

RESEARCH

Open Access



Identification of noval diagnostic biomarker for HFpEF based on proteomics and machine learning

Muyashaer Abudurexiti¹, Salamaiti Aimaier¹, Nuerdun Wupuer¹, Dongqin Duan¹, Aihaidan Abudouwayiti¹, Meiheriayi Nuermaiti¹ and Ailiman Mahemuti^{1*}

Abstract

Background Heart failure with preserved ejection fraction (HFpEF) is a complex syndrome that currently lacks effective biomarkers for early diagnosis and treatment. This study seeks to identify new potential biomarkers for HFpEF using proteomics and machine learning.

Methods Plasma samples were collected from 20 patients newly diagnosed age, sex, BMI matched HFpEF and 20 healthy controls (HCs). Proteomic analysis was performed using liquid chromatography-tandem mass spectrometry (LC-MS/MS) in data-independent acquisition mode. Differentially expressed proteins (DEPs) were identified and analyzed through enrichment analyses and protein-protein interaction (PPI) network construction. Machine learning methods, including LASSO regression and the Boruta algorithm were used to select candidate biomarkers. The diagnostic value of these proteins was assessed using receiver operating characteristic (ROC) curves and nomogram construction. Expression of candidate proteins was analyzed in immune cells and tissues. Finally, enzyme-linked immunosorbent assay (ELISA) was used to validate the plasma levels of selected proteins.

Results A total of 34 DEPs were identified between HFpEF patients and HCs. Enrichment analyses revealed involvement in acute-phase response and immune pathways. PPI network analysis identified nine hub proteins. Machine learning methods narrowed the candidates to four potential biomarkers: SERPINA1, AFM, SERPINA3, and ITIH4. Among these, SERPINA3 showed the highest diagnostic value with an area under the ROC curve (AUC) of 0.835. ELISA validation confirmed that plasma SERPINA3 levels were significantly elevated in HFpEF patients compared to HCs ($p < 0.0001$).

Conclusions Our findings suggest that SERPINA3 could serve as a biomarker for HFpEF. Elevated plasma levels of SERPINA3 in HFpEF patients suggest its utility in early diagnosis and may provide insights into the disease's pathogenesis.

Keywords HFpEF, SERPINA3, Proteomics, Machine learning, Biomarker

*Correspondence:

Ailiman Mahemuti
xinjiangailiman@163.com

¹Department of Heart Failure, First Affiliated Hospital of Xinjiang Medical University, Urumqi, Xinjiang 830054, China



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. 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-nc-nd/4.0/>.

Introduction

Heart failure with preserved ejection fraction (HFpEF) affects approximately half of all patients with heart failure (HF), leading to substantial morbidity, mortality, and impaired quality of life [1, 2]. Alarming, its prevalence is increasing, and it carries an estimated 5-year mortality rate of up to 75% [3]. Therefore, HFpEF is widely acknowledged as one of the most pressing challenges in cardiovascular medicine, requiring urgent attention from the scientific community [4].

Diagnosing HFpEF is challenging because its symptoms often overlap with those of other comorbid conditions such as hypertension, obesity, and diabetes mellitus [5]. Current diagnostic criteria rely on a combination of clinical presentation, echocardiographic findings, and elevated levels of natriuretic peptides, which may not always be sufficiently specific or sensitive [6]. Therefore, there is a critical need to identify novel biomarkers that can improve the accuracy of HFpEF diagnosis and provide insights into its underlying molecular mechanisms.

Proteomics provides a powerful method for comprehensively analyzing protein expression profiles in biological samples, enabling the discovery of potential biomarkers and therapeutic targets [7]. Recent advancements in proteomic technologies—especially data-independent acquisition (DIA)-based liquid chromatography-tandem mass spectrometry (LC-MS/MS)—have improved the sensitivity and reproducibility of protein detection [8, 9]. Moreover, machine learning algorithms can efficiently process complex proteomic data, helping to identify key proteins associated with disease states [10].

SERPINA3, also known as alpha-1-antichymotrypsin, is a serine protease inhibitor involved in inflammatory processes and has been implicated in various cardiovascular diseases [11, 12]. However, its role in HFpEF remains unclear. In this study, we aim to identify proteins that are differentially expressed in the plasma of HFpEF patients compared to healthy controls (HC) using proteomic analysis and machine learning techniques. Our focus is on evaluating the potential of SERPINA3 as a biomarker for HFpEF. By validating our findings with enzyme-linked immunosorbent assay (ELISA) and conducting *in vitro* experiments, we hope to establish a foundation for improved diagnostic methods and a better understanding of HFpEF pathogenesis.

Materials and methods

Data collection and analyzation

All subjects provided informed consent, and ethical approval for research involving human subjects was obtained from the Ethics Committee of the First Affiliated Hospital of Xinjiang Medical University (IRB-K202403-12). All research procedures complied with the

Declaration of Helsinki. The participants were individuals aged between 18 and 85 who were newly diagnosed with HFpEF upon admission, prior to the initiation of any treatment. All participants had been diagnosed with HFpEF according to established consensus criteria [5, 13–15]. The inclusion criteria required participants to exhibit symptoms and signs of exertional dyspnea corresponding to New York Heart Association (NYHA) class II or III, and to have heart failure with a left ventricular ejection fraction (LVEF) of $\geq 50\%$. Additionally, participants needed to meet at least two of the following conditions:

1. Elevated NT-proBNP (N-terminal pro-B-type natriuretic peptide) levels ≥ 125 pg/mL.
2. Echocardiographic evidence of structural heart abnormalities or diastolic dysfunction.
3. An E/e' ratio ≥ 9 .

The exclusion criteria ruled out individuals with any of the following conditions: congenital heart defects, LVEF $< 40\%$, heart failure categorized as mid-range ejection fraction (EF) ($40\text{--}50\%$), hypertrophic cardiomyopathy, prior cardiac transplantation, constrictive pericarditis, severe valvular disorders, or infiltrative or restrictive cardiomyopathies. The control group consisted of individuals who displayed no evident signs of heart failure, characterized by an LVEF $\geq 50\%$ and NT-proBNP concentrations < 125 pg/mL.

At the time of admission, plasma samples were obtained from patients newly diagnosed with HFpEF. To reduce the impact of dietary variables, patients fasted overnight prior to sample collection. Blood samples were drawn into vacuum-sealed tubes containing EDTA to prevent coagulation. After centrifugation at 3000 rpm for 10 min, the supernatant was carefully collected and stored at -80°C for subsequent analysis [4].

Proteome analysis

To investigate plasma proteins, we conducted LC-MS/MS analysis using data-independent acquisition (DIA)-based proteomics. The samples were analyzed on a Q Exactive HF-X mass spectrometer (Thermo Fisher, Germany) coupled with an EASY-nLC 1200 UHPLC system (Thermo Fisher, Germany). Data collection was performed in DIA mode to ensure comprehensive protein profiling [8, 9].

Differentially expressed proteins (DEPs) screening

Differentially expressed proteins (DEPs) of significant interest were identified using a fold-change threshold of ≥ 1.5 or ≤ 0.67 and a *t*-test *p*-value of < 0.05 [4, 16]. The analysis was conducted using the Limma package in R.

Pathway and biological function enrichment analysis

To identify the signaling pathways associated with the DEPs in HFpEF, KEGG pathway analysis was conducted. Additionally, Gene Ontology (GO) functional enrichment analysis was performed to explore the biological processes (BP), cellular components (CC), and molecular functions (MF) of these DEPs [17]. This approach aimed to investigate the activated signaling pathways and biological functions in HFpEF.

PPI (protein–protein interactions) network analysis

To construct a robust and reliable protein–protein interaction (PPI) network, we utilized the STRING database (<https://cn.string-db.org/>). Cytoscape software was employed to identify proteins with significant interactions within the network [18].

Feature selection of characteristic biomarkers via two machine learning methods

Feature selection was initially performed using the LASSO algorithm implemented in the ‘glmnet’ package, with parameter tuning conducted via k-fold cross-validation. The LASSO algorithm, which incorporates a penalty parameter (λ), is effective at evaluating high-dimensional data [19]. We selected λ_{\min} to construct the model with the best fit while retaining more variables [20]. Subsequently, the Boruta algorithm was applied using the ‘Boruta’ package to confirm the relevance of the features and identify the final hub proteins [21]. The intersection of the two results served as the candidate hub proteins for diagnosis.

