



OPEN ACCESS

EDITED BY

Alexandre Keiji Tashima,
Federal University of São Paulo, Brazil

REVIEWED BY

José Thalles Lacerda,
Federal University of Pernambuco, Brazil
Chen-Yang Su,
McGill University and Génome Québec
Innovation Centre, Canada

*CORRESPONDENCE

Zhihai Chen

✉ chenzhihai0001@126.com

Ling Lin

✉ linling4012@163.com

[†]These authors have contributed
equally to this work and share
first authorship

RECEIVED 27 August 2025

ACCEPTED 22 October 2025

PUBLISHED 05 November 2025

CITATION

Zhao C, Ge Z, Wang R, Xu Y, Zhang T,
Jiang Z, Liu L, Lin L and Chen Z (2025) Serum
proteomics and machine learning identify
PSMD11 as a prognostic biomarker in severe
fever with thrombocytopenia syndrome.
Front. Immunol. 16:1693946.
doi: 10.3389/fimmu.2025.1693946

COPYRIGHT

© 2025 Zhao, Ge, Wang, Xu, Zhang, Jiang, Liu,
Lin and Chen. This is an open-access article
distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The
use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Serum proteomics and machine learning identify PSMD11 as a prognostic biomarker in severe fever with thrombocytopenia syndrome

Chenxi Zhao^{1†}, Ziruo Ge^{1†}, Ranran Wang^{1†}, Yanli Xu²,
Tingyu Zhang¹, Zhouling Jiang¹, Lu Liu¹, Ling Lin^{2*}
and Zhihai Chen^{1*}

¹National Key Laboratory of Intelligent Tracking and Forecasting for Infectious Diseases, Beijing Ditan Hospital, Capital Medical University, Beijing, China, ²Department of Infectious Diseases, Yantai Qishan Hospital, Yantai, Shandong, China

Background: Severe fever with thrombocytopenia syndrome (SFTS) is an emerging tick-borne viral disease associated with high mortality. This study aimed to characterize serum proteomic signatures linked to adverse outcomes and to identify prognostic biomarkers with potential translational value for patient management.

Methods: Serum samples from 55 survivors, 32 non-survivors, and 10 healthy controls were analyzed by data-independent acquisition-based proteomics. Differential abundance analysis, Gene Ontology (GO) enrichment, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis, and protein-protein interaction (PPI) network analyses with Markov clustering were conducted to characterize disease-associated proteins. XGBoost and Random Forest machine learning models were applied to prioritize candidate biomarkers, and discriminative performance was evaluated by the receiver operating characteristic (ROC) curve. Spearman correlation analyses were further used to examine associations between candidate proteins, clinical laboratory indicators, and viral load.

Results: Non-survivors exhibited 642 differentially abundant proteins (DAPs) compared with survivors. Functional enrichment and PPI network analyses revealed a proteasome-centered module overrepresented in non-survivors. XGBoost and Random Forest consistently prioritized four candidate biomarkers (PSMD11, IL1RL1, PSMC4, and IFIH1) with areas under the ROC curve of 0.847, 0.847, 0.843, and 0.791, respectively. PSMD11 emerged as the strongest predictor of adverse outcome and showed strong correlations with markers of organ injury and dysfunction such as lactate dehydrogenase ($r = 0.77$), thrombin time ($r = 0.76$), aspartate aminotransferase ($r = 0.75$), hydroxybutyrate dehydrogenase ($r = 0.74$), viral load ($r = 0.63$), and platelet count ($r = -0.57$) (all $p < 0.001$).

Conclusions: This study identified a proteasome-centered signature associated with adverse outcomes in SFTS, with PSMD11 emerging as a key prognostic biomarker. Its strong correlations with viral load and multi-organ injury support potential utility for early risk stratification and prognostic prediction, while also providing mechanistic insights into disease progression and a foundation for future translational research and therapeutic development.

KEYWORDS

severe fever with thrombocytopenia syndrome, proteomics, machine learning, PSMD11, SHAP

Introduction

Severe Fever with Thrombocytopenia Syndrome (SFTS) is an emerging tick-borne infectious disease caused by *Bandavirus dabiense* (family *Phenuiviridae*) (1), commonly known as SFTS virus (SFTSV), which was first identified in China in 2009 (2). Since then, SFTS has been increasingly reported across East and Southeast Asia, including South Korea, Japan, Vietnam, Thailand, and Pakistan, with both annual cases and geographic range continuing to expand (3–7). SFTS is characterized by acute high fever, thrombocytopenia, leukopenia, hemorrhagic manifestations, and gastrointestinal symptoms, and can progress to multiorgan dysfunction and death in severe cases (8). According to a recent meta-analysis, the overall pooled case-fatality rate was 7.80% (95% CI, 7.01–8.69%) (9), emphasizing the considerable mortality burden that persists despite improvements in supportive care. Favipiravir was approved for the treatment of SFTS for the first time in Japan in June 2024 (10). However, randomized evidence remains limited and there are still no globally licensed vaccines or widely accessible targeted antivirals for SFTS (11). The WHO's 2024 prioritization classified *Phenuiviridae* as high risk for Public Health Emergency of International Concern (PHEIC) and identified *Bandavirus dabiense* as a priority pathogen for research and development (12). These underscore the urgency of elucidating SFTS pathogenesis and accelerating biomarker-informed strategies for prognosis and therapeutic development.

Accumulating evidence indicates that dynamic changes in circulating biomarkers—including interleukin-6 (IL-6), IL-10, interferon- α (IFN- α), lactate dehydrogenase (LDH), ferritin, and C-C motif chemokine 20 (CCL20)—track with disease severity and mortality risk, reflecting the importance of host–pathogen interactions (13–15). High-throughput proteomics has accelerated both fundamental and applied research on emerging and re-emerging viral pathogens (16). By enabling quantitative assessment of protein abundance, proteomics provides direct insights into disease progression (17). Serum and plasma proteomic profiling has been widely applied to SARS-CoV-2, revealing host–pathogen interactions, elucidating the molecular mechanisms underlying COVID-19 pathology, and identifying candidate biomarkers for

risk stratification and therapeutic development (18–20). Machine learning (ML) has become an essential tool in clinical and translational research, facilitating the identification of diagnostic, prognostic, and predictive biomarkers (21). A recent study developed a machine learning model based on circulating protein signatures that accurately predicted COVID-19 severity, highlighting the value of proteomics-driven ML strategies for clinical risk assessment (22). Integrating proteomics with ML approaches offers a promising strategy to generate clinically relevant insights into emerging infectious diseases (23). Nevertheless, applications of such integrative analyses to SFTS remain scarce.

In this study, our objective was to integrate serum proteomics with interpretable ML to delineate molecular signatures of SFTS, identify candidate biomarkers with translational potential, and illuminate disease-relevant pathways. By providing a systems-level view of host–pathogen interactions, our work aims to improve risk stratification and inform the development of targeted interventions, thereby contributing to better clinical management of SFTS.

Methods

Study design and patients

This prospective cohort study was conducted at Yantai Qishan Hospital between May and September 2024. A total of 87 patients with laboratory-confirmed SFTS were enrolled at admission, comprising 55 survivors and 32 non-survivors. SFTS diagnosis was confirmed by real-time quantitative reverse transcription polymerase chain reaction (RT-qPCR) detection of SFTSV RNA. Inclusion criteria were age ≥ 18 years and admission within 7 days of symptom onset. Exclusion criteria were: co-infection with other acute or chronic infections; receipt of antitumor therapy; inability to obtain required samples; or incomplete essential clinical data. Ten age- and sex- matched healthy controls (HC) with negative SFTSV tests were included. The primary endpoint was defined as either recovery with discharge or in-hospital death.

The study was conducted in accordance with the principles of the Declaration of Helsinki and was approved by the Ethics

Committee of Beijing Ditan Hospital, Capital Medical University (No. DTEC-KY2022-022). Written informed consent was obtained from all participants or their relatives.

