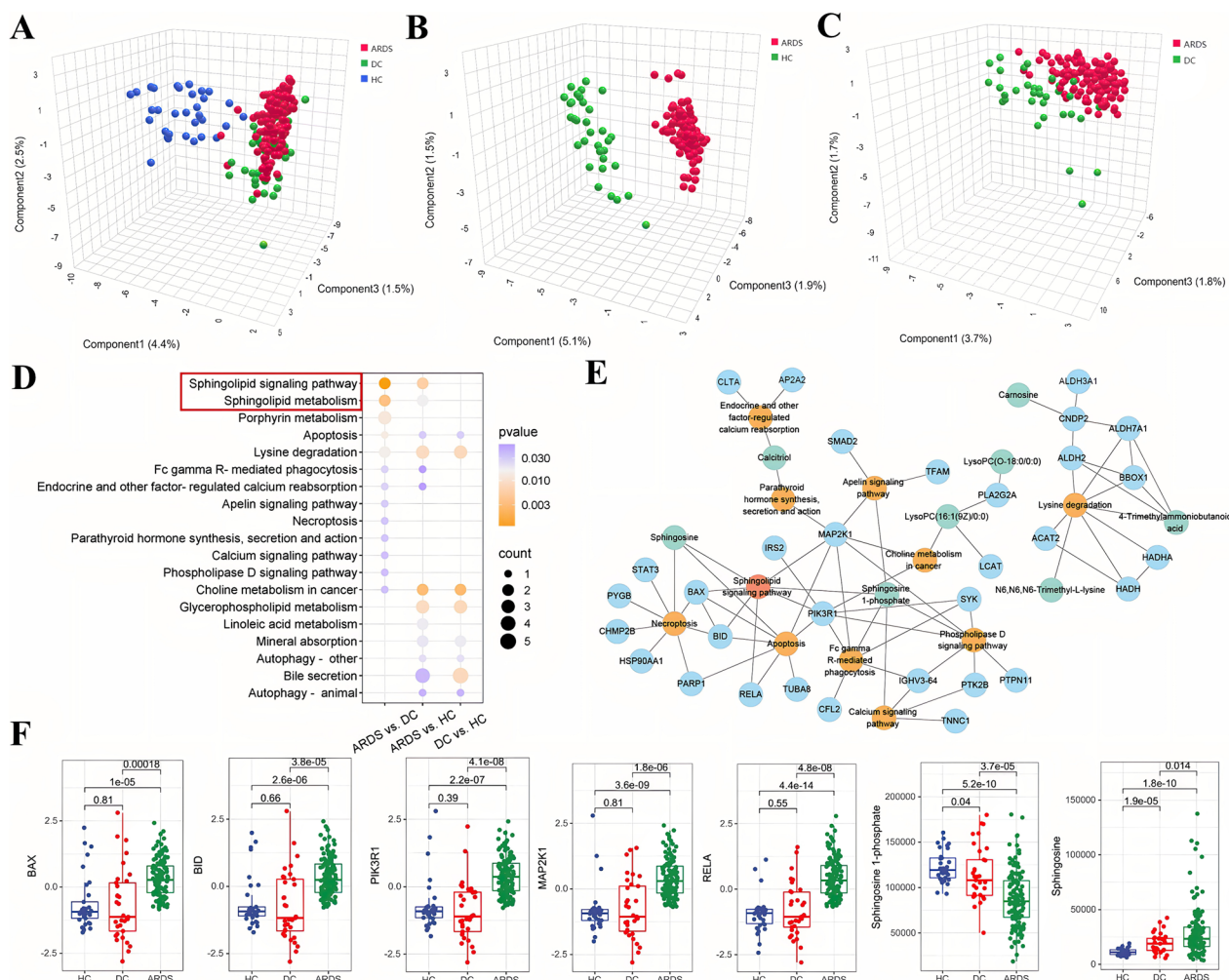


Fig. 3 The metabolic analysis and its interaction with proteomics in ARDS in the discovery cohort. **A–C** Partial least squares discriminant analysis (PLS-DA) showed the separation of ARDS patients from HC and DC groups. **D** Dot plot of enriched KEGG pathways identified by differentially abundant metabolites. Dot size indicates the number of enriched proteins, colors represent the significance of enrichment (*P*-value). **E** Cross-talk network of the significantly enriched pathway by integrating both proteomics and metabolomics. Blue dots represent proteins, green dots are metabolites, and orange dots are enriched pathways. Sphingolipid signaling pathway was highlighted in red. **F** The relative abundance of proteins and metabolites in sphingolipid signaling pathway across ARDS, HC, and DC groups

by a significant decrease in LysoPC (0:0/18:2(9Z, 12Z)) and LysoPC (0:0/18:0). Further exploration through the KEGG pathway notably highlighted the modulation of the sphingolipid signaling pathway and sphingolipid metabolism, which were distinctive in ARDS as compared to the non-ARDS groups (Fig. 3D). Importantly, Fig. 3E revealed that metabolite S1P linked the sphingolipid signaling pathway with several other signaling pathways, including the Apelin, calcium, and phospholipase D signaling pathways. Additionally, S1P levels were reduced, whereas sphingosine exhibited elevated levels in ARDS compared to both DC and HC groups (Fig. 3F).

We further conducted an integrated analysis of proteomics and metabolomics to elucidate the pathogenetic mechanisms underlying ARDS. A significant finding from our study was the identification of the sphingolipid signaling pathway as a key regulatory axis, influenced at both the protein and metabolite levels, pinpointing it as a hub pathway in ARDS pathogenesis (Fig. 3D and Table S2). Further network analysis demonstrated the critical role of the sphingolipid signaling pathway, positioning it centrally within the ARDS-specific regulatory network (Fig. 3E and Figs. S4–S5). Specifically, critical regulatory proteins such as BCL2 Associated X (BAX), BH3 Interacting Domain Death Agonist (BID), PIK3R1, MAP2K1, and NF- κ B Subunit (RELA) were upregulated in ARDS (Fig. 3F). Furthermore, MAP2K1, the hub protein in the network, linked the sphingolipid signaling pathway with several important pathways, including parathyroid hormone synthesis secretion and action, Apelin signaling pathway, choline metabolism, phospholipase D signaling pathway, Fc gamma R-mediated phagocytosis, and apoptosis. BAX and BID mediate interactions linking the sphingolipid signaling pathway with necroptosis and apoptosis, while RELA associates with the pathway in apoptosis. PIK3R1 connects the sphingolipid signaling pathway to Fc gamma R-mediated phagocytosis, apoptosis, phospholipase D signaling pathway, and choline metabolism (Fig. 3E).