Nomogram construction and receiver operating characteristic evaluation

To assess the significance of the candidate proteins in diagnosing HFpEF, we developed a nomogram using the ‘rms’ R package. This nomogram includes a “Points” scale, representing the score assigned to each candidate protein, and a “Total Points” scale, displaying the cumulative score of all protein [22]. The nomogram proved to be a crucial tool for predicting the diagnosis of HFpEF. To further evaluate the prognostic value of the candidate proteins and the nomogram, we conducted a ROC analysis, which provided the area under the curve (AUC) along with the 95% confidence interval (CI). An AUC value greater than 0.7 was considered indicative of high diagnostic efficacy [23].

Gene–gene interactions (GGI) and friends analysis

To investigate the functional roles of key genes and their interacting proteins, we used GeneMANIA (www.genemania.org) to construct a gene–gene interaction (GGI) network [24]. Semantic comparison of Gene Ontology (GO) annotations provides a quantitative method

for analyzing similarities between genes and genomes, serving as a crucial foundation for many bioinformatics analyses [25]. To further compare the similarities among differentially expressed genes (DEGs), we used the GOSemSim R package. This allowed us to calculate GO semantic similarity scores for DEGs by computing the geometric mean of their similarity scores across biological processes, molecular functions, and cellular components. Finally, we visualized the results using the ggplot2 R package [26].

Expression of feature genes in immune cells and tissue-specific analysis

To determine the expression levels of key proteins within immune cells, we analyzed their expression using the Human Protein Atlas (HPA) database (<https://www.proteinatlas.org/>). To further explore the expression levels of these genes across different tissues, we accessed data from the GTEx database to further explore how these proteins are expressed across different tissues, we accessed data from the GTEx database (<https://gtexportal.org/home/>).

Enzyme-linked immunosorbent assay (ELISA) of clinical blood samples

plasma samples were obtained from patients newly diagnosed with HFpEF. To reduce the impact of dietary variables, patients fasted overnight prior to sample collection. Blood samples were drawn into vacuum-sealed tubes containing EDTA to prevent coagulation. After centrifugation at 3000 rpm for 10 min, the supernatant was carefully collected and stored at -80°C for subsequent analysis. Plasma SERPINA3 levels were measured by ELISA kit SERPINA3(Human SERPINA3 ELISA Kit YX-E11217, sinobestbio), ITIH4 (Human ITIH4 ELISA Kit EK1670 BOSTER), SERPINA1 (Human Alpha 1 Antitrypsin/SERPINA1 ELISA Kit EK1634 BOSTER), AFM (Human Afamin/AFM ELISA Kit (EK1487 BOSTER) according to the manufacturer’s instructions.

Results

Baseline characteristics of patient study

Supplementary Table S1 showed the demographic and laboratory characteristics of patients in HFpEF and HC groups. There were no significant differences in terms of age, gender, BMI, and prevalence of hyperlipidemia ($p>0.05$). However, the prevalence of comorbid conditions such as coronary artery disease (CAD) ($p=0.002$) and diabetes mellitus (DM) ($p<0.01$) was significantly higher in the HFpEF group compared to the HCs. In contrast, parameters indicative of diastolic dysfunction were significantly elevated in the HFpEF group. The pulmonary artery (PA) pressure was higher in HFpEF patients compared to HCs ($p=0.013$). Additionally, the E/e’ ratio,

an echocardiographic marker of left ventricular filling pressures, was significantly higher in the HFpEF group than in the HC group ($p=0.002$). Comparative analysis of laboratory data revealed no significant differences in WBC, CRP, CKMB and CK ($p>0.05$) between the two groups. In contrast, Interleukin-6 (IL-6), Serum amyloid A (SAA), N-terminal pro-brain natriuretic peptide (NT-proBNP), creatinine (Cr), and Lipoprotein(a) [Lp (a)] levels were significantly higher in HFpEF patients compared to HCs ($p<0.05$).

Differential expression analysis

We used the limma package to analyze the differentially expressed proteins (DEPs) between normal samples and HFpEF samples. Our analysis revealed that 26 proteins were upregulated and 8 were downregulated, as illustrated in the volcano plots in Fig. 1A. To gain a comprehensive understanding of the expression patterns of DEPs, we employed hierarchical clustering. This method allowed us to visualize and interpret the relationships and similarities among DEPs based on their expression profiles. The hierarchical clustering heatmap enabled us to identify distinct clusters of protein expression, providing insights into the fundamental processes underlying the dysregulated protein expression typically observed in HFpEF, as shown in Fig. 1B.

Enrichment analysis of depts

We conducted gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses on the DEPs to better understand the signaling pathways and biological functions they are involved in. The GO enrichment results, displayed in Fig. 1C, identified 350 enriched GO functional entries, including 34 cellular components (CC), 44 molecular functions (MF), and 272 biological processes (BP). These pathways are primarily associated with the acute-phase response, humoral immune reaction, acute inflammatory response, complement activation, and the killing of cells from other organisms. The KEGG enrichment results, shown in Fig. 1D, revealed enrichment in 12 pathways. These pathways are mainly involved in the complement and coagulation cascade, regulation of the actin cytoskeleton, coronavirus disease (COVID-19), and systemic lupus erythematosus, among others.

PPI network analysis and identification of hub proteins

In this study, protein interaction networks of 34 differential proteins were constructed using the STRING database (confidence value >0.4). The results are shown in Fig. 2A. The PPI network has 31 nodes (Nodes) and 76 edges (Edges). In this study, MCODE plugin was selected for top connectivity gene clusters, and the TOP1 gene clusters were selected for follow-up study. The results are

shown in Fig. 2B, with 9 nodes and 35 edges in this network. Nine hub proteins were obtained: AFM, C9, ORM1, SERPINA1, A2M, F2, SERPINA3, ITIH4, and ITIH3.

Feature selection of characteristic biomarkers via two machine learning methods

We created two algorithms to choose potential genes for HFpEF from a total of 9 PPI-related DEPs. LASSO regression was then applied for candidate gene identification to distinguish HFpEF from HC. From the results, 6 potential candidate proteins were identified. These proteins are SERPINA1, AFM, SERPINA3, ITIH4, A2M, F2 (Fig. 3A, B). Subsequently, we constructed a Boruta algorithm to identify the final 5 hub proteins. These are SERPINA1, AFM, SERPINA3, ITIH4, ORM1 (Fig. 3C). We chose the proteins that were produced by each of the two distinct machine learning models, and then we intersected them. Because of this, we were only able to focus our research on 4 proteins, SERPINA1, AFM, SERPINA3, ITIH4 (Fig. 3D).

Diagnostic value of feature proteins

We calculated the ROC curves for the 4 feature proteins and the column line plot model.

to evaluate their diagnostic performance. Our primary focus was on the AUC (Area Under the Curve) value, according to the established criteria, an area under the ROC curve (AUC) of $0.9 \leq \text{AUC} < 1$ indicates excellent accuracy, $0.8 \leq \text{AUC} < 0.9$ indicates good accuracy, while $\text{AUC} = 0.5$ indicates no informative accuracy, SERPINA1 (AUC:0.815), SERPINA3 (0.835) showed higher AUC values (Fig. 4A). We further analyzed the expression patterns of the feature proteins in two groups. The results showed that, SERPINA1, SERPINA3, and ITIH4 exhibited significantly higher expression levels in HFpEF groups compared to HC groups (Fig. 4B).

Subsequently, A nomogram was constructed based on the prediction model (Fig. 4C). The prediction accuracy of the calibration curve and ROC curve model is evaluated, and the results are shown in Fig. 4D, E. The calibration curve indicates that the diagnostic error rate is low, and there is no obvious difference between the predicted value and the true value through HL test, and the C-index of the model reaches 0.7, which shows that the model has a good differentiation between HC and HFpEF group.