Clinical sample collection and definitions

For patients with SFTS, peripheral venous blood was collected within 24 hours of hospital admission using silica-coated clot activator tubes without gel separators. A total of 97 serum samples were obtained from 87 patients with SFTS and 10 healthy donors. Samples were centrifuged at $2,000 \times g$ for 10 minutes within 2 hours of collection to ensure complete serum separation. The resulting serum was aliquoted and stored at -80°C until analysis. Serum viral load was measured using a commercial RT-qPCR kit (Daan, Guangzhou, China) according to the manufacturer's instructions.

The acute phase, characterized by high fever and systemic symptoms, was defined as days 1–7 from symptom onset (24).

Data-independent acquisition proteomics

Serum samples were processed using a standardized DIA proteomics workflow. Briefly, 15 μL of serum was used as the starting volume and combined with 75 μL of Buffer 2 from the OmniProt Kit (OmniProt, China), followed by centrifugation at 4,000 rpm for 2 minutes at room temperature. The supernatant was incubated with 90 μL of depletion reagent at 32°C for 1 hour with gentle agitation (220 rpm) to remove high-abundance proteins, followed by centrifugation at $17,000 \times g$ for 10 minutes. The pellet was washed twice with Buffer 3 and re-centrifuged under the same conditions. The resulting proteins were solubilized in lysis buffer (6 M urea, 2 M thiourea), reduced with tris (2-carboxyethyl) phosphine (0.2 M, 4 μL , 32°C , 30 minutes), and alkylated with iodoacetamide (0.8 M, 4 μL , 32°C , 30 minutes). Proteins were digested overnight with sequencing-grade trypsin at 32°C with gentle shaking (220 rpm). The reaction was quenched with 30 μL of 10% trifluoroacetic acid, and peptides were centrifuged at $17,000 \times g$ for 10 minutes. The supernatant was transferred to a new tube, and the pH was adjusted to 2. Peptides were desalted using SOLA μ solid-phase extraction plates (Thermo Fisher Scientific, San Jose, USA) according to the manufacturer's instructions. The eluates were vacuum dried at 40°C , reconstituted in 0.1% formic acid, and quantified at 280 nm using a NanoDrop spectrophotometer. Peptide concentrations were adjusted to 0.2 $\mu\text{g}/\mu\text{L}$ to ensure equal loading across all samples. For DIA analysis, 1 μL of each digest (equivalent to 200 ng of peptides) was injected into a Vanquish Neo UHPLC system coupled to an Orbitrap Astral mass spectrometer (Thermo Fisher Scientific, San Jose, USA). Peptides were separated using a 24-minute liquid chromatography gradient. Raw DIA data were processed with DIA-NN (version 1.8.1) in single-pass (library-free) mode for protein identification and quantification.

Samples were analyzed in 11 experimental batches under strictly standardized procedures, including identical reagent lots, uniform sample preparation, and consistent instrument calibration. Each batch contained one pooled quality control (QC) sample and one biological replicate sample to monitor intra- and inter-batch consistency. The pooled QC sample was prepared by mixing equal aliquots from all experimental samples.

Proteomic data analysis

Proteins were identified and quantified using library-free DIA-NN searches against a human reference database. Proteins with more than 70% missing values across all samples were excluded prior to downstream analysis, and the remaining missing values were imputed using the k-nearest neighbors (KNN) algorithm after quality control. Outliers were defined as values exceeding three times the interquartile range (IQR) above the upper quartile or below the lower quartile. Protein intensities were $\log_2$ -transformed and Z-score normalized. Differential protein abundance was analyzed using the limma package in R, applying moderated t-statistics with empirical Bayes shrinkage. *p*-values were adjusted using the Benjamini-Hochberg method to control the false discovery rate (FDR). Differentially abundant proteins (DAPs) were defined as $\text{FDR} < 0.05$ and a fold change (FC) > 2 or < 0.5 (i.e., absolute $\log_2\text{FC} > 1$). Gene Ontology (GO) annotation categorized DAPs into biological process (BP), cellular component (CC), and molecular function (MF). Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment was performed to identify significantly enriched pathways among DAPs. Protein-protein interaction (PPI) networks via the search tool for retrieval of interacting genes/proteins (STRING) database (version 12.0) to explore interactions among proteins, a combined score > 0.7 was applied to ensure a high-confidence interaction threshold in the identified interactions and visualized in Cytoscape (version 3.10.3).

Machine learning algorithms

Discriminative proteins among the DAPs were identified using two supervised machine learning algorithms, eXtreme Gradient Boosting (XGBoost) and Random Forest. The XGBoost model was trained using a 10-fold cross-validation procedure with a fixed random seed to ensure reproducibility. The model was optimized for binary classification based on the log-loss objective function, and hyperparameters were tuned to balance model complexity and mitigate overfitting. Feature importance was evaluated using SHapley Additive exPlanations (SHAP), with the mean absolute SHAP value was used to rank the relative importance of DAPs. The Random Forest model was trained using the caret framework with 10-fold cross-validation and a fixed random seed to ensure reproducibility. Variable importance for each protein was determined by the mean decrease in the Gini index, averaged over folds and normalized for comparison across features.

Statistical analysis

Clinical continuous variables were summarized as mean $\pm$ standard deviation (SD) or median (IQR) and compared with Student's t-test or Mann–Whitney U test. Categorical variables were presented as counts (percentages) and compared using the χ^2 test or Fisher's exact test. Multivariate patterns in proteomic data were explored by principal component analysis (PCA). Predictive performance of selected proteins was evaluated using receiver operating characteristic (ROC) and precision-recall (PR) analyses, reporting the area under the ROC curve (AUC) and the area under the PR curve (AUPRC) with 95% confidence intervals (CI). Additional performance metrics, including sensitivity, specificity, precision, accuracy, and F1 score, were visualized using radar charts. Correlations between selected DAPs and clinical laboratory parameters were analyzed using Spearman's rank correlation and partial Spearman correlation controlling for age. All statistical analyses were performed in R software (version 4.4.2) using the following packages: limma, xgboost, SHAPforxgboost, randomForest, caret, pROC, PRROC, clusterProfiler, and ggplot2. Statistical significance was defined as p value (or FDR) < 0.05 .

Results

Clinical characteristics of participants

The study enrolled 87 patients with SFTS, comprising 55 survivors (SA) and 32 non-survivors (NS), along with 10 age- and sex-matched HC. The SA had an average age of 65.38 ± 9.55 years, while the NS had significantly higher age of 70.6 ± 10.3 years ($p = 0.023$). No significant differences in sex distribution were observed between SA and NS ($p = 0.868$), or between SFTS patients and HC ($p = 0.690$). Compared with HC, patients with SFTS had significantly lower white blood cell (WBC) counts ($p < 0.001$), platelet counts ($p < 0.001$), and albumin levels ($p < 0.001$), as well as higher alanine aminotransferase (ALT) ($p < 0.001$), aspartate aminotransferase (AST) ($p < 0.001$), and creatinine levels ($p < 0.001$). Within the SFTS cohort, NS exhibited significantly lower platelet counts ($p < 0.001$) and albumin levels ($p < 0.001$), and higher AST ($p < 0.001$) and creatinine ($p < 0.001$) compared with SA. No significant differences were found in WBC counts ($p = 0.403$), hemoglobin ($p = 0.758$), mean platelet volume (MPV) ($p = 0.867$), or total bilirubin (TBil) ($p = 0.060$) between the survivor and non-survivor groups (Table 1).

Identification of differentially abundant proteins

A total of 97 serum samples were collected from 87 patients with SFTS and 10 HC. DIA proteomics quantified 5,541 proteins. PCA demonstrated distinct clustering among SA, NS, and HC (Figure 1a). Venn analysis identified 108 DAPs shared across NS vs SA, SA vs HC, and NS vs HC comparisons (Figure 1b).

Hierarchical clustering of DAPs revealed a clear pattern, with upregulated proteins enriched in NS and downregulated proteins more prevalent in HC (Figure 1c). In total, 642 DAPs were identified between NS and SA, comprising 617 upregulated and 25 downregulated proteins. The top five upregulated DAPs in the volcano plot are CCL20, CDKN1A interacting zinc finger protein 1 (CIZ1), proteasome 26S subunit, ATPase 4 (PSMC4), proteasome 26S subunit, non-ATPase 11 (PSMD11), and nuclear receptor corepressor 1 (NCOR1) (Figure 1d). The HC group had a total of 1316 DAPs, including 1234 upregulated and 82 downregulated proteins, compared with the NS and SA groups (Figure 1e).