Dysregulation of biological functions and alterations in metabolites LysoPCs and S1P in deceased patients at ARDS onset

To further investigate the proteomic alterations associated with prognosis in ARDS, we conducted a comparative analysis between ARDS survivors and non-survivors (Fig. 4A). Differential expression analysis identified 40 proteins with significant variations ($|\text{fold change}| \geq 1.5$ and $\text{FDR} < 0.05$) (Table S3). The most significantly altered proteins between deceased and survived ARDS were Radixin (RDX) and Moesin (MSN), which play crucial roles in linking actin to the plasma membrane (Fig. 4B). Subsequent GSEA revealed several highly ranked

molecular pathways markedly enriched across various biological functions such as energy metabolism, immune response, proteasome function, gap junction communication, calcium signaling, and hematopoietic cell lineage differentiation (Fig. 4C and Supporting Information). To refine our selection of key protein candidates, we spotlighted the top 25 proteins exhibiting the highest connectivity degrees in PPI network (Additional file 7). Among these, Glyceraldehyde-3-Phosphate Dehydrogenase (GAPDH), Heat Shock Protein 90 Alpha Family Class A Member 1 (HSP90AA1), and Enolase 1 (ENO1) emerged as central hub proteins within the network (Fig. 4D–E). This finding aligned with the GSEA results, as GAPDH and ENO1 were critical enzymes in glycolysis; moreover, HSP90AA1 was the housekeeping protein that aids protein folding and has intrinsic ATPase activity.

At the global metabolic level, the PCA plot demonstrated a slight separation between the ARDS-deceased and survived groups (Fig. 4F). The 45 DAMs were most enriched in signaling pathways, such as the sphingolipid signaling pathway, phospholipase D signaling pathway, calcium signaling pathway, and Apelin signaling pathway (Fig. 4G and Additional file 8). LysoPC (O-18:0/0:0), LysoPC (15:0/0:0), and S1P were prominently enriched in these pathways. Furthermore, the levels of LysoPC and S1P were significantly reduced in the ARDS-deceased group (Fig. 4H). More importantly, correlation analysis between these three metabolites and the severity of ARDS, as quantified by the SOFA score, revealed negative correlations ($R = -0.47$ for LysoPC (O-18:0/0:0); $R = -0.50$ for LysoPC (15:0/0:0); $R = -0.36$ for S1P) (Fig. 4I), suggesting the potential metabolome mechanism underlying the heightened mortality risk in ARDS non-survivors.

Biomarker panel for early prediction of ARDS prognosis

Further, a prognosis model was constructed to predict the outcome of ARDS early. We obtained 36 candidate biomarkers using LASSO (Fig. S6A–B), Boruta, and DAPs. Finally, 31 proteins with 174 peptides were targeted by PRM assay in the external validation cohort (Additional file 9), and a total of 22 candidate proteins with peptides ≥ 2 were retained, in which 8 proteins maintained consistent significance in both discovery and validation cohorts (Fig. 5A–B). Among the eight proteins, six proteins were significantly upregulated in deceased patients in both discovery and validation cohorts, including Vascular Cell Adhesion Molecule 1 (VCAM1), Lactate Dehydrogenase B (LDHB), MSN, Filaggrin 2 (FLG2), Lamin A/C (LMNA), and Lipopolysaccharide Binding Protein (LBP), while two proteins, Transgelin 2 (TAGLN2) and Mannose Binding Lectin 2 (MBL2) were consistently downregulated in the deceased ARDS group (Fig. 5B).