Gene-gene interactions (GGI) network analyses for feature proteins

Gene-gene interaction network (GGI) of 4 feature proteins was constructed using GeneMANIA (<http://genemania.org>), TOP 20 proteins were selected as key nodes for display, and the function of correlation TOP 7 was shown in Fig. 5A. The results revealed that 4 feature proteins are related to secretory granule lumen, peptidase inhibitor

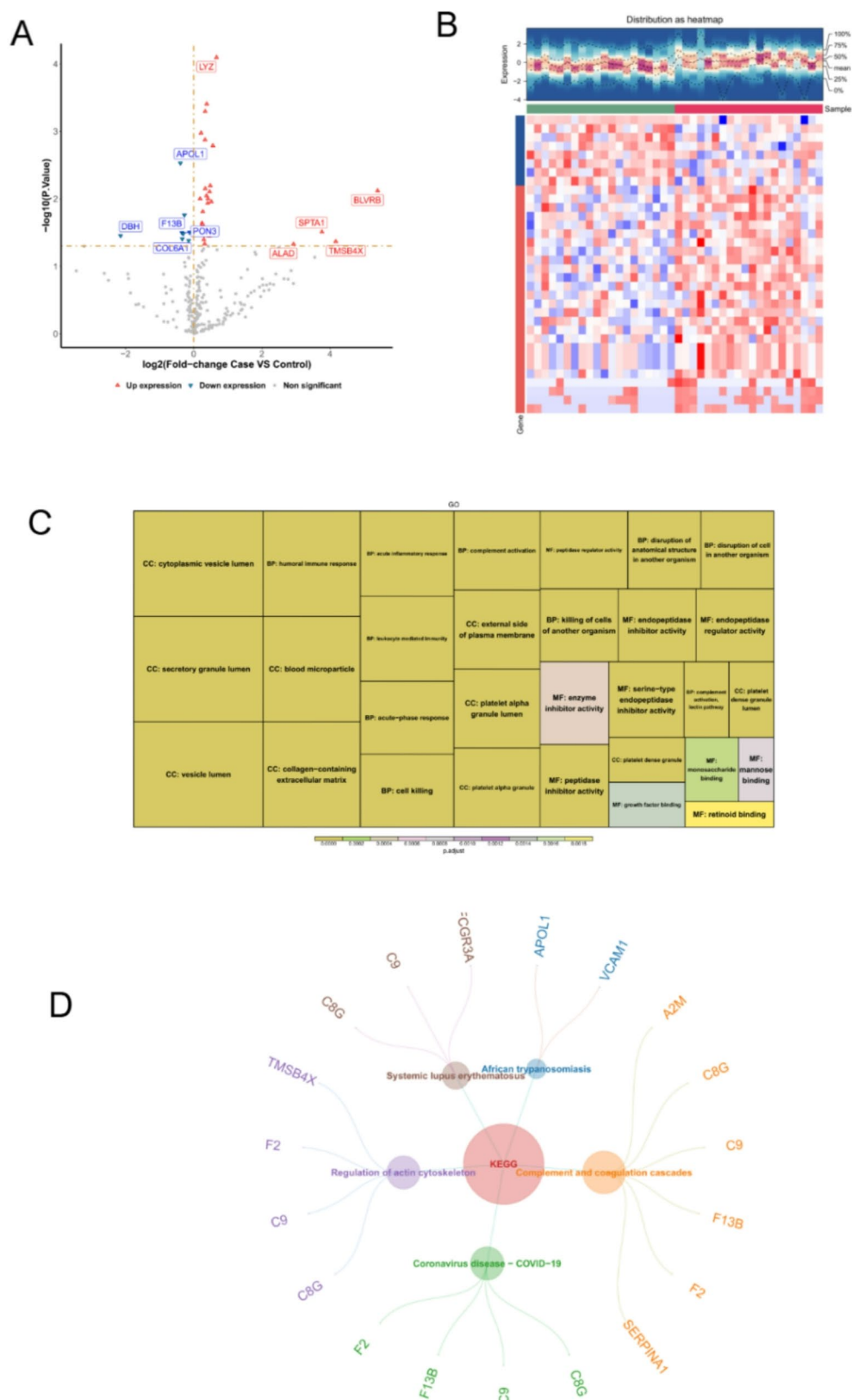


Fig. 1 Proteomics analysis in patients with HFpEF compared to HC. **(A)** Volcano plot of the DEPs. **(B)** Heatmap displaying the DEPs. **(C-D)** Functional enrichment analysis based on DEPs. **(C)** GO pathway analysis (biological process, cellular component, and molecular function) of top10 pathways. **(D)** KEGG pathway analysis of top5 pathways

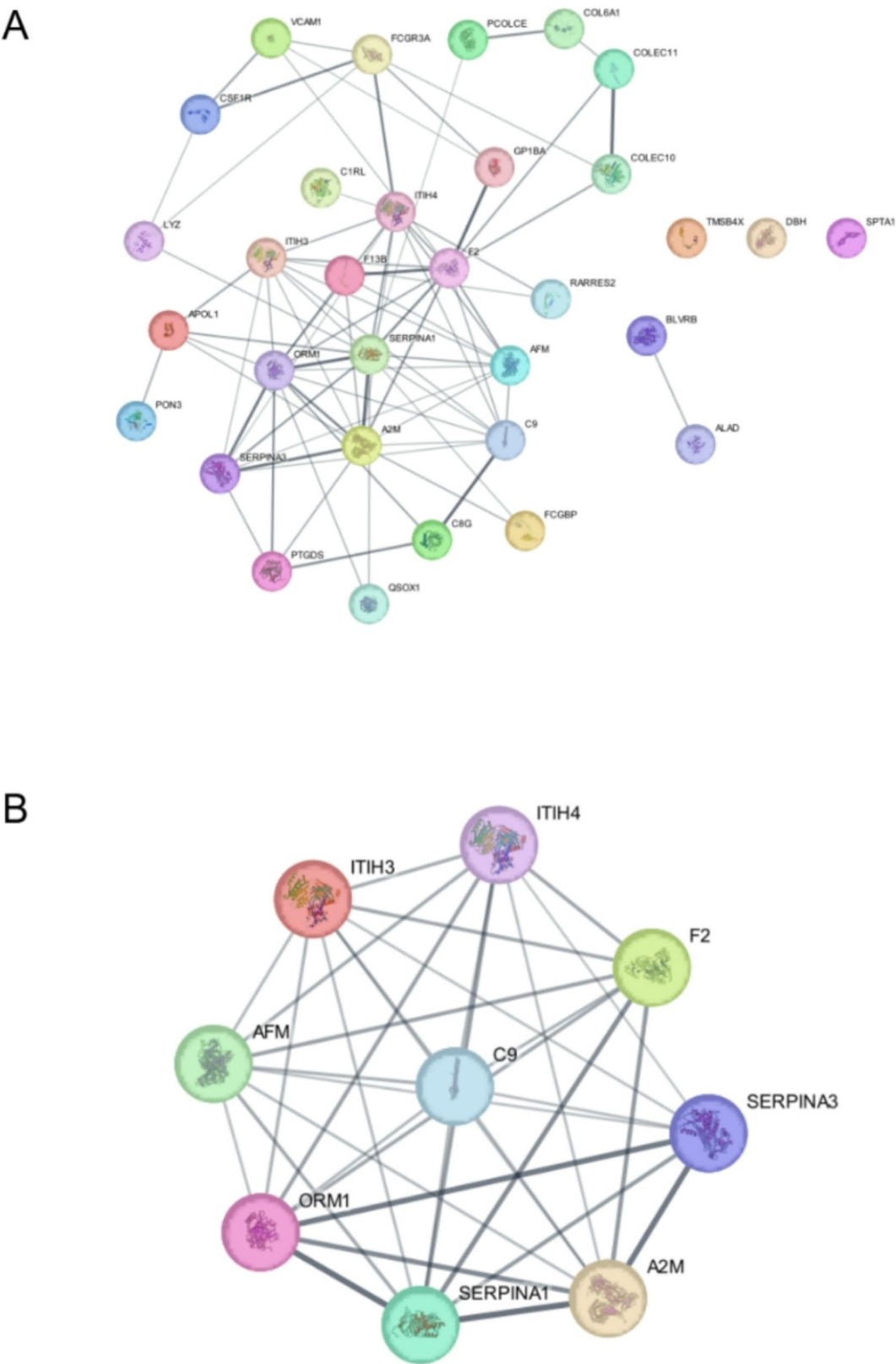


Fig. 2 The results of an analysis of the PPI network and identifies hub proteins. **(A)** PPI network associated with DEPs. **(B)** The proteins that serve as nodes in a PPI network