GO enrichment and KEGG pathway analysis

To explore the potential biological functions and pathways associated with the observed protein abundance differences, GO and KEGG enrichment analyses were performed. DAPs between NS and SA were significantly enriched for BP terms related to ribonucleoprotein complex and ribosome biogenesis, and for CC terms including the preribosome and proteasome accessory complex (Figure 2a; Supplementary Table S1). KEGG analysis highlighted several pathways, with amyotrophic lateral sclerosis (ALS) ranking highest by gene ratio (Figure 2b; Supplementary Table S2). These findings implicate perturbations in protein homeostasis and RNA/protein biogenesis in poor outcomes.

Network-based clustering and KEGG pathway enrichment of candidate proteins

We constructed a STRING PPI network by integrating the top 50 proteins ranked by SHAP values from the XGBoost model and the top 50 proteins ranked by feature importance from the Random Forest classifier, yielding 85 nodes and 178 edges (clustering coefficient = 0.379; PPI enrichment $p < 1.0 \times 10^{-16}$). The resulting PPI network was clustered using the Markov Clustering (MCL) algorithm implemented in Cytoscape, revealing 11 distinct protein clusters (Supplementary Table S3). The most prominent cluster contained PSMD11 and multiple 26S proteasome subunits (PSMC4, PSMC2, PSMC3, PSMC6, PSMD7), pointing to coordinated proteasome involvement (Figure 3a). A Sankey plot mapped clusters to significantly enriched KEGG pathways, with several PSMD-containing clusters converging on the proteasome pathway (Figure 3b). These results suggest proteasome-associated pathways may play a central role in SFTS pathogenesis.

Screening proteomic biomarkers via machine learning

To explore potential prognostic biomarkers associated with poor outcomes in patients with SFTS, we trained XGBoost and Random Forest models using DAPs between the NS and SA. The

TABLE 1 The baseline clinical and laboratory characteristics of the participants.

	Health control (n=10)	SFTS patients			<i>p</i> ^b value	<i>p</i> ^a value
		Total (n=87)	Survivor group (n=55)	Non-survivor group (n=32)		
Age (years)	61.4 ± 8.6	67.3 ± 10.1	65.38 ± 9.55	70.6 ± 10.3	0.023	0.068
Male, n (%)	5 (50)	33 (37.9)	20 (36.4)	13 (40.6)	0.868	0.690
WBC (×10 ⁹ /L)	5.8 (5.2, 6.3)	2.1 (1.6, 3.1)	2.0 (1.5, 2.8)	2.1 (1.7, 3.1)	0.403	<0.001
Hemoglobin (g/L)	135.5 (130.0, 137.0)	143.0 (133.0, 152.0)	141.0 (133.0, 152.0)	144.5 (133.0, 151.8)	0.758	0.034
Platelet (×10 ⁹ /L)	271.5 (237.5, 290.0)	65.0 (51.0, 82.0)	74.0 (61.0, 88.0)	51.5 (43.8, 65.8)	<0.001	<0.001
MPV (fL)	9.83 ± 0.66	10.46 ± 1.05	10.47 ± 1.13	10.44 ± 0.90	0.867	0.018
ALT (U/L)	14.0 (11.2, 25.1)	52.2 (34.2, 98.1)	43.3 (33.2, 90.2)	66.9 (46.0, 137.5)	0.059	<0.001
AST (U/L)	17.9 (15.2, 24.0)	122.3 (55.5, 213.7)	71.0 (49.1, 141.5)	170.2 (114.9, 456.9)	<0.001	<0.001
Albumin (g/L)	46.6 (45.6, 48.1)	34.9 (30.8, 37.4)	35.8 ± 4.2	31.8 ± 4.4	<0.001	<0.001
TBil (umol/L)	12.3 (11.0, 15.2)	9.3 (8.2, 11.3)	9.0 (7.9, 10.9)	10.5 (8.5, 14.7)	0.060	0.040
Creatinine (umol/L)	50.4 (47.0, 53.0)	74.0 (60.5, 94.2)	65.1 (57.0, 78.7)	99.9 (75.8, 152.8)	<0.001	<0.001

p^a: Comparisons between healthy controls and patients with SFTS.
p^b: Comparisons between survivors and non-survivors among patients with SFTS.
WBC, white blood cell. ALT, alanine aminotransferase. AST, aspartate aminotransferase. TBil, total bilirubin. MPV, mean platelet volume.

XGBoost model interpreted via SHAP ranked IL1RL1, CA3, PSMD11, EIF2B5, IFIH1, SRSF3, SAA2, PSMC6, PSMC4, and NFKB1 as the top features (Figure 4a), while the Random Forest model highlighted IL1RL1, PSMD11, PSMC6, IFIH1, ATP6V1D, CCL20, PSMD13, SRSF3, BMP2K and LECT2 (Figure 4b). ROC analysis of four candidates (IL1RL1, PSMD11, PSMC4, IFIH1) showed AUCs of 0.847 (95% CI: 0.761–0.932), 0.847 (95% CI: 0.766–0.928), 0.843 (95% CI: 0.763–0.924), and 0.791 (95% CI: 0.681–0.902), respectively (Figure 4c). PR curves yielded AUPRCs of 0.771 (95% CI: 0.721–0.821), 0.731 (95% CI: 0.681–0.781), 0.747 (95% CI: 0.697–0.797), and 0.762 (95% CI: 0.712–0.812), respectively (Figure 4d). Comprehensive metric comparisons including AUC, AUPRC, sensitivity, specificity, precision, accuracy, and F1 score are shown in a radar chart (Figure 4e). Furthermore, violin plots demonstrated significantly elevated abundant of all four proteins in NS compared with SA (Figures 4f–i, *p* < 0.0001). Age-adjusted logistic regression analysis was performed to verify the robustness of these associations (Supplementary Figure S1).