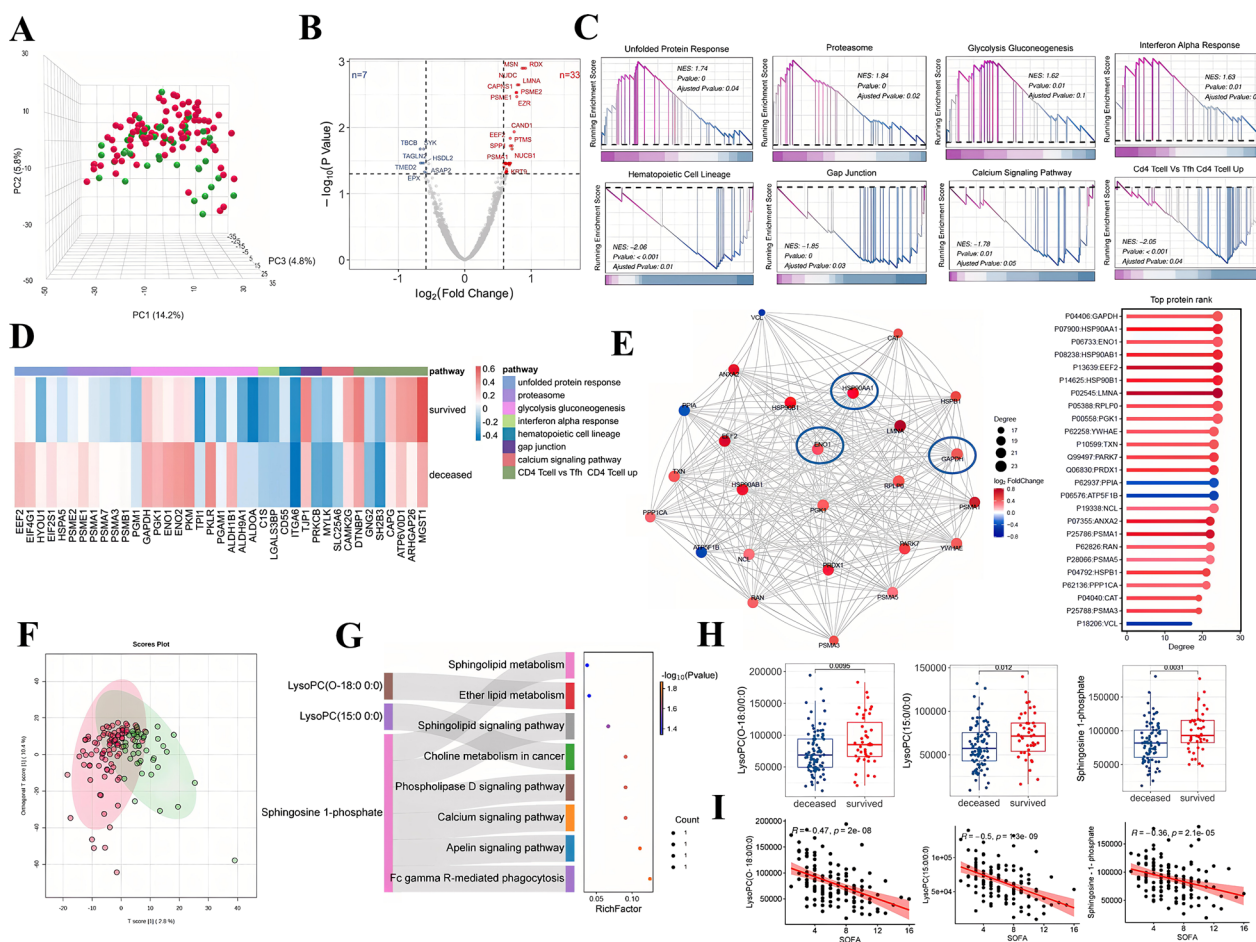


Fig. 4 The multi-omics analysis compared deceased ARDS with survived ARDS. **A** PCA plot displays the proteome separation in deceased and survived ARDS. Each dot represents a sample, with red dots for deceased ARDS samples and green dots for survived ARDS samples. **B** Volcano plot displaying differentially abundant proteins between deceased and survived ARDS. *P* values were adjusted using the Benjamini–Hochberg method. **C** GSEA analysis showed up- and down-regulated pathways in deceased ARDS compared with survived ARDS. **D** A heatmap showed representative proteins linked to the pathways highlighted in **C**. The mean difference (deceased ARDS–survived ARDS) of proteins in pathways larger than 0.1 were kept. **E** Left: PPI network analysis of DAPs between deceased ARDS and survived ARDS. The significance threshold was defined as $P < 0.05$. The sliding ball chart on the right shows the network degree and \log_2FC of the corresponding proteins in the PPI network. **(F)** PLS-DA score plots between deceased and survived ARDS for metabolomics. **G** The Sankey diagram showed the enriched pathways and their associated DAMs. **H** The abundance of LysoPC (O-18:0/0:0), LysoPC (15:0/0:0), and S1P between deceased and survived ARDS groups. Statistical significance was determined using the FDR-adjusted *P*-value. *P* values were calculated by *Wilcoxon's* rank sum test. **I** Correlations between metabolites mentioned in **H** and the SOFA scores. The correlation coefficients (simplified as *R*) and rho values of correlations are indicated

These markers were selected as an eight-protein panel for early identification of deceased patients of ARDS.

In the discovery cohort, we first compared five state-of-the-art machine learning classifiers using the protein model. The Glm model was the final classifier due to its overall superior performance, evidenced by a 10-fold cross-validated ROC–AUC of 0.893 (95% CI 0.837–0.949) and a sensitivity of 0.920 (Fig. 5C). In comparison, the clinical risk model, comprising the SOFA score, P/F ratio, and PLT from LASSO feature selection (Fig. S6C–D), yielded in a ROC-AUC of 0.784 (95% CI 0.703–0.866) and a sensitivity of 0.885 (Fig. 5D).

Combination of parameters in both models resulted in a ROC-AUC of 0.890 (95% CI 0.834–0.946) and a sensitivity of 0.931 (Fig. 5E). The protein model performed significantly better than the clinical risk model (DeLong test, $P < 0.001$), whereas the combination of both models did not exhibit statistical superiority over the protein model alone (DeLong test, $P = 0.970$). Subsequent validation of the Glm classifier using the protein model resulted in a ROC-AUC of 0.802 (95% CI 0.739–0.865) and a sensitivity of 0.835 (Fig. 5F) in the external cohort, whereas the clinical risk model resulted in a ROC-AUC of 0.738 (95% CI 0.655–0.811) and a sensitivity of 0.788