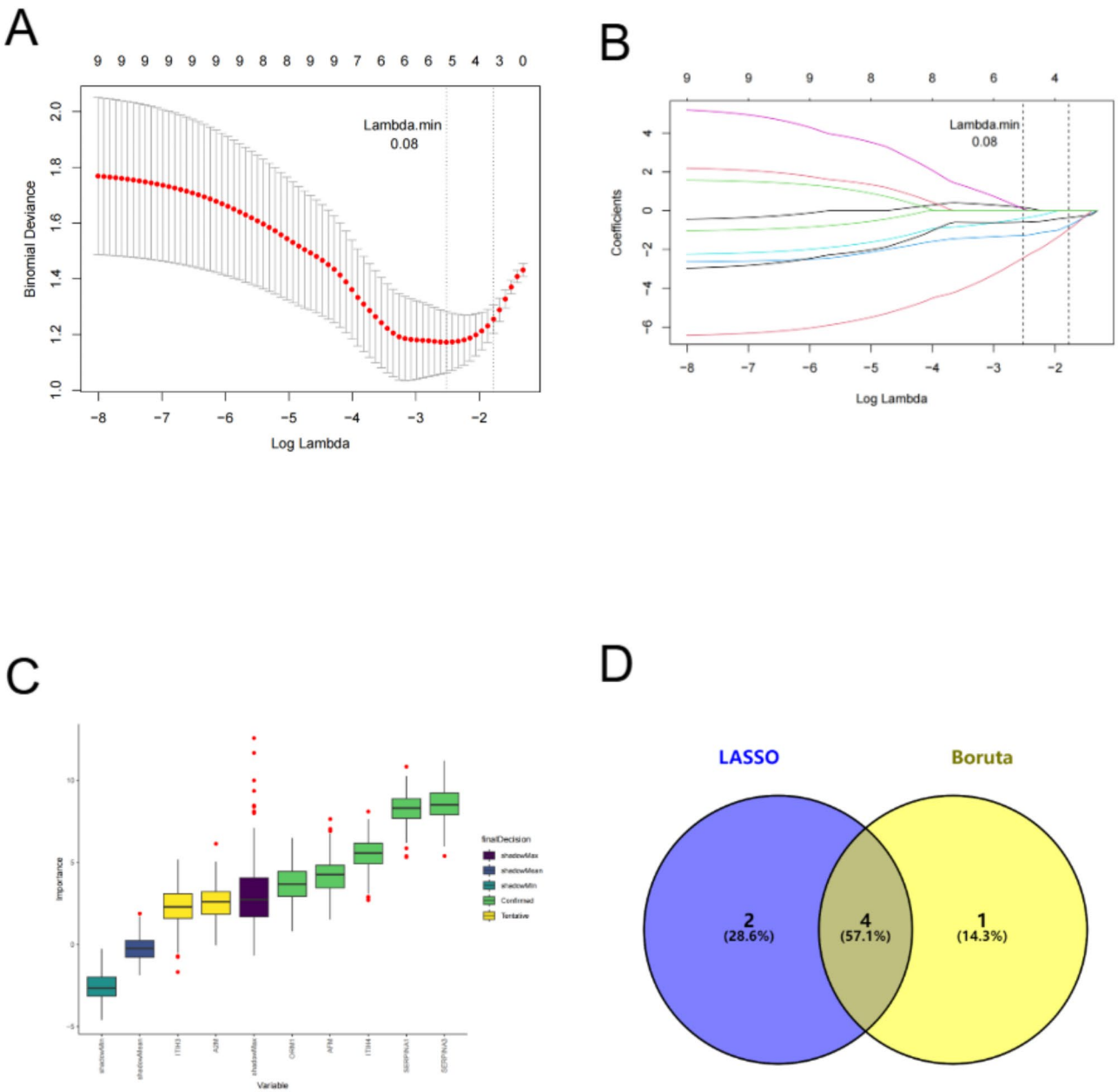


Fig. 3 Machine learning in identifying key diagnosis proteins for HFpEF. **(A-B)** 9 different iterations of the LASSO model's cross-validation process for adjusting parameter selection. Each curve represents a single protein. LASSO analysis of the coefficients. Plotted at the best lambda are vertical dashed lines. **(C)** Boruta Algorithm feature score. **(D)** The intersection of proteins of the above two algorithms is shown in the Venn diagram

activity, endopeptidase regulatory activity, peptidase regulatory activity, enzyme inhibitor activity, serine-type peptidase activity, and endopeptidase inhibitor activity.

Friends analysis for feature proteins

In this study, four feature genes were used for Friends analysis, and the ranking of functional semantic similarity among each gene is shown in Fig. 5B. The results showed that proteins SERPINA1 and SERPINA3 have high semantic similarity among these four genes, indicating that they may play more important functions.

Expression of feature proteins in immune cells and tissue-specific analysis

Expression of feature proteins in immune cells

To elucidate the expression levels of feature proteins in immune cells, we analyzed these levels using the Human Protein Atlas (HPA) database (<https://www.proteinatlas.org/>). The analysis revealed that the proteins AFM and SERPINA3 exhibit minimal expression in immune cells. In contrast, the protein ITIH4 is predominantly expressed in monocytes and T cell groups, while the

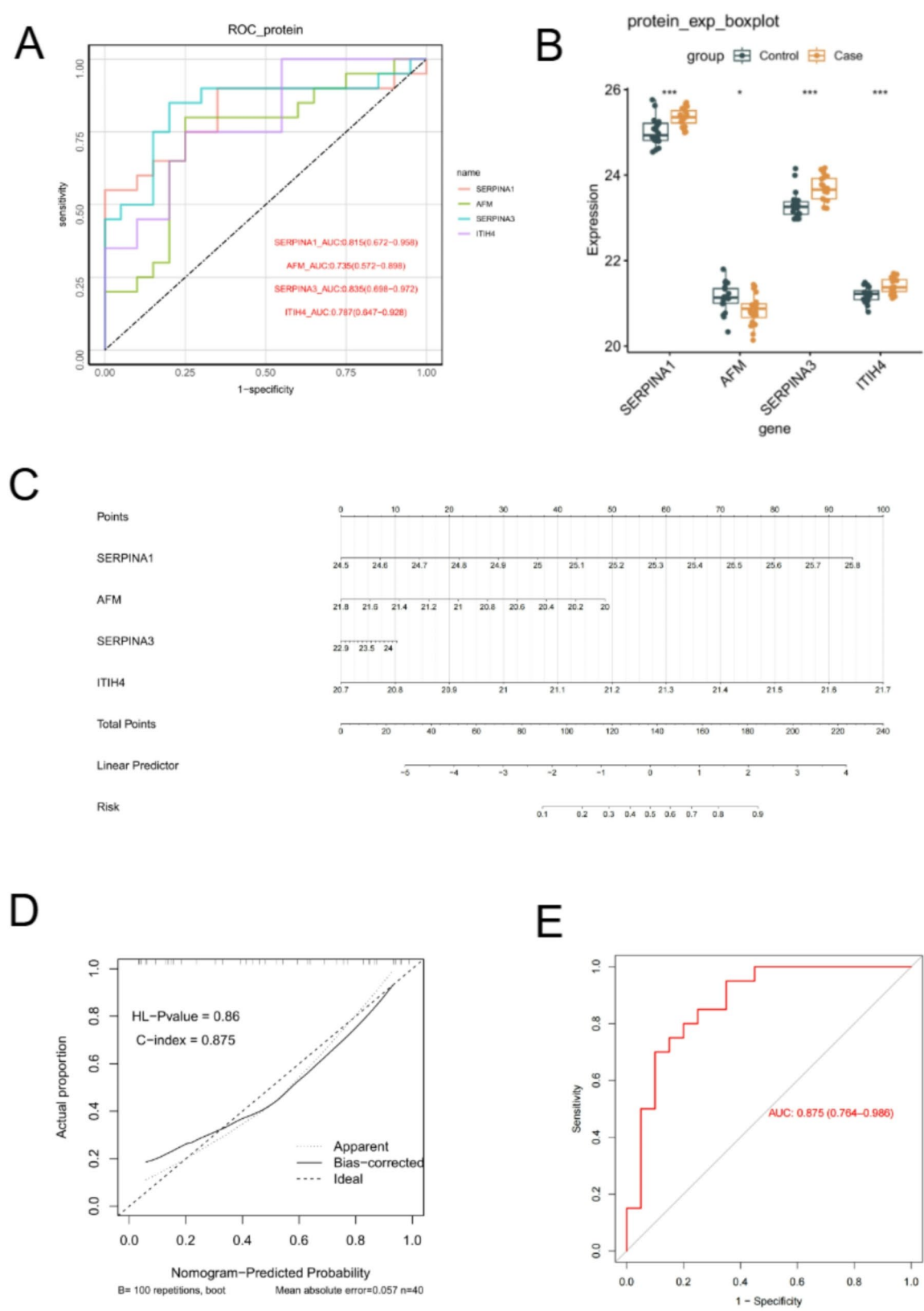


Fig. 4 Construction of the nomogram and the diagnosis value assessment. **(A)** The ROC curve of each candidate protein (SERPINA1、AFM、SERPINA3、ITIH4). **(B)** Boxplot of protein expression of the four candidate proteins. **(C)** Nomogram for diagnosis HFpEF. **(D)** Nomogram correction curve. **(E)** Predictive accuracy of the ROC curve model