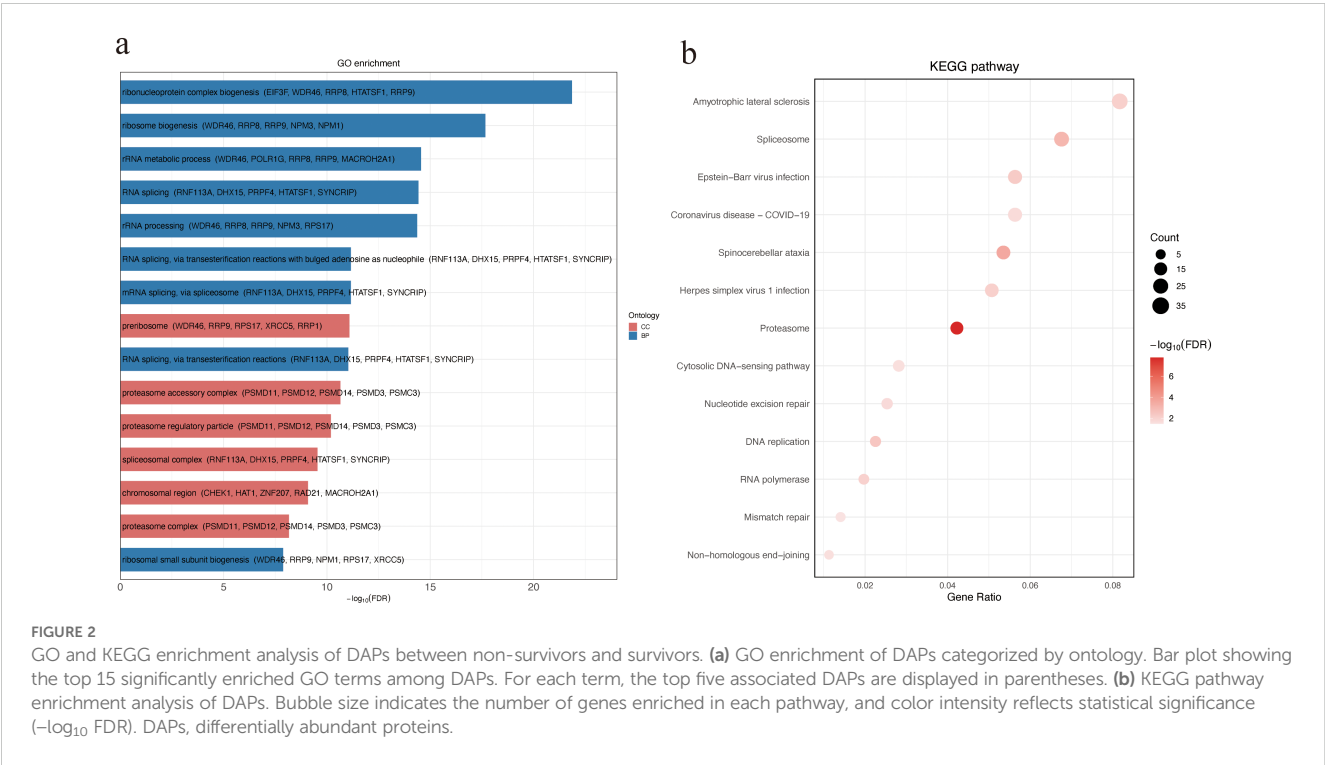
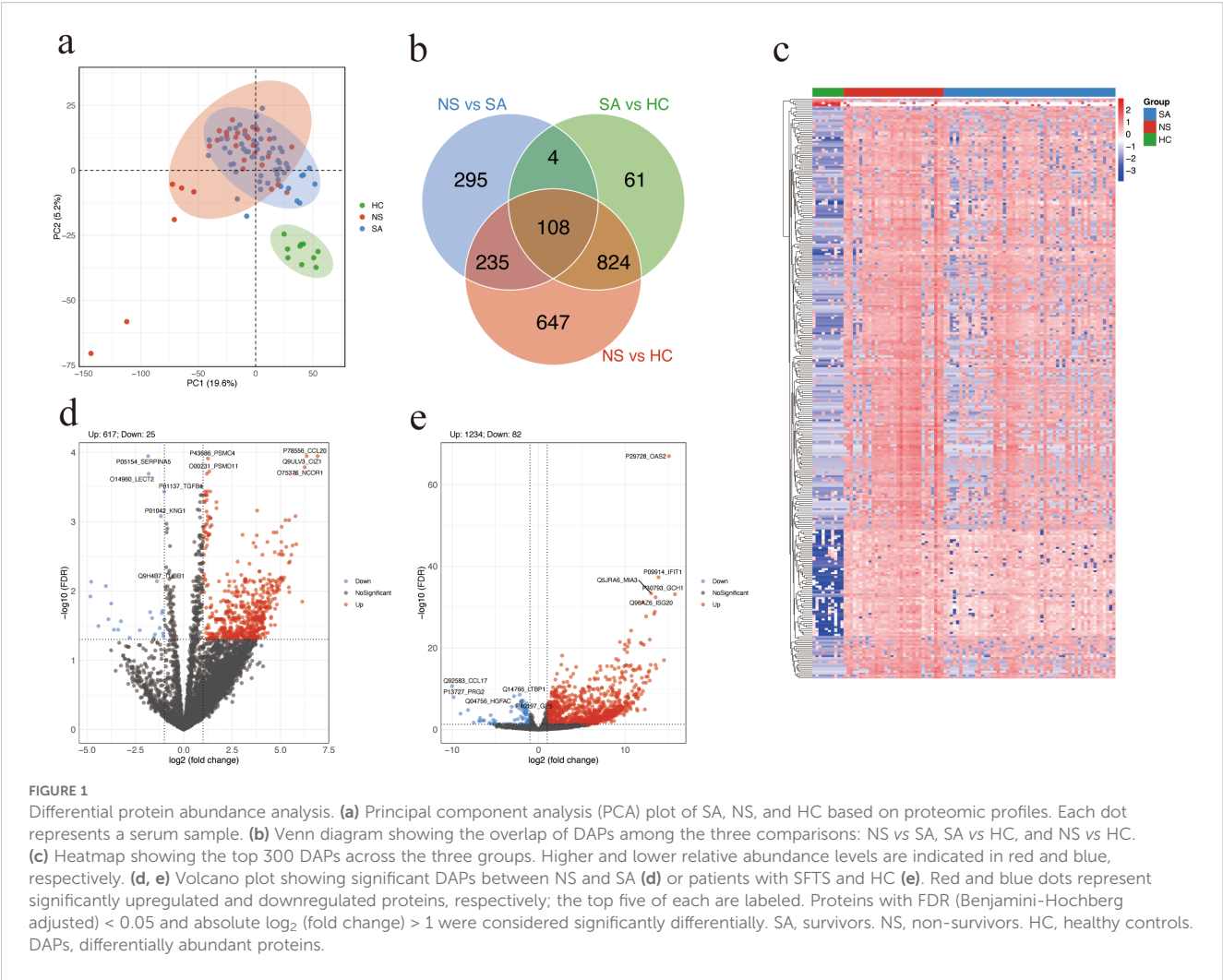
Correlation of candidate proteins with clinical indicators

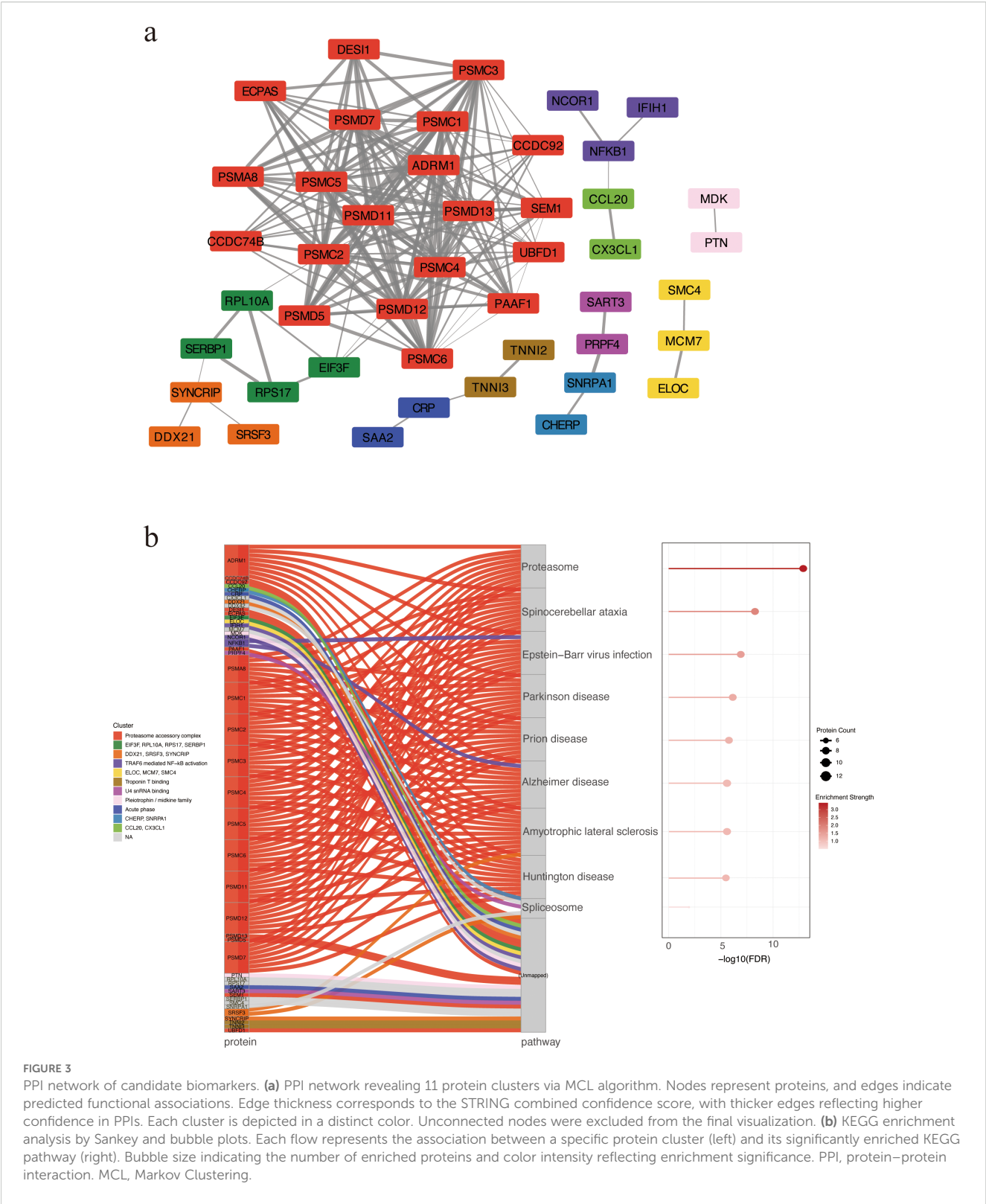
Age-adjusted partial Spearman correlation analysis was performed to assess associations between candidate protein levels and clinical laboratory parameters. Correlation heatmaps revealed broad associations of the four candidates (PSMD11, IL1RL1, IFIH1, PSMC4) with clinical laboratory parameters (Figure 5a). PSMD11 showed the strongest and most consistent relationships with hepatic dysfunction, myocardial injury, coagulation dysfunction,