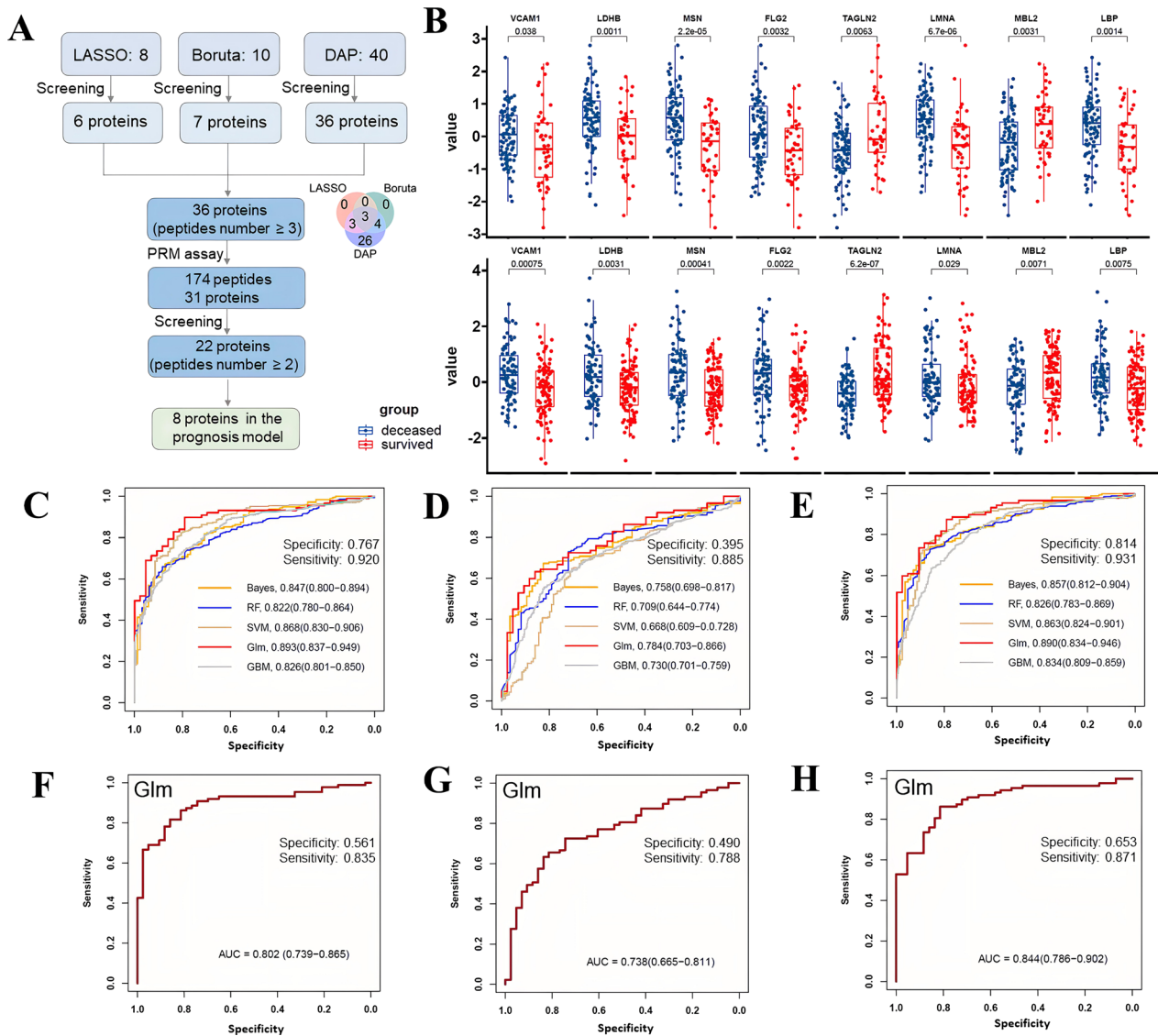


Fig. 5 Prognostic model development and validation for ARDS. **A** The flowchart details the methodology for screening candidate biomarkers. **B** The expression levels of eight proteins between the deceased and survived ARDS in both discovery (the upper panel) and validation cohorts (the lower panel). Receiver operating characteristic curve of protein **C**, clinical **D**, and combined model **E** in the discovery cohort and the validation cohort **F–H**. The 95% confidence interval is shown between brackets. The sensitivity and specificity of the Glm model were also provided. AUC, area under the curve

(Fig. 5G). The combination of both models led to a ROC-AUC of 0.844 (95% CI 0.786–0.902) (Fig. 5H). In the validation cohort, the protein model also outperformed the clinical risk model (DeLong test, $P = 0.008$), and a combination of both models was superior to the protein model alone (DeLong test, $P = 0.006$).

Discussion

This was the first study to integrate proteomics and metabolomics to investigate differences in the proteome, metabolome, and related biological pathways in adult

ARDS patients. Our findings identified eight proteins as early candidates for distinguishing deceased ARDS cases from survivors. These candidate proteins, along with the constructed prognostic model, were validated in an independent external cohort, revealing their potential utility in clinical prognosis.

This integrative analysis offers a global perspective on the proteomic and metabolic landscapes of ARDS, highlighting the potential involvement of the sphingolipid signaling pathway, specifically the role of S1P, a molecule synthesized through the phosphorylation of

sphingosine. Indeed, S1P was a sphingolipid that helped improve glycocalyx integrity by inhibiting syndecan-1 shedding [24]. S1P activated the S1P1 receptor, which attenuated the activity of metalloproteinases causing syndecan-1 ectodomain shedding [24, 25]. In addition, previous studies have suggested that S1P played a significant role in regulating vascular permeability, barrier integrity, and inflammation in ARDS [26–32]. S1P was protective against acute lung injury by regulating endothelial cell barrier integrity through binding to S1P receptors [30, 31, 33]. Murine and canine models revealed that S1P exerted potent beneficial systemic effects, significantly reducing vascular leak, LPS-induced lung edema formation, and inflammatory lung injury [28, 29]. Therefore, targeting the S1P or S1P1 receptor activation may be a potential therapeutic option for inhibiting glycocalyx degradation in ARDS. Similarly, decreased serum S1P levels have been associated with septic shock severity [34] and a significant association was identified between sphingolipid metabolism and outcomes in trauma patients [35]. These findings indicated that lower S1P levels correlated with increased disease severity in ARDS, although our study has not established mechanistic data.

In addition, in our study, ARDS patients exhibited higher levels of MAP2K1 (i.e., MEK1). MAP2K1, involved in the sphingolipid signaling pathway, was identified as the hub protein in the proteome-metabolome cross-talk network. Furthermore, Jonathan et al. found S1P signals regulated lymphatic endothelial cells permeability and junction molecule expression through MAPKs [36]. Meanwhile, MEK1 has been implicated in various signaling networks that affect tumor necrosis factor (TNF) mediated antiapoptotic signaling [37] and inflammatory profile [38]. Therefore, MEK inhibition has important therapeutic implications for ARDS by potentially disrupting the sphingolipid signaling pathway and its associated biological processes.

Notably, the concentrations of LysoPC species such as LysoPC (O-18:0/0:0) and LysoPC (15:0/0:0) in the deceased ARDS groups were decreased compared with the survived group. LysoPC is a lipid mediator derived from membrane phosphatidylcholine, and it is known to contribute to inflammation by increasing chemokine production and activating endothelium, neutrophils, monocytes, macrophages, and lymphocytes [39]. However, the role of LysoPCs in ARDS has not yet been clearly elucidated. In sepsis patients, serum LysoPC levels were lower than those in controls, and among them, LysoPC (16:0) and LysoPC (18:0) were decreased [40]. Lower concentrations of serum or plasma LysoPC predicted worse outcomes [41]. Our findings align with these observations, revealing a comparable reduction in serum LysoPC

levels in ARDS, akin to that observed in sepsis. Therefore, LysoPCs could be important metabolite markers for differentiating the diagnosis and prognosis of ARDS, which needs further validation and mechanism exploration.