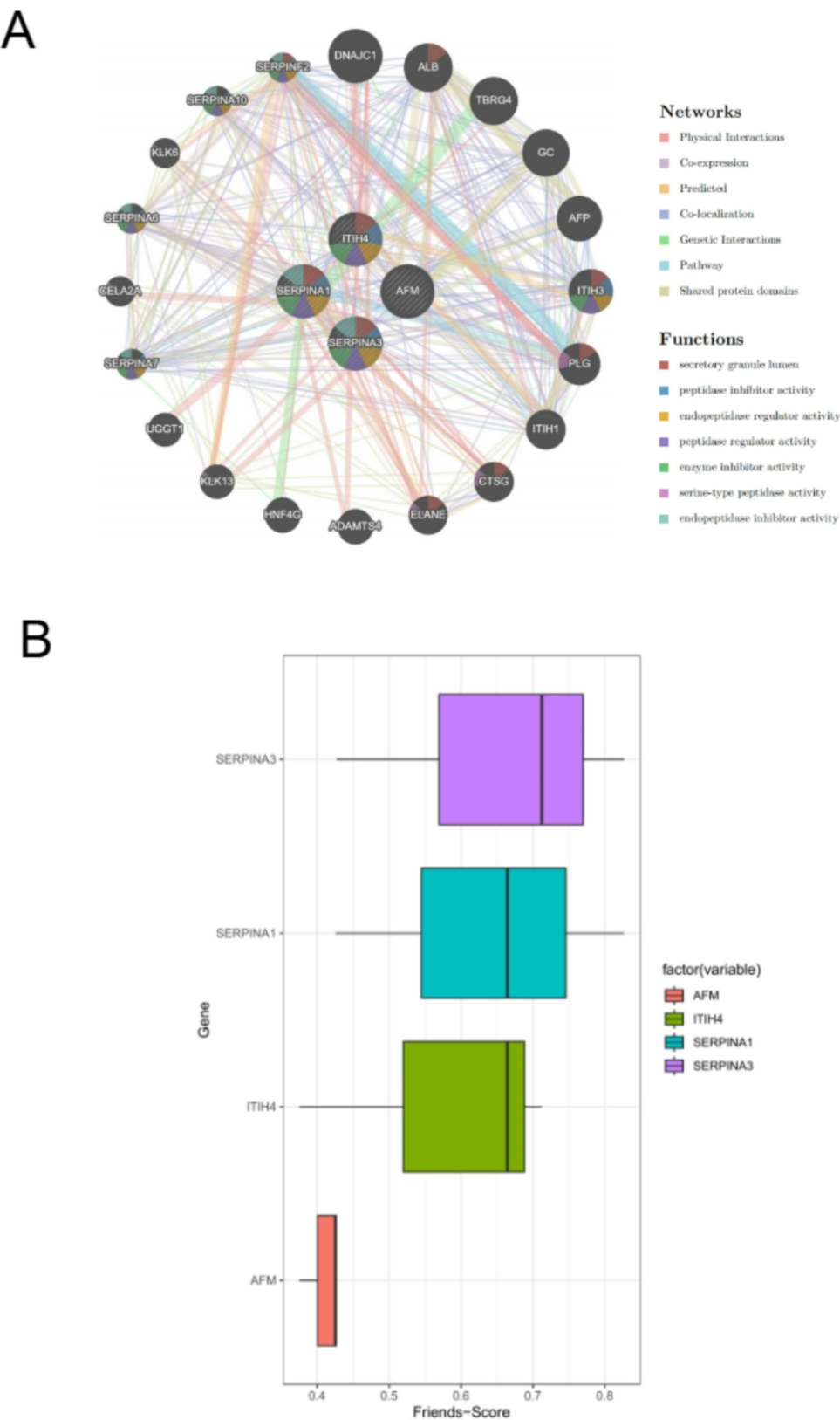


Fig. 5 GGI network and Friends analyses for feature proteins. **(A)** GGI network of DEPs, TOP 20 proteins were selected as key nodes for display. **(B)** Friends Analysis of DEPs, Boxplot of the similarity distribution between proteins

protein SERPINA1 shows significant expression primarily in monocyte groups (Fig. 6A, B).

Expression of feature proteins in tissue-specific analysis

The tissue-specific expression of proteins provides insights into the expression patterns of certain proteins in specific tissues. These patterns are intimately linked to the functions and potential disorders of these tissues or organs, thereby offering a deeper understanding of the affected tissues or organs during disease progression. To delve into the expression levels of 4 feature proteins across various tissues, we utilized the GTEx database (<https://gtexportal.org/home/>) to retrieve this

information. The findings indicate that all the feature proteins identified are predominantly expressed in liver tissues(Figure7A-D).

Plasma levels of SERPINA1, AFM, SERPINA3, and ITIH4 in HFpEF patients and healthy controls

we measured the plasma concentrations of SERPINA3, AFM, SERPINA3, and ITIH4 in HFpEF Patients and HC subjects using Elisa. The results are summarized in Fig. 8A-D.The level of plasma SERPINA3 was significantly higher in HFpEF group than that in HC group ($P < 0.0001$) (Fig. 8A). Plasma SERPINA1 and AFM (Afa-min) levels were higher in HFpEF patients than in HC

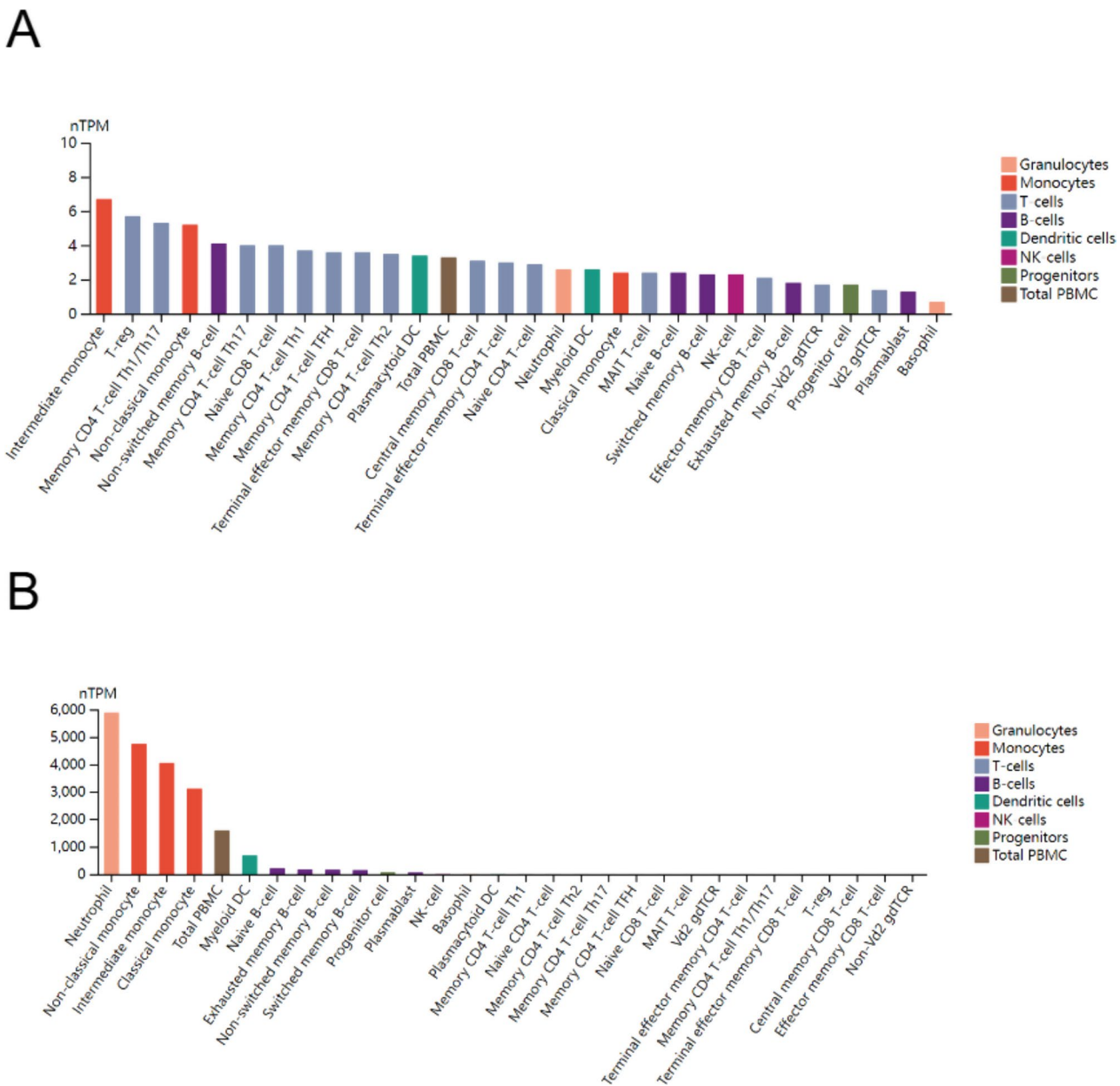


Fig. 6 Expression analysis of hub proteins in immune cells. **(A)** Expression of ITIH 4 in immune cells. **(B)** Expression of SERPINA1 in immune cells