inflammation, and viral load. The strongest positive correlations were observed with LDH (*r* = 0.77), thrombin time (TT; *r* = 0.76), AST (*r* = 0.75), and hydroxybutyrate dehydrogenase (HBDH; *r* = 0.74). PSMD11 showed moderate to strong positive correlations with viral load (*r* = 0.63), alpha-L-fucosidase (AFU; *r* = 0.63), adenosine deaminase (ADA; *r* = 0.62), high-sensitivity cardiac troponin I (hs-cTnI; *r* = 0.61), IL-10 (*r* = 0.59), ALT (*r* = 0.59), creatine kinase (CK; *r* = 0.57), activated partial thromboplastin time (APTT; *r* = 0.56), and D-dimer (*r* = 0.53). Conversely, PSMD11 showed significant negative correlations with platelet count (*r* = −0.57), prealbumin (*r* = −0.55), and albumin (*r* = −0.52) (all *p* < 0.001; Figure 5b). These findings support a close linkage between PSMD11 and disease severity and multisystem pathophysiology in SFTS. Moreover, we performed a multivariable logistic regression model including age and these laboratory indicators, PSMD11 remained significantly associated with the outcome (OR = 4.20, 95% CI = 1.19–19.5, *p* = 0.039; Supplementary Table S4).

Discussion

In this prospective clinical cohort, we employed DIA-based serum proteomics and integrated network-based analyses. By combining PPI mapping with MCL clustering, we delineated a coherent proteasome-associated module enriched in DAPs. Multiple lines of evidence from differential abundance, enrichment analyses, model explainability, and network mapping implicated a proteasome-centered axis as the dominant correlate of adverse outcomes in SFTS. This finding provides important mechanistic insight into how proteasome dysregulation may contribute to disease progression. To mitigate potential bias from





the modest sample size and to reduce model overfitting, we further applied two complementary supervised ML algorithms-XGBoost and Random Forest for feature prioritization within DAPs and used interpretable feature-attribution metrics to rank their importance (25). Because age differed significantly between survivors and non-survivors, we further performed sensitivity analyses to account for this potential confounder. Four proteins (PSMD11, IL1RL1, IFIH1, and PSMC4) consistently exhibited the most pronounced and outcome-relevant alterations, with PSMD11 emerging as the most robust and stable signal. These results link proteomic alterations to

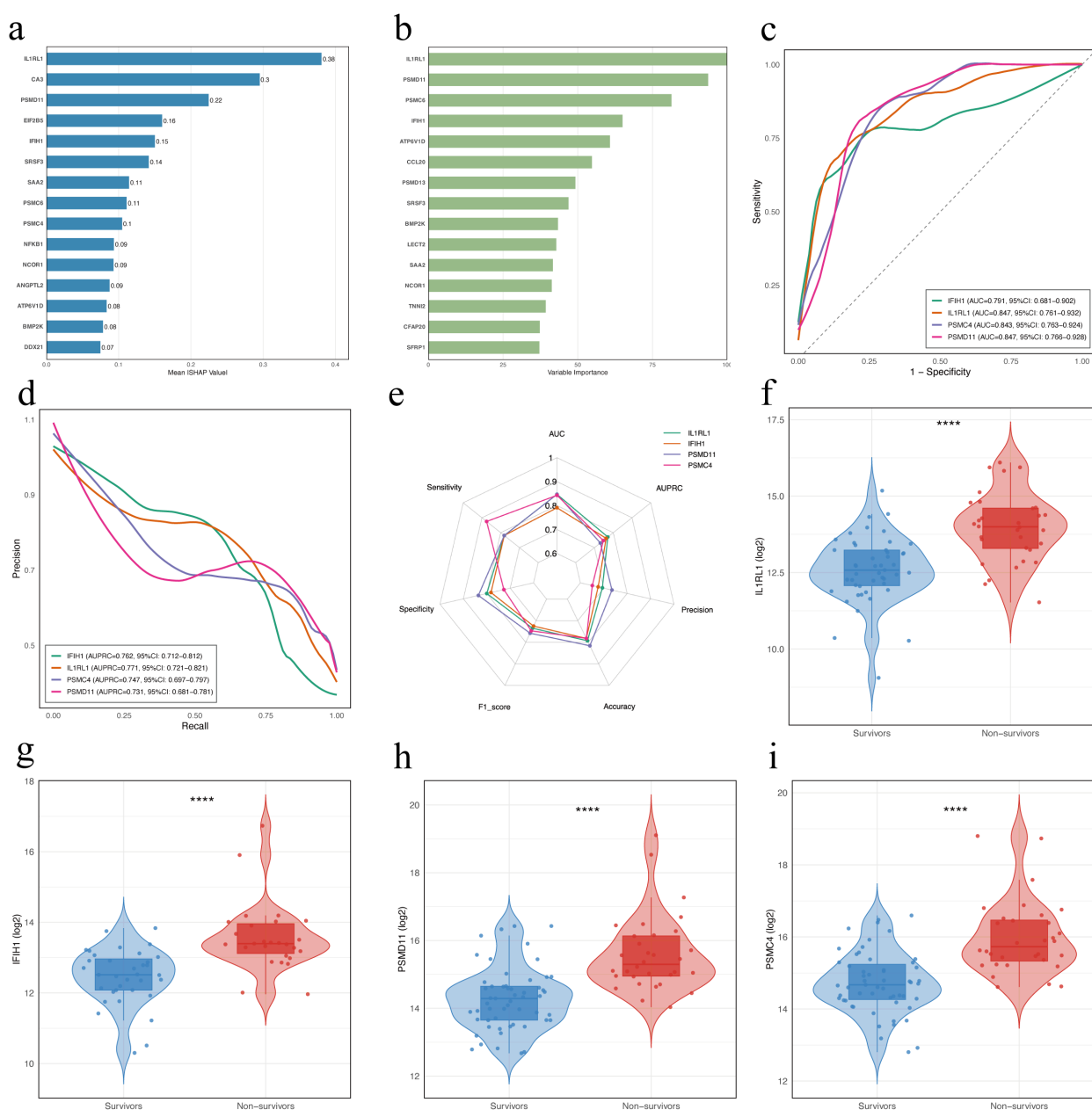


FIGURE 4

Identification and evaluation of potential prognostic biomarkers in patients with SFTS using machine learning. **(a)** Feature importance of DAPs in the XGBoost model. Bars indicate the mean absolute SHAP values averaged over 10-fold cross-validation. **(b)** Variable importance of proteins ranked by a Random Forest classifier trained with 10-fold cross-validation, using DAPs ($|\log_2FC| > 1$, FDR < 0.05) as input features. **(c)** ROC curves for the four proteins. **(d)** Precision-recall (PR) curves for the four proteins. **(e)** Radar chart comparing diagnostic performance metrics of the four proteins. **(f–i)** Violin and box plots showing abundance levels of IL1RL1, IFIH1, PSMC4, and PSMD11 in survivors vs. non-survivors. Statistical significance was assessed using the Wilcoxon rank-sum test; **** indicates $p < 0.0001$. DAPs, differentially abundant proteins.

clinical outcomes in SFTS and highlight PSMD11 and the proteasome pathway as promising candidates for prognostic assessment and potential translational research. In addition, this study extends current strategies for prognostic biomarker discovery in SFTS and highlights a set of novel candidate biomarkers that warrant further investigation. These key DAPs can serve not only as effective predictors of adverse outcomes but also as indicators of severe-related pathological mechanisms.