Moreover, our analysis presented a comprehensive view of proteomic alterations distinguishing deceased ARDS from survived ARDS, the elucidation of the functional roles and molecular mechanisms of these proteins in ARDS lays the groundwork for potential therapeutic innovations targeting energy metabolism, such as glycolysis. One previous study showed that the increase of pro-inflammatory cytokines including pro-IL-1 β and TNF- α preceded the activation of glycolysis in macrophages during LPS-induced ALI, suggesting that inflammation may induce glycolysis [42]. In COVID-19 ARDS, T cells are exhausted and skewed towards glycolysis, with a concomitant reduction in mitochondrial dependence, probably as a result of reduced oxygenation of the pulmonary tissue [43, 44]. Consistently, we identified upregulated glycolysis, interferon- α response, and activated CD4⁺ T cells in ARDS, indicating that glycolysis is coordinated with inflammation, beyond just generating energy and producing building blocks for cellular survival and signal transduction [45].

In this study, our serum biomarkers and prognosis model could indeed be a promising prognosis predictor of ARDS patients compared to traditional clinical models. Therefore, we proposed that the protein model based on serum proteomics has shown promise in identifying high-risk populations for deceased ARDS. Interestingly, a previous study revealed VCAM1 overexpressed on endothelial cell-derived extracellular vesicles during sepsis, facilitating the activation of the NF- κ B pathway by interacting with integrin subunit alpha 4 (ITGA4) on the monocyte surface thereby regulating monocyte differentiation [46]. Furthermore, the biomarker MBL2 has been identified as a critical determinant of mortality risk among patients with severe pneumococcal infections who exhibit MBL deficiency [47]. ‘MBL deficiency’ appears to play a role in increasing the susceptibility to severe infections in patients receiving stem cell transplantation [48, 49], highlighting the pathogenic significance of this innate immune defense protein. Other candidate biomarkers, for example, LDHB was engaged in glycolysis [50], MSN was expressed in the endothelial cell and was important for cell-cell recognition and signaling and for cell movement [51]. LBP is the LPS binding protein, the LBP-LPS complex initiates a signal cascade that triggers the secretion of pro-inflammatory cytokines [52]. Abnormal LMNA results in nuclear structural abnormalities and mesenchymal tissue damage [53], which are critical in ARDS pathogenesis due to their roles in cellular integrity and repair mechanisms. FLG2 is

involved in epidermal maturation [54], a process that may impact the barrier function of the alveolar epithelium in ARDS. TAGLN2 facilitating the formation of intracellular cytoskeleton structures [55], which are essential for maintaining cell shape and motility, and crucial in the inflammatory response seen in ARDS. These biomarkers can be categorized into several groups: epithelial injury, endothelial damage, imbalanced immune system, mitochondrial dysfunction, and disrupted cytoskeleton, which engaged in the overall ARDS progress. However, the specificity of these eight proteins for predicting ARDS prognosis should be confirmed through either singular or multiple assessments, including enzyme-linked immunosorbent assay (ELISA), proximity extension assay (PEA), or bead-based immunoassays, which would benefit its application in future clinical practice.

We acknowledge several limitations of the present study. First, while we identified eight proteins associated with adverse prognosis in ARDS, further studies are required to explore the underlying mechanisms by which these proteins influence disease outcomes, particularly focusing on their specific roles within the immune and pathology systems. Second, longitudinal multi-omics studies are needed to explore the dynamic features of ARDS and the temporal changes in prognostic biomarkers. Such studies will provide a more comprehensive understanding of the disease progression and the potential for early intervention. Third, integrating metabolomics and proteomics presents significant methodological challenges, such as data harmonization and handling missing values, which may introduce biases and affect the robustness of our findings. Lastly, the identification of intracellular metabolites or proteins in the circulation does not necessarily indicate changes in cellular or tissue processes. These findings may also reflect cell death and the release of intracellular contents, which should be taken into account when interpreting the results.

In conclusion, this comprehensive study highlights the importance of the sphingolipid signaling pathway in unraveling pathogenesis of ARDS. Protein MAP2K1 and metabolite S1P play critical roles in this pathway. We show that a panel of eight proteins is superior to a clinical prognostic model in predicting ARDS-deceased events. These findings have significant implications for risk assessment and potential guidance of therapeutic strategies in the management of ARDS.

Abbreviations

| | |
|-----|----------------------------------|
| ICU | Intensive care unit |
| S1P | Sphingosine 1-phosphate |
| PRM | Parallel reaction monitoring |
| DIA | Data-independent acquisition |
| PCA | Principal component analysis |
| DAP | Differentially abundant proteins |

| | |
|-------|---|
| DAM | Differentially abundant metabolites |
| GSEA | Gene set enrichment analysis |
| LASSO | Least absolute shrinkage and selection operator |
| Bayes | Naïve Bayes |
| RF | Random forest |
| Glm | Generalized linear model |
| SVM | Supporting vector machine |
| GBM | Gradient boosting machine |
| PPI | Protein–protein interaction |
| FDR | False discovery rate |
| FC | Fold change |
| NES | Normalized enrichment score |
| ROC | Receiver operating characteristic |
| AUC | Area under the curve |

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13054-024-05000-3>.

Additional file 1.

Additional file 2.

Additional file 3.

Additional file 4.

Additional file 5.

Additional file 6.

Additional file 7.

Additional file 8.

Additional file 9.

Additional file 10.

Acknowledgements

We thank all the participants. We gratefully acknowledge the LuMing Biological Technology Co., Ltd. (Shanghai) for the support of 4D-DIA proteomic and global metabolic sequencing. We also acknowledged the iProteome Biological Technology Co., Ltd. (Shanghai) for the support of the PRM assay.