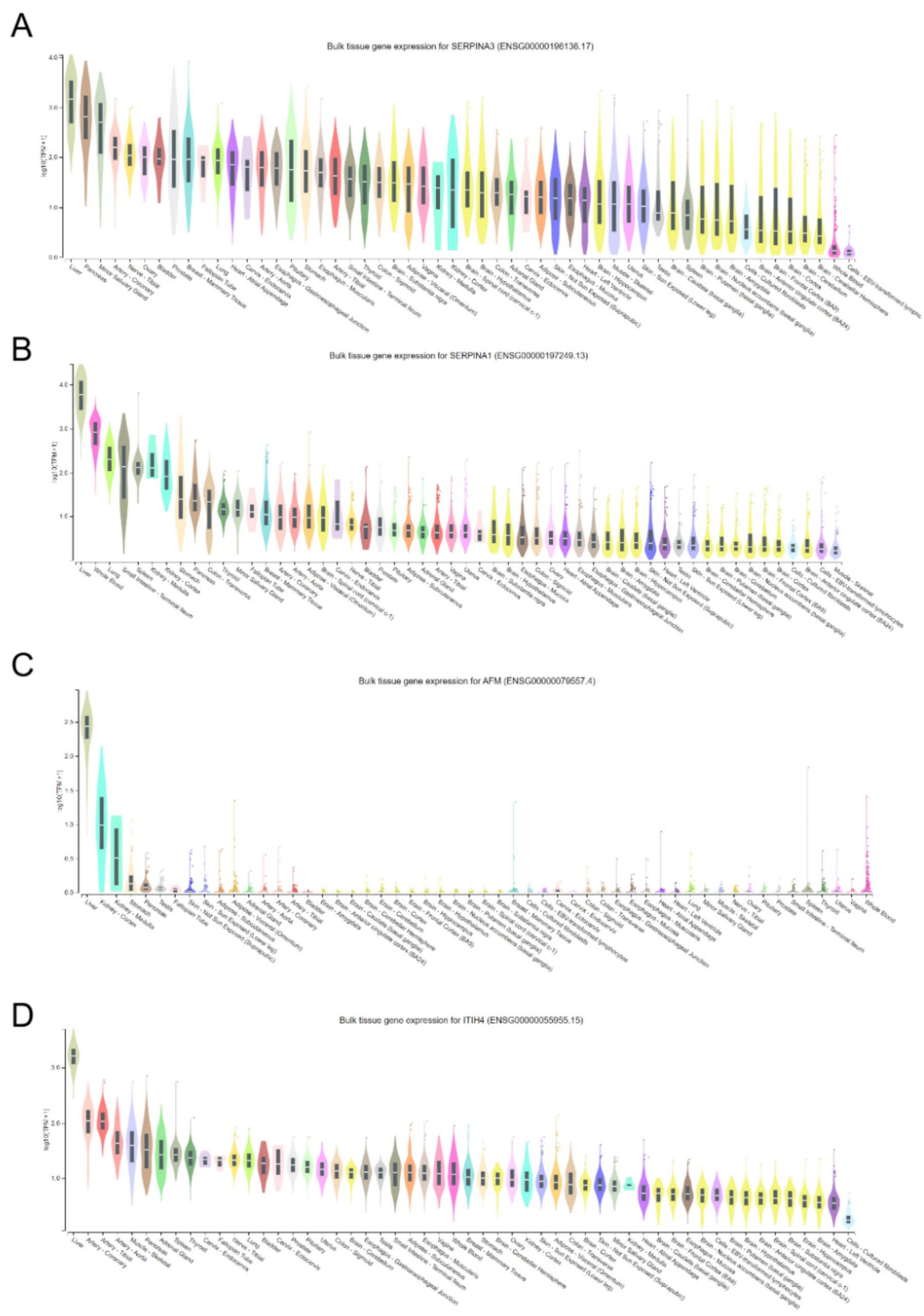


Fig. 7 Expression of feature proteins in tissue-specific analysis. **(A)** Expression of SERPINA3 in tissue-specific analysis. **(B)** Expression of SERPINA1 in tissue-specific analysis. **(C)** Expression of AFM in tissue-specific analysis. **(D)** Expression of ITIH4 in tissue-specific analysis

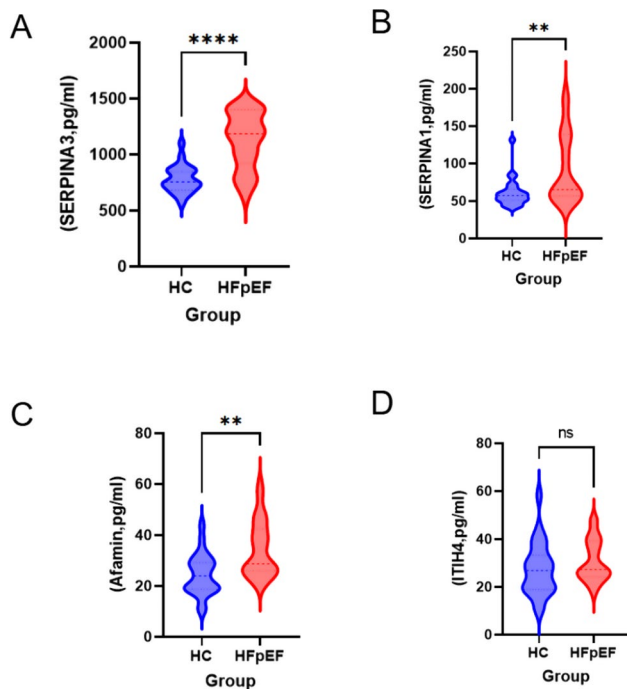


Fig. 8 The expression profile analysis of feature proteins. **(A)** SERPINA3 expression in clinical plasma samples. **(B)** SERPINA1 expression in clinical plasma samples. **(C)** AFM expression in clinical plasma samples. **(D)** ITIH4 expression in clinical plasma samples

group ($p < 0.001$) (Fig. 8B-C); These results demonstrate that SERPINA3 is significantly altered in HFpEF patients compared to HCs, suggesting potential involvement in the disease process and utility as biomarkers for HFpEF. In contrast, ITIH4 levels did not differ significantly, indicating that it may not play a major role or may require further investigation (Figure 8D).

Discussion

This study utilized a combination of proteomics and machine learning to identify SERPINA3 as a potential biomarker for HFpEF. Through differential expression analysis, we identified 34 DEPs between HFpEF patients and healthy controls, with 26 upregulated and 8 down-regulated proteins. Enrichment analyses revealed that these DEPs are involved in acute-phase responses and immune-related pathways, suggesting a significant role of inflammation in HFpEF pathogenesis.

The construction of a PPI network allowed us to identify nine hub proteins with significant interactions. Applying LASSO regression and the Boruta algorithm, we narrowed the candidates to four key proteins: SERPINA1, AFM, SERPINA3, and ITIH4. Among these, SERPINA3 showed the highest diagnostic value, with an AUC of 0.835, indicating good accuracy in distinguishing HFpEF patients from HCs.

Our findings align with previous studies that have implicated SERPINA3 in cardiovascular diseases [11,

27]. SERPINA3 is known to be involved in inflammatory responses and protease inhibition, which are processes relevant to HFpEF pathophysiology. Inflammation is a key component in the pathogenesis of cardiovascular diseases (CVDs), including atherosclerosis, coronary artery disease (CAD), and heart failure (HF) [28].