Previous studies have identified higher viral load, prolonged TT and APTT, as well as elevated levels of D-dimer, AST, ALT, LDH, CK, and cTnI as risk factors for adverse outcomes in patients with SFTS (26–28). Lower platelet count and albumin have also been linked to increased mortality (29). SFTS is characterized by multi-organ involvement and damage (30). While individual clinical laboratory indicators can typically reflect injury to only one or two specific organs, they fail to capture the systemic complexity of

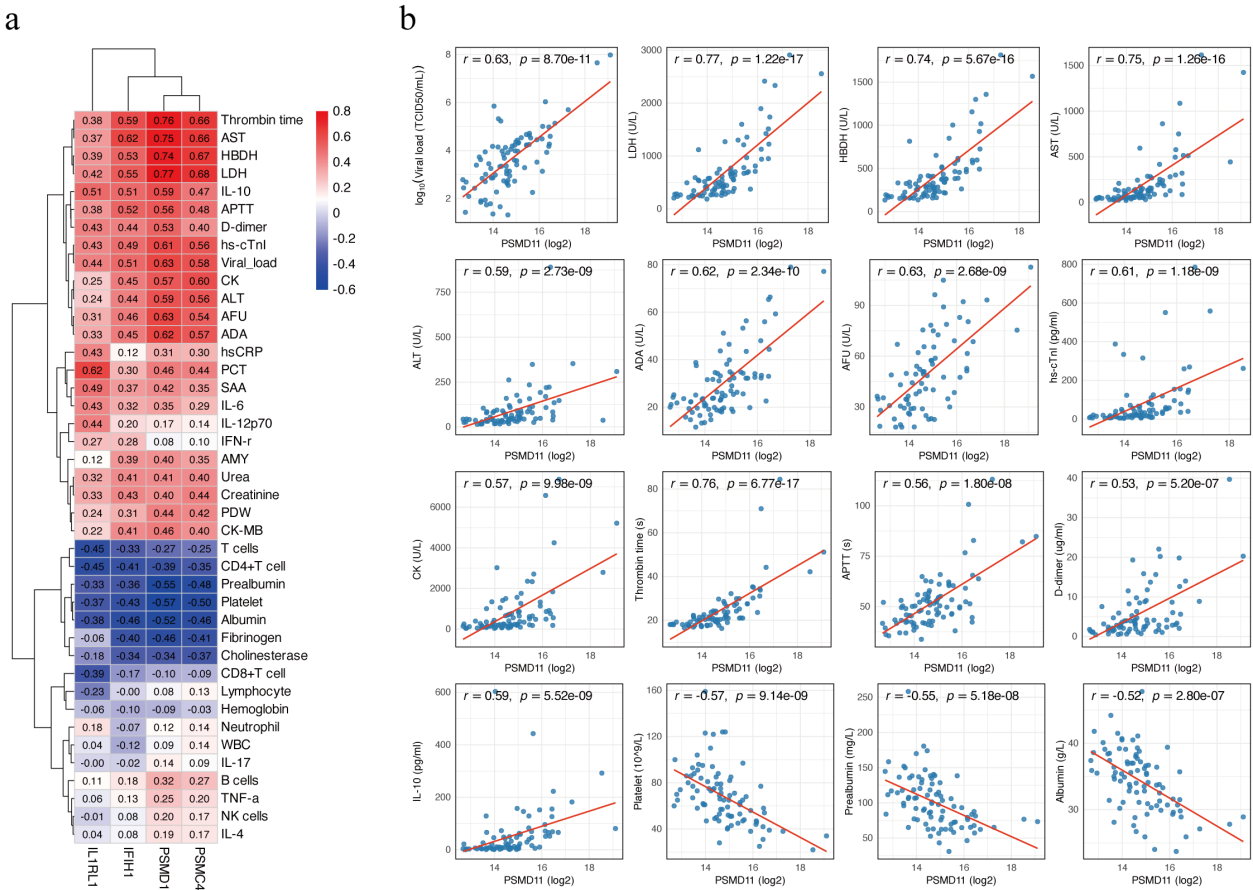


FIGURE 5 Correlation between candidate biomarkers and clinical parameters. **(a)** Heatmap of age-adjusted partial Spearman correlation coefficients between four proteins (PSMD11, IL1RL1, IFIH1, and PSMC4) and clinical laboratory parameters. Positive correlations are shown in red, and negative correlations in blue. **(b)** Scatter plots showing age-adjusted correlations between PSMD11 and clinical parameters. Viral load values were log₁₀-transformed prior to analysis, while other variables were analyzed on their original scales. Partial Spearman correlation coefficients (r) and p -values are shown on each panel, and the red line represents the linear regression fit.

the disease. In our study, age-adjusted partial Spearman correlation analysis revealed that PSMD11 was significantly associated with a wide range of key clinical parameters, including markers of hepatic and myocardial injury, coagulation dysfunction, systemic inflammation, and viral load. These associations indicate that PSMD11 aligns with established clinical indicators of poor prognosis and may serve as a composite marker capturing multiple dimensions of disease severity in patients with SFTS.

PSMD11 is a non-ATPase lid subunit of the 19S regulatory particle that stabilizes 26S proteasome assembly and modulates substrate engagement (31). Ectopic abundance of PSMD11 increases proteasome assembly and activity in cells with relatively low basal proteasome function, whereas PSMD11 knockdown reduces the number of assembled proteasomes (32). PSMD11 serves as a dynamically regulated node that enhances 26S proteasome activity and maintains proteostasis under cellular stress (33). A recent study demonstrated that PSMD11 loss-of-function disrupts 26S proteasome assembly and triggers a persistent type I interferon signature through integrated stress response protein kinase R (34). In glomerular mesangial cells, miR-451 significantly inhibited inflammation and proliferation by downregulating PSMD11 and NF- κ B p65, and transfection of miR-451 mimics significantly decreased levels of IL-1 β , IL-6, and IL-8 (35). Consistently, our study also revealed significantly higher NF- κ B abundance in non-survivors of SFTS. Previous studies have shown that an lncRNA, highly expressed in resting CD4⁺ T lymphocytes, recruits an HIV-1 regulatory protein to a PSMD11-containing ubiquitin-proteasome complex, thereby promoting its degradation and potentially contributing to HIV-1 latency (36). In addition, PSMD11 has been implicated in tumorigenesis partly through the modulation of tumor metabolism-related pathways. In hepatocellular carcinoma, its abundance is elevated and correlates with pathological stage and histologic grade (37). In lung carcinoma A549 cells, PSMD11 overexpression promotes proliferation, migration, invasion, and xenograft growth, while also altering immune-cell infiltration within the tumor microenvironment (38). In pancreatic ductal adenocarcinoma, PSMD11 and its related subunit PSMD14 are overexpressed, with higher abundance levels correlating with increased tumor malignancy and disease progression (39), and another study demonstrated that PSMD11 may serve as potential prognostic and diagnostic biomarkers in patients with early-stage disease (40). A previous study also reported that PSMD11 may play an important role in the metastasis of melanoma (41). Bortezomib is a potent proteasome inhibitor that has been widely used in the treatment of hematologic malignancies, acting through inhibition of the chymotrypsin-like site of the 20S proteolytic core within the 26S proteasome, which in turn induces cell-cycle arrest and apoptosis (42). Rather than being confined to a specific disease context, the dysregulation of PSMD11 appears to reflect broader cellular and immune processes. PSMD11 may be a multifunctional protein that contributes to the multi-organ dysfunction characteristic of severe SFTS, thereby

underscoring its potential relevance to disease pathophysiology and prognosis.