Author contributions

SZJ and WWB conceived the study; LMN, WWB, and SZJ designed the study; SZJ, WWB, and WXL supervised this project; SJ, SJF, SY, LS, DHL, LLL, CC, MW, and WXL performed the processing of samples and generation of data; LMN and XFX analyzed the data; LMN and XFX drafted the original paper; WWB, SZJ, and WXL reviewed the paper. All authors read and approved the final manuscript.

Funding

This work was supported by the National Key Research and Development Program of China (2021YFC2501800, 2023YFC0872500, and 2023YFC3043507), the National Natural Science Foundation of China (82072214), and the Science and Technology of Shanghai Committee (23Y31900100 and 21MC1930400). This work also was supported by the Shanghai Municipal Science and Technology Major Project (ZD2021CY001), and Shanghai New Three-year Action Plan for Public Health (GWVI-11.1-03).

Availability of data and materials

Data and research materials used in this study are available upon request to qualified researchers for purposes of replication, further analysis, and academic collaboration.

Declarations

Ethics approval and consent to participate

The lead author and guarantor certify that this manuscript presents a truthful, precise, and complete report of the study conducted; it ensures no significant

facets of the research have been excluded, and any deviations from the initial study design and registration are thoroughly accounted for and elucidated.

Competing interests

The authors declare no competing interests.

Author details

¹Shanghai Institute of Infectious Disease and Biosecurity, School of Public Health, Fudan University, Shanghai, China. ²Department of Emergency Medicine, Zhongshan Hospital, Fudan University, Shanghai, China. ³Department of Emergency Medicine, Minhang Hospital, Fudan University, Shanghai, China. ⁴Department of Respiratory Medicine, Pudong Hospital, Fudan University, Shanghai, China. ⁵School of Public Health, Fudan University, Shanghai, China. ⁶Department of Respiratory Medicine, Renji Hospital, Shanghai Jiaotong University School of Medicine, Shanghai, China. ⁷Institute of Emergency Rescue and Critical Care, Fudan University, Shanghai, China.