Molecular mechanisms of immune regulatory proteins in HFpEF pathogenesis

The differential expression patterns of key proteins (SERPINA1, AFM, SERPINA3, ITIH4) in immune cells among HFpEF patients underscore the central role of systemic inflammation and immune homeostasis dysregulation in disease progression. ITIH4 (inter-alpha-trypsin inhibitor heavy chain 4), functioning as an extracellular matrix stabilizer, participates in immune-mediated myocardial remodeling through its specific expression in monocytes/macrophages and T lymphocytes [29]. Clinical studies have demonstrated monocytic infiltration in HFpEF myocardial tissue, where secreted pro-inflammatory cytokines (e.g., IL-6, TNF- α) contribute to reduced ventricular compliance by activating fibroblast-myofibroblast transition and promoting collagen deposition [30]. ITIH4 may counteract protease activity in these cells, regulating tissue injury and repair. SERPINA1 (Alpha-1-antitrypsin) is a serine protease inhibitor with anti-inflammatory properties [31]. During this process, ITIH4 may exert bidirectional regulatory effects on myocardial injury-repair dynamics by inhibiting matrix metalloproteinase activity [32]. SERPINA1 (alpha-1-antitrypsin), a member of the serine protease inhibitor family, protects the myocardial extracellular matrix from excessive degradation through monocyte-specific suppression of hydrolytic enzymes such as neutrophil elastase. Notably, dysregulated SERPINA1 expression due to genetic polymorphisms has been positively correlated with the severity of left ventricular diastolic dysfunction in HFpEF [31]. The low expression levels of AFM and SERPINA3 in immune cells suggest indirect mechanisms: AFM modulates systemic inflammatory states via hepatic lipid metabolism regulation, while SERPINA3, as an acute-phase reactant, reflects persistent subclinical inflammatory activation in HFpEF patients through elevated plasma concentrations [31, 32].

Liver-Specific protein expression and systemic effects

Gene expression profiling from the GTEx database reveals hepatic-specific overexpression of these four proteins, highlighting the liver's pivotal role as an inflammatory regulatory hub in HFpEF pathophysiology. Under chronic low-grade inflammatory stimuli, hepatocytes establish a cardio-hepatic inflammatory axis by secreting acute-phase proteins (e.g., SERPINA3, ITIH4), which exacerbate myocardial pathology through three

key pathways: ① Direct activation of Toll-like receptor signaling in cardiac tissue, promoting M1 macrophage polarization [33]; ② Inhibition of vascular endothelial nitric oxide synthase activity, leading to microvascular endothelial dysfunction and oxidative stress accumulation [34]; ③ Enhancement of fibroblast anabolism via the TGF- β /Smad pathway, driving interstitial fibrosis [35]. Intriguingly, plasma SERPINA3 concentrations show significant positive correlations with left ventricular mass index, suggesting dual mechanisms in myocardial stiffness modulation: maintaining collagen network stability through matrix metalloproteinase inhibition while promoting collagen synthesis via STAT3 pathway activation [35]. This hepatocardiac crosstalk perpetuates a vicious cycle between systemic inflammation and localized myocardial remodeling, offering novel insights for targeted HFpEF therapies [36].

Elevated levels of SERPINA3 have been observed in patients with acute myocardial infarction and unstable angina, suggesting its involvement in acute coronary events [11, 37]. In a recent study, Wouter C. Meijers et al. found that, serpinA3 is elevated in human patients with chronic HF ($n=101$) compared with healthy subjects ($n=180$; $P<0.001$) [38]. In heart failure with reduced ejection fraction (HFrEF), studies have shown alterations in protease inhibitor levels, including SERPINA3, which may reflect the inflammatory and remodeling processes occurring in the failing heart [39–41]. However, the specific role of SERPINA3 in HFpEF not fully elucidated. While direct studies linking SERPINA3 specifically to HFpEF are limited, the involvement of SERPINA3 in inflammatory pathways provides a plausible connection [42]. Given that systemic and myocardial inflammation are critical in HFpEF pathogenesis, SERPINA3 may play a role through its regulation of protease activity and influence on inflammatory processes.

Clinically, SERPINA3 could enhance diagnostic accuracy when combined with existing biomarkers (e.g., NT-proBNP) and imaging modalities. Its elevation may reflect inflammatory pathways central to HFpEF, enabling targeted anti-inflammatory therapies [43]. Furthermore, longitudinal monitoring of SERPINA3 could aid in prognostication and tracking disease progression. Future studies should explore its integration into multi-marker panels to optimize diagnostic algorithms.

The use of DIA-based LC-MS/MS provided a high-throughput and sensitive method to detect protein expression differences between HFpEF patients and HCs [44]. Coupling this with machine learning enabled us to handle the complex dataset effectively, highlighting the value of integrating advanced computational tools in biomedical research [45].

However, our study has several limitations. The sample size, though sufficient for pilot biomarker screening, may

limit generalizability, which may affect the generalizability of the findings. Larger cohorts are needed to validate SERPINA3's diagnostic utility across diverse populations. Additionally, while in vitro experiments suggested an association between SERPINA3 and HFpEF, further studies are required to elucidate the causal relationships and underlying mechanisms. Investigating SERPINA3's role in animal models of HFpEF could provide more comprehensive insights.

Future research should focus on exploring the pathways through which SERPINA3 influences HFpEF development and progression. Longitudinal studies assessing SERPINA3 levels over time could determine its potential as a prognostic marker. Moreover, evaluating the efficacy of interventions targeting SERPINA3 may open new avenues for therapeutic strategies.

Conclusions

This study highlights SERPINA3 as a promising biomarker for HFpEF. The integration of proteomics and machine learning offers a powerful approach to unravel complex disease mechanisms and identify clinically relevant biomarkers. By enhancing diagnostic accuracy and understanding of HFpEF pathogenesis, these findings hold potential for improving patient outcomes through targeted interventions.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12953-025-00242-7>.

Supplementary Material 1

Acknowledgements

Not applicable.

Author contributions

M.A: Writing– original draft, Data curation, Methodology, Conceptualization. S.A: Methodology, Investigation, Formal analysis. N.W: Software, Resources, Project administration. D.D: Supervision. A.A: Project administration. M.N: Data curation, Conceptualization. A.M: Writing– review & editing, Methodology, Investigation.

Funding

This study was supported by the Key R&D Program of Xinjiang Uygur Autonomous Region. [Grant No. 2022B03023-4].

Data availability

Sequence data that support the findings of this study have been deposited in the ProteomeXchange Consortium (<https://proteomecentral.proteomexchange.org/cgi/GetDataset?ID=PXD053164>) via the iProX partner repository with the dataset identifier PXD053164.

Declarations

Ethics approval and consent to participate

The study was approved by the Ethics Committee of the First Affiliated Hospital of Xinjiang Medical University (IRB-K202403-12). All participants provided written informed consent.

Consent for publication

All authors have reviewed and approved the manuscript for submission.

Competing interests

The authors declare no competing interests.