This study has certain limitations. First, the sample size was relatively modest and derived from a single center, which may restrict the generalizability of our findings. Second, the lack of an independent external validation cohort limits the confirmation of robustness and the reproducibility of PSMD11 as a prognostic biomarker. Third, although PSMD11 emerged as a promising candidate, our results are associative and require functional validation to establish causality. Finally, only proteomic data were analyzed without integration of other omics layers or longitudinal sampling. Future multicenter studies with larger, externally validated cohorts and mechanistic experiments are warranted to confirm these findings and further elucidate the biological role of PSMD11 in SFTS.

In conclusion, our proteomic study identified a proteasome-centered axis associated with adverse outcomes in SFTS, with PSMD11 emerging as a prognostic biomarker closely linked to key clinical parameters and multi-organ dysfunction. These findings not only extend current strategies for biomarker discovery but also provide mechanistic insights into disease progression, thereby providing a foundation for future translational research and therapeutic exploration in SFTS.

Data availability statement

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Ethics statement

The studies involving humans were approved by The Ethics Committee of Beijing Ditan Hospital, Capital Medical University (No. DTEC-KY2022-022). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

CZ: Conceptualization, Data curation, Formal analysis, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing. ZG: Data curation, Formal analysis, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing. RW: Data curation, Formal analysis, Investigation, Writing – original draft. YX: Data curation, Investigation, Project administration, Writing – original draft. TZ: Data curation, Formal analysis, Investigation, Writing – original draft. ZJ: Data curation, Investigation, Writing – original draft. LuL:

Data curation, Investigation, Writing – original draft. LiL: Project administration, Supervision, Writing – review & editing. ZC: Conceptualization, Funding acquisition, Project administration, Supervision, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research and/or publication of this article. This work was supported by the National Key Research and Development Program of China (2022YFF1203201), and the National Natural Science Foundation of China (82072295).

Acknowledgments

We sincerely thank all the patients who participated in this study and the doctors and nurses who cared for the patients.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Taxon Details | ICTV . Available online at: https://ictv.global/taxonomy/taxondetails?taxnode_id=202400166&taxon_name=Bandavirus%20dabieense (Accessed August 18, 2025).
2. Yu X-J, Liang M-F, Zhang S-Y, Liu Y, Li J-D, Sun Y-L, et al. Fever with thrombocytopenia associated with a novel bunyavirus in China. *N Engl J Med*. (2011) 364:1523–32. doi: 10.1056/NEJMoa1010095
3. Kim K-H, Yi J, Kim G, Choi SJ, Jun KI, Kim N-H, et al. Severe fever with thrombocytopenia syndrome, South Korea, 2012. *Emerg Infect Dis*. (2013) 19:1892–4. doi: 10.3201/eid1911.130792
4. Takahashi T, Maeda K, Suzuki T, Ishido A, Shigeoka T, Tominaga T, et al. The first identification and retrospective study of severe fever with thrombocytopenia syndrome in Japan. *J Infect Dis*. (2014) 209:816–27. doi: 10.1093/infdis/jit603
5. Tran XC, Yun Y, Van An L, Kim S-H, Thao NTP, Man PKC, et al. Endemic severe fever with thrombocytopenia syndrome, Vietnam. *Emerg Infect Dis*. (2019) 25:1029–31. doi: 10.3201/eid2505.181463
6. Rattanakomol P, Khongwicht S, Linsuwanon P, Lee KH, Vongpunsawad S, Poovorawan Y. Severe fever with thrombocytopenia syndrome virus infection, Thailand, 2019–2020. *Emerg Infect Dis*. (2022) 28:2572–4. doi: 10.3201/eid2812.221183
7. Zohaib A, Zhang J, Saqib M, Athar MA, Hussain MH, Chen J, et al. Serologic evidence of severe fever with thrombocytopenia syndrome virus and related viruses in Pakistan. *Emerg Infect Dis*. (2020) 26:1513–6. doi: 10.3201/eid2607.190611
8. Li H, Lu Q-B, Xing B, Zhang S-F, Liu K, Du J, et al. Epidemiological and clinical features of laboratory-diagnosed severe fever with thrombocytopenia syndrome in China, 2011–17: a prospective observational study. *Lancet Infect Dis*. (2018) 18:1127–37. doi: 10.1016/S1473-3099(18)30293-7
9. Cui H, Shen S, Chen L, Fan Z, Wen Q, Xing Y, et al. Global epidemiology of severe fever with thrombocytopenia syndrome virus in human and animals: a systematic review and meta-analysis. *Lancet Reg Health West Pac*. (2024) 48:101133. doi: 10.1016/j.lanwpc.2024.101133
10. Qiao L, Zhang T-J, Feng Y, Yang L, Cai P, Liu S-W, et al. Case Report: Successful treatment of severe fever with thrombocytopenia syndrome associated with hemophagocytic lymphohistiocytosis by preemptively using favipiravir and methylprednisolone. *Front Med*. (2025) 12:1566719. doi: 10.3389/fmed.2025.1566719

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fimmu.2025.1693946/full#supplementary-material>

11. Kubo K, Toyama M, Kobayashi K, Komiya N, Nakayama T. Re-evaluating a single-arm trial of favipiravir for severe fever with thrombocytopenia syndrome in Japan: A proposal to use real-world data as external controls. *PloS Negl Trop Dis*. (2025) 19:e0013279. doi: 10.1371/journal.pntd.0013279
12. Pathogens prioritization: a scientific framework for epidemic and pandemic research preparedness . Available online at: <https://www.who.int/publications/m/item/pathogens-prioritization-a-scientific-framework-for-epidemic-and-pandemic-research-preparedness> (Accessed August 18, 2025).
13. Wang G, Ge H-H, Hu L, Guo P-J, Cui N, Zhu C-L, et al. Severe fever with thrombocytopenia syndrome complicated by haemophagocytic lymphohistiocytosis: a retrospective cohort study. *Clin Microbiol Infect*. (2024) 30:558–60. doi: 10.1016/j.cmi.2024.01.009
14. Song L, Zou W, Wang G, Qiu L, Sai L. Cytokines and lymphocyte subsets are associated with disease severity of severe fever with thrombocytopenia syndrome. *Virology*. (2024) 21:126. doi: 10.1186/s12985-024-02403-0
15. Zhang Y, Li L, Liu Y, Zhang W, Peng W, Zhang S, et al. Identification of CCL20 as a prognostic predictor for severe fever with thrombocytopenia syndrome based on plasma proteomics. *J Infect Dis*. (2024) 230(3):741–753. doi: 10.1093/infdis/jiae039
16. Loeb K, Lemaille C, Frederick C, Wallace HL, Kindrachuk J. Harnessing high-throughput OMICS in emerging zoonotic virus preparedness and response activities. *Biochim Biophys Acta Mol Basis Dis*. (2024) 1870:167337. doi: 10.1016/j.bbadis.2024.167337
17. Chen C, Wang J, Pan D, Wang X, Xu Y, Yan J, et al. Applications of multi-omics analysis in human diseases. *MedComm*. (2023) 4:e315. doi: 10.1002/mco2.315
18. Nie X, Qian L, Sun R, Huang B, Dong X, Xiao Q, et al. Multi-organ proteomic landscape of COVID-19 autopsies. *Cell*. (2021) 184:775–791.e14. doi: 10.1016/j.cell.2021.01.004
19. Shen B, Yi X, Sun Y, Bi X, Du J, Zhang C, et al. Proteomic and metabolomic characterization of COVID-19 patient sera. *Cell*. (2020) 182:59–72.e15. doi: 10.1016/j.cell.2020.05.032
20. Cen X, Wang F, Huang X, Jovic D, Dubee F, Yang H, et al. Towards precision medicine: Omics approach for COVID-19. *Biosaf Health*. (2023) 5:78–88. doi: 10.1016/j.bshealth.2023.01.002