Received: 4 May 2024 Accepted: 24 June 2024

Published online: 02 July 2024

References

- Chang YJ, Yoo HJ, Kim SJ, Lee KH, Lim CM, Hong SB, Koh Y, Huh JW. A targeted metabolomics approach for sepsis-induced ARDS and its sub-phenotypes. *Crit Care*. 2023. <https://doi.org/10.1186/s13054-023-04552-0>.
- Namendys-Silva SA, Gutiérrez-Villaseñor A, Romero-González JP. Hospital mortality in mechanically ventilated COVID-19 patients in Mexico. *Intensive Care Med*. 2020;46(11):2086–8.
- Kneyber MCJ, Khemani RG, Bhalla A, Blokpoel RGT, Cruces P, Dahmer MK, Emeriaud G, Grunwell J, Ilia S, Katira BH, et al. Understanding clinical and biological heterogeneity to advance precision medicine in paediatric acute respiratory distress syndrome. *Lancet Resp Med*. 2023;11(2):197–212.
- Shen B, Yi X, Sun YT, Bi XJ, Du JP, Zhang C, Quan S, Zhang FF, Sun R, Qian LJ, et al. Proteomic and metabolomic characterization of COVID-19 patient sera. *Cell*. 2020;182(1):59–72.e15.
- Shu T, Ning WS, Wu D, Xu JQ, Han QQ, Huang MH, Zou XJ, Yang QY, Yuan Y, Bie YY, et al. Plasma proteomics identify biomarkers and pathogenesis of COVID-19. *Immunity*. 2020;53(5):1108–1122.e5.
- Gorman EA, O’Kane CM, McAuley DF. Acute respiratory distress syndrome 2022 2 Acute respiratory distress syndrome in adults: diagnosis, outcomes, long-term sequelae, and management. *Lancet*. 2022;400(10358):1157–70.
- Li MB, Parker BL, Pearson E, Hunter B, Cao J, Koay YC, Guneratne O, James DE, Yang J, Lal S, et al. Core functional nodes and sex-specific pathways in human ischaemic and dilated cardiomyopathy. *Nat Commun*. 2020. <https://doi.org/10.1038/s41467-020-16584-z>.
- Xu RH, Wang JR, Zhu QQ, Zou C, Wei ZH, Wang H, Ding Z, Meng MJ, Wei HM, Xia SJ, et al. Integrated models of blood protein and metabolite enhance the diagnostic accuracy for Non-Small Cell Lung Cancer. *Biomark Res*. 2023. <https://doi.org/10.1186/s40364-023-00497-2>.
- Ali RO, Quinn GM, Umarova R, Haddad JA, Zhang GY, Townsend EC, Scheuing L, Hill KL, Gewirtz M, Rampertaap S, et al. Longitudinal multi-omics analyses of the gut-liver axis reveals metabolic dysregulation in hepatitis C infection and cirrhosis. *Nat Microbiol*. 2023;8(1):12–27.
- Wozniak JM, Mills RH, Olson J, Caldera JR, Sepich-Poore GD, Carrillo-Terrazas M, Tsai CM, Vargas F, Knight R, Dorrestein PC, et al. Mortality risk profiling of *Staphylococcus aureus* bacteremia by multi-omic serum analysis reveals early predictive and pathogenic signatures. *Cell*. 2020;182(5):1311–1327.e1314.
- Wang Y, Huang XL, Li F, Xia XB, Jia N, Fu J, Liu S, Zhang J, Ge HY, Huang SY, et al. Serum-integrated omics reveal the host response landscape for severe pediatric community-acquired pneumonia. *Crit Care*. 2023. <https://doi.org/10.1186/s13054-023-04378-w>.
- Cao SR, Li HQ, Xin JY, Jin ZH, Zhang ZY, Li JW, Zhu YK, Su L, Huang PP, Jiang L, et al. Identification of genetic profile and biomarkers involved in acute respiratory distress syndrome. *Intensive care Med*. 2023;50(1):49–55.
- Ranieri VM, Rubenfeld GD, Thompson BT, Ferguson ND, Caldwell E, Fan E, Camporota L, Slutsky AS, Force ADT. Acute respiratory distress syndrome: the berlin definition. *Jama-J Am Med Assoc*. 2012;307(23):2526–33.
- Blume JE, Manning WC, Troiano G, Hornburg D, Figa M, Hesterberg L, Platt TL, Zhao XY, Cuaresma RA, Everley PA, et al. Rapid, deep and precise profiling of the plasma proteome with multi-nanoparticle protein corona. *Nat Commun*. 2020. <https://doi.org/10.1038/s41467-020-17033-7>.
- Xing XH, Cai LS, Ouyang JH, Wang F, Li ZM, Liu MX, Wang YC, Zhou Y, Hu E, Huang CL, et al. Proteomics-driven noninvasive screening of circulating serum protein panels for the early diagnosis of hepatocellular carcinoma. *Nat Commun*. 2023. <https://doi.org/10.1038/s41467-023-44255-2>.
- Tang SS, Wang YZ, Luo RK, Fang RD, Liu YF, Xiang H, Ran P, Tong YX, Sun MJ, Tan SB, et al. Proteomic characterization identifies clinically relevant subgroups of soft tissue sarcoma. *Nat Commun*. 2024;15(1):1381. <https://doi.org/10.1038/s41467-024-45306-y>.
- Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003;13(11):2498–504.
- Szklarczyk D, Gable AL, Nastou KC, Lyon D, Kirsch R, Pyysalo S, Doncheva NT, Legeay M, Fang T, Bork P, et al. The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res*. 2021;49(18):10800–10800.
- Chin CH, Chen SH, Wu HH, Ho CW, Ko MT, Lin CY. Identifying hub objects and sub-networks from complex interactome. *Bmc Syst Biol*. 2014;8:S11.
- Hänzelmann S, Castelo R, Guinney J. GSEA: gene set variation analysis for microarray and RNA-Seq data. *Bmc Bioinf*. 2013;14:1–15.
- Bojic S, Kotur-Stevuljevic J, Kalezic N, Jelic-Ivanovic Z, Stefanovic A, Palibrk I, Memon L, Kalaba Z, Stojanovic M, Simic-Ogrizovic S. Low paraoxonase 1 activity predicts mortality in surgical patients with sepsis. *Dis Markers*. 2014;2014: 427378.
- Lin HS, Liu QSJ, Zhao L, Liu ZQ, Cui HH, Li PH, Fan HJ, Guo LQ. Circulating pulmonary-originated epithelial biomarkers for acute respiratory distress syndrome: a systematic review and meta-analysis. *Int J Mol Sci*. 2023;24(7):6090.
- Zhao JP, Yu H, Liu YD, Gibson SA, Yan ZQ, Xu X, Gaggari A, Li PK, Li CL, Wei S, et al. Protective effect of suppressing STAT3 activity in LPS-induced acute lung injury. *Am J Physiol-Lung C*. 2016;311(5):L868–80.
- Zeng Y, Adamson RH, Curry FRE, Tarbell JM. Sphingosine-1-phosphate protects endothelial glycocalyx by inhibiting syndecan-1 shedding. *Am J Physiol-Heart C*. 2014;306(3):H363–72.
- Uchimido R, Schmidt EP, Shapiro NI. The glycocalyx: a novel diagnostic and therapeutic target in sepsis. *Crit Care*. 2019;23:1–12.
- Piotti A, Novelli D, Meessen JM, Ferlicca D, Coppolecchia S, Marino A, Salati G, Savioli M, Grasselli G, Bellani G, et al. Endothelial damage in septic shock patients as evidenced by circulating syndecan-1, sphingosine-1-phosphate and soluble VE-cadherin: a substudy of ALBIOS. *Crit Care*. 2021;25(1):1–12.
- Natarajan V, Dudek SM, Jacobson JR, Moreno-Vinasco L, Huang LS, Abassi T, Mathew B, Zhao YT, Wang LC, Bittman R, et al. Sphingosine-1-phosphate, FTY720, and sphingosine-1-phosphate receptors in the pathobiology of acute lung injury. *Am J Resp Cell Mol*. 2013;49(1):6–17.
- McVerry BJ, Peng XQ, Hassoun PM, Sammani S, Simon BA, Garcia JGN. Sphingosine 1-phosphate reduces vascular leak in murine and canine models of acute lung injury. *Am J Resp Crit Care*. 2004;170(9):987–93.
- Peng XQ, Hassoun PM, Sammani S, McVerry BJ, Burne MJ, Rabb H, Pearce D, Tuder RM, Garcia JGN. Protective effects of sphingosine 1-phosphate in murine endotoxin-induced inflammatory lung injury. *Am J Resp Crit Care*. 2004;169(11):1245–51.
- Sun XG, Ma SF, Wade MS, Acosta-Herrera M, Villar J, Pino-Yanes M, Zhou T, Liu B, Belvitch P, Moitra J, et al. Functional promoter variants in sphingosine 1-phosphate receptor 3 associate with susceptibility to sepsis-associated acute respiratory distress syndrome. *Am J Physiol-Lung C*. 2013;305(7):L467–77.
- Sammani S, Moreno-Vinasco L, Mirzapoiazova T, Singleton PA, Chiang ET, Evenoski CL, Wang T, Mathew B, Husain A, Moitra J, et al. Differential effects of sphingosine 1-phosphate receptors on airway and vascular barrier function in the murine lung. *Am J Resp Cell Mol*. 2010;43(4):394–402.
- Naz F, Arish M. Battling COVID-19 pandemic: sphingosine-1-phosphate analogs as an adjunctive therapy? *Front Immunol*. 2020;11:546310.