Received: 1 February 2025 / Accepted: 26 March 2025

Published online: 08 April 2025

References

- Borlaug BA, Sharma K, Shah SJ, Ho JE. Heart failure with preserved ejection fraction. *J Am Coll Cardiol*. 2023;81(18):1810–34.
- Bauersachs J, Soltani S. Herzinsuffizienz: Leitlinien-Update der ESC 2023. *Herz*. 2023;49(1):19–21.
- Liu H, Magaye R, Kaye DM, Wang BH. Heart failure with preserved ejection fraction: the role of inflammation. *Eur J Pharmacol*. 2024, 980.
- Abudurexiti M, Abuduhaliq R, Naman T, Wupuer N, Duan D, Keranmu M, Mahemuti A. Integrated proteomic and metabolomic profiling reveals novel insights on the inflammation and immune response in HFpEF. *BMC Genomics*. 2024;25(1):676.
- von Haehling S, Assmus B, Bekfani T, Dworatzek E, Edelmann F, Hashemi D, Hellenkamp K, Kempf T, Raake P, Schütt KA, et al. Heart failure with preserved ejection fraction: diagnosis, risk assessment, and treatment. *Clin Res Cardiol*. 2024;113(9):1287–305.
- Desai AS, Lam CSP, McMurray JJV, Redfield MM. How to manage heart failure with preserved ejection fraction. *JACC: Heart Fail*. 2023;11(6):619–36.
- Abudureyimu M, Luo X, Jiang L, Jin X, Pan C, Yu W, Ge J, Zhang Y, Ren J. FBXL4 protects against HFpEF through Drp1-Mediated regulation of mitochondrial dynamics and the downstream SERCA2a. *Redox Biol*. 2024, 70.
- Liu X, Liu X, Wang Y, Sun H, Guo Z, Tang X, Li J, Xiao X, Zheng S, Yu M et al. Proteome characterization of glaucoma aqueous humor. *Mol Cell Proteom*. 2021, 20.
- Xu M, Deng J, Xu K, Zhu T, Han L, Yan Y, Yao D, Deng H, Wang D, Sun Y, et al. In-depth serum proteomics reveals biomarkers of psoriasis severity and response to traditional Chinese medicine. *Theranostics*. 2019;9(9):2475–88.
- Bermea KC, Lovell JP, Hays AG, Goerlich E, Vungarala S, Jani V, Shah SJ, Sharma K, Adamo L. A machine Learning-Derived score to effectively identify heart failure with preserved ejection fraction. *JACC: Adv*. 2024, 3(7).
- Li B, Lei Z, Wu Y, Li B, Zhai M, Zhong Y, Ju P, Kou W, Shi Y, Zhang X et al. The association and pathogenesis of SERPINA3 in coronary artery disease. *Front Cardiovasc Med*. 2021, 8.
- Zhao J, Pan J. Circulating Serpina3 might be a new potential biomarker to predict the clinical outcomes in AML. *Int J Cardiol*. 2020, 312.
- Paulus WJ. H(2)FPEF score: at last, a properly validated diagnostic algorithm for heart failure with preserved ejection fraction. *Circulation*. 2018;138(9):871–3.
- Pieske B, Tschöpe C, de Boer RA, Fraser AG, Anker SD, Donal E, Edelmann F, Fu M, Guazzi M, Lam CSP, et al. How to diagnose heart failure with preserved ejection fraction: the HFA-PEFF diagnostic algorithm: a consensus recommendation from the heart failure association (HFA) of the European society of cardiology (ESC). *Eur Heart J*. 2019;40(40):3297–317.
- McDonagh TA, Metra M, Adamo M, Gardner RS, Baumbach A, Böhm M, Burri H, Butler J, Čelutkienė J, Chioncel O, et al. Corrigendum to: 2021 ESC guidelines for the diagnosis and treatment of acute and chronic heart failure: developed by the task force for the diagnosis and treatment of acute and chronic heart failure of the European society of cardiology (ESC) with the special contribution of the heart failure association (HFA) of the ESC. *Eur Heart J*. 2021;42(48):4901.
- Zhang H, Wang L, Yin D, Zhou Q, Lv L, Dong Z, Shi Y. Integration of proteomic and metabolomic characterization in atrial fibrillation-induced heart failure. *BMC Genomics*. 2022;23(1):789.
- Zhou X, Liang B, Lin W, Zha L. Identification of MACC1 as a potential biomarker for pulmonary arterial hypertension based on bioinformatics and machine learning. *Comput Biol Med*. 2024, 173.
- Guan S, Xu Z, Yang T, Zhang Y, Zheng Y, Chen T, Liu H, Zhou J. Identifying potential targets for preventing cancer progression through the PLA2G1B Recombinant protein using bioinformatics and machine learning methods. *Int J Biol Macromol*. 2024, 276.
- Ranstam J, Cook JA. LASSO regression. *Br J Surg*. 2018;105(10):1348–1348.
- Chen C, Hou J, Tanner JJ, Cheng J. Bioinformatics methods for mass Spectrometry-Based proteomics data analysis. *Int J Mol Sci*. 2020, 21(8).
- Kong C, Zhu Y, Xie X, Wu J, Qian M. Six potential biomarkers in septic shock: a deep bioinformatics and prospective observational study. *Front Immunol*. 2023;14:1184700.
- Balachandran VP, Gonen M, Smith JJ, DeMatteo RP. Nomograms in oncology: more than Meets the eye. *Lancet Oncol*. 2015;16(4):e173–180.
- Wang K, Li Y, Lin J. Identification of diagnostic biomarkers for osteoarthritis through bioinformatics and machine learning. *Heliyon*. 2024, 10(6).
- Zhang W, Landback P, Gschwend AR, Shen B, Long M. New genes drive the evolution of gene interaction networks in the human and mouse genomes. *Genome Biol*. 2015, 16(1).
- Lin S, Yan J, Wang W, Luo L. STAT3-Mediated ferroptosis is involved in Sepsis-Associated acute respiratory distress syndrome. *Inflammation*. 2024;47(4):1204–19.
- Tian S, Wu L, Zheng H, Zhong X, Yu X, Wu W. Identification of autophagy-related genes in neuropathic pain through bioinformatic analysis. *Hereditas*. 2023, 160(1).
- Delrue L, Vanderheyden M, Beles M, Paolisso P, Di Gioia G, Dierckx R, Verstrecken S, Goethals M, Heggremont W, Bartunek J. Circulating SERPINA3 improves prognostic stratification in patients with a de Novo or worsened heart failure. *ESC Heart Fail*. 2021;8(6):4780–90.
- Jiang Y, Zhang Y, Zhao C. Integrated gene expression profiling analysis reveals SERPINA3, FCN3, FREM1, MNS1 as candidate biomarkers in heart failure and their correlation with immune infiltration. *J Thorac Disease*. 2022;14(4):1106–19.
- Ravindran A, Holappa L, Niskanen H, Skovorodkin I, Kaisto S, Beter M, Kiema M, Selvarajan I, Nurminen V, Aavik E, et al. Translatome profiling reveals Itih4 as a novel smooth muscle cell-specific gene in atherosclerosis. *Cardiovasc Res*. 2024;120(8):869–82.
- Pihl R, Jensen RK, Poulsen EC, Jensen L, Hansen AG, Thøgersen IB, Dobó J, Gál P, Andersen GR, Enghild JJ et al. ITIH4 acts as a protease inhibitor by a novel inhibitory mechanism. *Sci Adv*. 2021, 7(2).
- Insenser M, Vilarrasa N, Vendrell J, Escobar-Morreale HF. Remission of diabetes following bariatric surgery: plasma proteomic profiles. *J Clin Med*. 2021, 10(17).
- Park J, Kim H, Kim SY, Kim Y, Lee JS, Dan K, Seong MW, Han D. In-depth blood proteome profiling analysis revealed distinct functional characteristics of plasma proteins between severe and non-severe COVID-19 patients. *Sci Rep*. 2020;10(1):22418.
- Tschöpe C, Van Linthout S. New insights in (inter)cellular mechanisms by heart failure with preserved ejection fraction. *Curr Heart Fail Rep*. 2014;11(4):436–44.
- Rech M, Barandiarán Aizpurua A, van Empel V, van Bilsen M, Schroen B. Pathophysiological Understanding of HFpEF: MicroRNAs as part of the puzzle. *Cardiovasc Res*. 2018;114(6):782–93.
- Yang D, Liu HQ, Liu FY, Tang N, Guo Z, Ma SQ, An P, Wang MY, Wu HM, Yang Z, et al. The roles of noncardiomyocytes in cardiac remodeling. *Int J Biol Sci*. 2020;16(13):2414–29.
- Castillo EC, Vázquez-Garza E, Yee-Trejo D, García-Rivas G, Torre-Amione G. What is the role of the inflammation in the pathogenesis of heart failure?? *Curr Cardiol Rep*. 2020;22(11):139.
- Wu D, Guo M, Robinson CV. Connecting single-nucleotide polymorphisms, glycosylation status, and interactions of plasma Serine protease inhibitors. *Chem*. 2023;9(3):665–81.
- Meijers WC, Maglione M, Bakker SJJ, Oberhuber R, Kieneker LM, de Jong S, Haubner BJ, Nagengast WB, Lyon AR, van der Vegt B, et al. Heart failure stimulates tumor growth by Circulating factors. *Circulation*. 2018;138(7):678–91.
- Zhao L, Guo Z, Wang P, Zheng M, Yang X, Liu Y, Ma Z, Chen M, Yang X. Proteomics of epicardial adipose tissue in patients with heart failure. *J Cell Mol Med*. 2019;24(1):511–20.
- Zhou L, Peng F, Li J, Gong H. Exploring novel biomarkers in dilated cardiomyopathy-induced heart failure by integrated analysis and in vitro experiments. *Experimental Therapeutic Med*. 2023, 26(1).
- Attachaipanich T, Chattipakorn SC, Chattipakorn N. Current evidence regarding the cellular mechanisms associated with cancer progression due to cardiovascular diseases. *J Translational Med*. 2024, 22(1).
- Hage C, Michaëlsson E, Linde C, Donal E, Daubert J-C, Gan L-M, Lund LH. Inflammatory biomarkers predict heart failure severity and prognosis in patients with heart failure with preserved ejection fraction. *Circulation: Cardiovasc Genet*. 2017, 10(1).

43. Correction to et al. Sun. A Novel Regulatory Mechanism of Smooth Muscle α -Actin Expression by NRG-1/circACTA2/miR-548f-5p Axis. *Circ Res*. 2017;121:628–635. <https://doi.org/10.1161/CIRCRESAHA.117.311441>. *Circ Res* 2021, 128(1):e25.
44. Lou R, Cao Y, Li S, Lang X, Li Y, Zhang Y, Shui W. Benchmarking commonly used software suites and analysis workflows for DIA proteomics and phosphoproteomics. *Nat Commun* 2023, 14(1).
45. Verdonk C, Verdonk F, Dreyfus G. How machine learning could be used in clinical practice during an epidemic. *Crit Care* 2020, 24(1).

Publisher's note

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