21. Reel PS, Reel S, Pearson E, Trucco E, Jefferson E. Using machine learning approaches for multi-omics data analysis: A review. *Biotechnol Adv.* (2021) 49:107739. doi: 10.1016/j.biotechadv.2021.107739
22. Su C-Y, Zhou S, Gonzalez-Kozlova E, Butler-Laporte G, Brunet-Ratnasingham E, Nakanishi T, et al. Circulating proteins to predict COVID-19 severity. *Sci Rep.* (2023) 13:6236. doi: 10.1038/s41598-023-31850-y
23. Boada P, Fatou B, Belperron AA, Sigdel TK, Smolen KK, Wurie Z, et al. Longitudinal serum proteomics analyses identify unique and overlapping host response pathways in Lyme disease and West Nile virus infection. *Front Immunol.* (2022) 13:1012824. doi: 10.3389/fimmu.2022.1012824
24. Liu S, Chai C, Wang C, Amer S, Lv H, He H, et al. Systematic review of severe fever with thrombocytopenia syndrome: virology, epidemiology, and clinical characteristics. *Rev Med Virol.* (2014) 24:90–102. doi: 10.1002/rmv.1776
25. Li R, Li L, Xu Y, Yang J. Machine learning meets omics: applications and perspectives. *Brief Bioinform.* (2022) 23:bbab460. doi: 10.1093/bib/bbab460
26. Zhang Z, Hu X, Du Q, Mo P, Chen X, Luo M, et al. Clinical characteristics and outcomes of disseminated intravascular coagulation in patients with severe fever with thrombocytopenia syndrome. *BMC Infect Dis.* (2025) 25:508. doi: 10.1186/s12879-025-10900-y
27. Wang Y, Song Z, Wei X, Yuan H, Xu X, Liang H, et al. Clinical laboratory parameters and fatality of Severe fever with thrombocytopenia syndrome patients: A systematic review and meta-analysis. *PloS Negl Trop Dis.* (2022) 16:e0010489. doi: 10.1371/journal.pntd.0010489
28. Liang B, Xu L, Li M, Wang H, Lu S, Fan L, et al. The association between elevated myocardial injury-related biomarker (TnI) and increased mortality in patients with severe fever with thrombocytopenia syndrome*. *Crit Care Med.* (2024) 52:1509. doi: 10.1097/CCM.00000000000006367
29. Yan S, Ding X, Gao Q, Zhao L, Li C, Sun Z, et al. Prediction of mortality by clinical laboratory parameters in severe fever with thrombocytopenia syndrome: A meta-analysis. *Trop Med Infect Dis.* (2025) 10:193. doi: 10.3390/tropicalmed10070193
30. Peng C, Hao Y, Yuan Y, Ma W, Zhang D, Kong J, et al. Bandavirus dabiense: A review of epidemiology, clinical characteristics, pathophysiology, treatment and prevention. *Virulence.* (2025) 16:2520343. doi: 10.1080/21505594.2025.2520343
31. Lokireddy S, Kukushkin NV, Goldberg AL. cAMP-induced phosphorylation of 26S proteasomes on Rpn6/PSMD11 enhances their activity and the degradation of misfolded proteins. *Proc Natl Acad Sci U.S.A.* (2015) 112:E7176–85. doi: 10.1073/pnas.1522332112
32. Vilchez D, Boyer L, Morante I, Lutz M, Merkwirth C, Joyce D, et al. Increased proteasome activity determines human embryonic stem cell identity. *Nature.* (2012) 489:304–8. doi: 10.1038/nature11468
33. Ahammed MS, Wang X. Promoting proteostasis by cAMP/PKA and cGMP/PKG. *Trends Mol Med.* (2025) 31:224–39. doi: 10.1016/j.molmed.2024.10.006
34. Deb W, Rosenfelt C, Vignard V, Papendorf JJ, Möller S, Wendlandt M, et al. PSMD11 loss-of-function variants correlate with a neurobehavioral phenotype, obesity, and increased interferon response. *Am J Hum Genet.* (2024) 111:1352–69. doi: 10.1016/j.ajhg.2024.05.016
35. Wei H, Li J, Li Y, Song J. MicroRNA-451 inhibits inflammation and proliferation of glomerular mesangial cells through down-regulating PSMD11 and NF- κ B p65. *Biosci Rep.* (2019) 39:BSR20191455. doi: 10.1042/BSR20191455
36. Li J, Chen C, Ma X, Geng G, Liu B, Zhang Y, et al. Long noncoding RNA NRON contributes to HIV-1 latency by specifically inducing tat protein degradation. *Nat Commun.* (2016) 7:11730. doi: 10.1038/ncomms11730
37. Zhang C, Xu T, Ji K, Cao S, Cao Y, Ai J, et al. An integrative analysis reveals the prognostic value and potential functions of PSMD11 in hepatocellular carcinoma. *Mol Carcinog.* (2023) 62(9):1355–1368. doi: 10.1002/mc.23568
38. Huang Q, Tian R, Yu J, Du W. Identification of PSMD11 as a novel cuproptosis- and immune-related prognostic biomarker promoting lung adenocarcinoma progression. *Cancer Med.* (2024) 13:e7379. doi: 10.1002/cam4.7379
39. Yang Y-H, Xing Z-H, Wang H, Zhang C, Liu Y-B, Bai Q-Q, et al. PSMD11 and PSMD14 may serve as novel biomarkers for the prognosis of pancreatic ductal adenocarcinoma. *Front Oncol.* (2025) 15:1555649. doi: 10.3389/fonc.2025.1555649
40. Zhou C, Li H, Han X, Pang H, Wu M, Tang Y, et al. Prognostic value and molecular mechanisms of proteasome 26S subunit, non-ATPase family genes for pancreatic ductal adenocarcinoma patients after pancreaticoduodenectomy. *J Invest Surg Off J Acad Surg Res.* (2022) 35:330–46. doi: 10.1080/08941939.2020.1863527
41. Chen J, Wu F, Shi Y, Yang D, Xu M, Lai Y, et al. Identification of key candidate genes involved in melanoma metastasis. *Mol Med Rep.* (2019) 20:903–14. doi: 10.3892/mmr.2019.10314
42. Sogbein O, Paul P, Umar M, Chaari A, Batuman V, Upadhyay R. Bortezomib in cancer therapy: Mechanisms, side effects, and future proteasome inhibitors. *Life Sci.* (2024) 358:123125. doi: 10.1016/j.lfs.2024.123125

Glossary

SFTS	severe fever with thrombocytopenia syndrome	PCA	principal component analysis
GO	Gene Ontology	XGBoost	eXtreme Gradient Boosting
KEGG	Kyoto Encyclopedia of Genes and Genomes	SHAP	SHapley Additive exPlanations
PPI	protein–protein interaction	PR	precision-recall
ROC	receiver operating characteristic	AUC	area under the receiver operating characteristic curve
DAPs	differentially abundant proteins	AUPRC	area under the precision-recall curve
SFTSV	severe fever with thrombocytopenia syndrome virus	CI	confidence intervals
PHEIC	Public Health Emergency of International Concern	WBC	white blood cell
IL	interleukin	ALT	alanine aminotransferase
IFN	interferon	AST	aspartate aminotransferase
LDH	lactate dehydrogenase	MPV	mean platelet volume
CCL20	C-C motif chemokine 20	TBil	total bilirubin
ML	machine learning	CIZ1	CDKN1A interacting zinc finger protein 1
HC	healthy control	PSMC4	proteasome 26S subunit, ATPase 4
DIA	data-independent acquisition	PSMD11	proteasome 26S subunit, non-ATPase 11
RT-qPCR	reverse transcription quantitative polymerase chain reaction	NCOR1	nuclear receptor corepressor 1
LC-MS	liquid chromatography–mass spectrometry	ALS	amyotrophic lateral sclerosis
KNN	k-nearest neighbors	MCL	Markov Clustering
IQR	interquartile range	TT	thrombin time
FDR	false discovery rate FC, fold change	HBDH	hydroxybutyrate dehydrogenase
BP	biological process	ADA	adenosine deaminase
CC	cellular component	AFU	alpha-L-fucosidase
MF	molecular function	hs-cTnI	high-sensitivity cardiac troponin I
STRING	search tool for retrieval of interacting genes/proteins	CK	creatine kinase
SD	standard deviation	APTT	activated partial thromboplastin time.