33. Singleton PA, Dudek SM, Ma SF, Garcia JGN. Transactivation of sphingosine 1-phosphate receptors is essential for vascular barrier regulation—novel role for hyaluronan and CD44 receptor family. *J Biol Chem*. 2006;281(45):34381–93.
34. Winkler MS, Nierhaus A, Holzmann M, Mudersbach E, Bauer A, Robbe L, Zahrte C, Geffken M, Peine S, Schwedhelm E, et al. Decreased serum concentrations of sphingosine-1-phosphate in sepsis. *Crit Care*. 2015;19:1–8.
35. Cyr A, Zhong YJ, Reis SE, Namas RA, Amoscato A, Zuckerbraun B, Sperry J, Zamora R, Vodovotz Y, Billiar TR. Analysis of the plasma metabolome after trauma, novel circulating sphingolipid signatures, and in-hospital outcomes. *J Am Coll Surgeons*. 2021;232(3):276–287e1.
36. Xiong YB, Piao WJ, Brinkman CC, Li L, Kulinski JM, Olivera A, Cartier A, Hla T, Hippen KL, Blazar BR, et al. CD4 T cell sphingosine 1-phosphate receptor (S1PR)1 and S1PR4 and endothelial S1PR2 regulate afferent lymphatic migration. *Sci Immunol*. 2019;4(33):eaav1263.
37. Kilpatrick LE, Sun S, Mackie DM, Baik F, Li HY, Korchak HM. Regulation of TNF mediated antiapoptotic signaling in human neutrophils: role of δ -PKC and ERK1/2. *J Leukocyte Biol*. 2008;83(3):797–797.
38. Wan P, Zhang SM, Ruan ZH, Liu XL, Yang G, Jia YL, Li YK, Pan P, Wang WB, Li G, et al. AP-1 signaling pathway promotes pro-IL-1 β transcription to facilitate NLRP3 inflammasome activation upon influenza A virus infection. *Virulence*. 2022;13(1):502–13.
39. Amunugama K, Pike DP, Ford DA. The lipid biology of sepsis. *J Lipid Res*. 2021;62:100090.
40. Drobnik W, Liebisch G, Audebert FX, Fröhlich D, Glück T, Vogel P, Rothe G, Schmitz G. Plasma ceramide and lysophosphatidylcholine inversely correlate with mortality in sepsis patients. *J Lipid Res*. 2003;44(4):754–61.
41. Park DW, Kwak DS, Park YY, Chang Y, Huh JW, Lim CM, Koh Y, Song DK, Hong SB. Impact of serial measurements of lysophosphatidylcholine on 28-day mortality prediction in patients admitted to the intensive care unit with severe sepsis or septic shock. *J Crit Care*. 2014;29(5):882e.5–882e.11.
42. Ning L, Zou SS, Bo W, Lin HQ. Targeting immunometabolism against acute lung injury. *Clin Immunol*. 2023;249:109289.
43. Karagiannis F, Peukert K, Surace L, Michla M, Nikolka F, Fox M, Weiss P, Feuerborn C, Maier P, Schulz S, et al. Impaired ketogenesis ties metabolism to T cell dysfunction in COVID-19. *Nature*. 2022;609(7928):801–7.
44. Osuchowski MF, Winkler MS, Skirecki T, Cajander S, Shankar-Hari M, Lachmann G, Monneret G, Venet F, Bauer M, Brunkhorst FM, et al. The COVID-19 puzzle: deciphering pathophysiology and phenotypes of a new disease entity. *Lancet Resp Med*. 2021;9(6):622–42.
45. O'Neill LAJ, Kishton RJ, Rathmell J. A guide to immunometabolism for immunologists. *Nat Rev Immunol*. 2016;16(9):553–65.
46. Wang L, Tang Y, Tang JJ, Liu X, Zi SF, Li SL, Chen HB, Liu AR, Huang W, Xie JF, et al. Endothelial cell-derived extracellular vesicles expressing surface VCAM1 promote sepsis-related acute lung injury by targeting and reprogramming monocytes. *J Extracell Vesicles*. 2024;13(3):e12423.
47. Eisen DP, Dean MM, Boermeester MA, Fidler KJ, Gordon AC, Kronborg G, Kun JFJ, Lau YL, Payeras A, Valdimarsson H, et al. Low serum mannose-binding lectin level increases the risk of death due to pneumococcal infection. *Clin Infect Dis*. 2008;47(4):510–6.
48. Osthoff M, Rovó A, Stern M, Danner D, Gratwohl A, Tichelli A, Trendelenburg M. Mannose-binding lectin levels and major infections in a cohort of very long-term survivors after allogeneic stem cell transplantation. *Haematol-Hematol J*. 2010;95(8):1389–96.
49. Mullighan CG, Heatley SL, Danner S, Dean MM, Doherty K, Hahn U, Bradstock KF, Minchinton R, Schwarer AP, Szer J, et al. Mannose-binding lectin status is associated with risk of major infection following myeloablative sibling allogeneic hematopoietic stem cell transplantation. *Blood*. 2008;112(5):2120–8.
50. Beltinger C. LDHA and LDHB are dispensable for aerobic glycolysis in neuroblastoma cells while promoting their aggressiveness. *J Biol Chem*. 2019;294(1):66–66.
51. Lagresle-Peyrou C, Luce S, Ouchani F, Soheili TS, Sadek H, Chouteau M, Durand A, Pic I, Majewski J, Brouzes C, et al. X-linked primary immunodeficiency associated with hemizygous mutations in the moesin (MSN) gene. *J Allergy Clin Immunol*. 2016;138(6):1681–1689e.8.
52. Gabarin RS, Li MS, Zimmel PA, Marshall JC, Li YM, Zhang HB. Intracellular and extracellular lipopolysaccharide signaling in sepsis: avenues for novel therapeutic strategies. *J Innate Immun*. 2021;13(6):323–32.
53. He LG, Liao JQ, Liu Z, Wang T, Zhou Y, Wang TF, Lei BY, Zhou GQ. Multi-omic analysis of mandibuloacral dysplasia type A patient iPSC-derived MSC senescence reveals miR-311 as a novel biomarker for MSC senescence. *Hum Mol Genet*. 2023;32(19):2872–86.
54. Agren M, Litman T, Eriksen JO, Schjerling P, Bzorek M, Gjerdrum LMR. Gene expression linked to reepithelialization of human skin wounds. *Int J Mol Sci*. 2022;23(24):15746.
55. Wang LD, Tan H, Huang YL, Guo MY, Dong YX, Liu CX, Zhao H, Liu Z. TAGLN2 promotes papillary thyroid carcinoma invasion via the Rap1/PI3K/AKT axis. *Endocr-Relat Cancer*. 2023;30(1):e210352.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.