Check for updates

*CORRESPONDENCE
Yitao Xue
✉ xytsdzydfy@126.com
Yan Li
✉ liyan88130@163.com

SMOC2, OGN, FCN3, and SERPINA3 could be biomarkers for the evaluation of acute decompensated heart failure caused by venous congestion

Yiding Yu[1], Huajing Yuan[1], Quancheng Han[1], Jingle Shi[1], Xiujuan Liu[2], Yitao Xue[2]* and Yan Li[2]*

[1]Shandong University of Traditional Chinese Medicine, Jinan, China, [2]Affiliated Hospital of Shandong University of Traditional Chinese Medicine, Jinan, China

Background: Venous congestion (VC) sets in weeks before visible clinical decompensation, progressively increasing cardiac strain and leading to acute heart failure (HF) decompensation. Currently, the field lacks a universally acknowledged gold standard and early detection methods for VC.

Methods: Using data from the GEO database, we identified VC's impact on HF through key genes using Limma and STRING databases. The potential mechanisms of HF exacerbation were explored via GO and KEGG enrichment analyses. Diagnostic genes for acute decompensated HF were discovered using LASSO, RF, and SVM-REF machine learning algorithms, complemented by single-gene GSEA analysis. A nomogram tool was developed for the diagnostic model's evaluation and application, with validation conducted on external datasets.

Results: Our findings reveal that VC influences 37 genes impacting HF via 8 genes, primarily affecting oxygen transport, binding, and extracellular matrix stability. Four diagnostic genes for HF's pre-decompensation phase were identified: SMOC2, OGN, FCN3, and SERPINA3. These genes showed high diagnostic potential, with AUCs for each gene exceeding 0.9 and a genomic AUC of 0.942.

Conclusions: Our study identifies four critical diagnostic genes for HF's pre-decompensated phase using bioinformatics and machine learning, shedding light on the molecular mechanisms through which VC worsens HF. It offers a novel approach for clinical evaluation of acute decompensated HF patient congestion status, presenting fresh insights into its pathogenesis, diagnosis, and treatment.

KEYWORDS

heart failure, venous congestion, machine learning, bioinformatics, nomogram

Introduction

Despite adherence to guideline-directed medical therapies, the prognosis for heart failure (HF) patients remains suboptimal. Individuals in advanced stages of HF often necessitate recurrent hospital admissions and enduring pharmacological intervention, imposing significant demands on healthcare resources ([1]). The primary catalyst for these hospitalizations is the clinical manifestations of venous congestion (VC). A diminution in venous capacitance coupled with an augmented return of venous blood to the heart escalates preload. This increase in preload, when juxtaposed with

compromised cardiac functionality, precipitates a transition from chronic HF to acute decompensated heart failure (ADHF) (2, 3).

HF precipitates fluid retention and VC, transitioning the body from a low-pressure, healthy biological system to a high-pressure, pathological state. This shift forces organs to operate under significantly elevated venous and interstitial pressures, challenging their physiological functions. Traditional views attribute congestion primarily to deteriorating cardiovascular function, with treatment strategies aimed at achieving euvolemic status and congestion alleviation through diuretic therapy (4). However, emerging evidence suggests that venous congestion may initiate weeks prior to the clinical manifestations of decompensation that necessitate medical intervention (5). Biomechanical forces, such as increased venous pressure and shear stress, can induce endothelial cells to undergo a phenotypic transformation towards pro-oxidative, pro-inflammatory, and vasoconstrictive states (6). Biomechanical stress triggers endothelial cells to secrete cytokines, including TNF-α, IL-6, and VCAM-1 (7, 8). Yet, the intricate linkage and precise mechanisms underlying endothelial activation due to VC in the context of HF require further investigation. A deeper understanding of these mechanisms is critical for evaluating the clinical efficacy and optimizing the application of diuretic therapies. Concurrently, identifying pivotal biomarkers could pave the way for early detection of VC—a domain where a definitive diagnostic gold standard is currently absent. By identifying and targeting these key markers, there is potential to preemptively shift treatment modalities, thereby reinstating endothelial cell stability and preventing the progression to a decompensated state in heart failure patients.

This study analyzed the mRNA data sets of VC and HF in the GEO database and identified differentially expressed genes (DEGs) through the limma package. We performed immune infiltration analysis on the VC dataset to assess whether venous pressure affects immune cell composition. Subsequently, we obtained the interaction relationship between VC and HF differential genes through the STRING database and established a protein-protein interaction network (PPI). Through PPI, we obtained the key genes that HF is affected by VC. We performed enrichment analysis on these genes to obtain the impact of VC on the biological processes and functions of HF. Subsequently, we screened key genes through three machine learning models including least absolute shrinkage and selection operator (LASSO), random forest (RF), and support vector machine recursive feature elimination (SVM-REF). We established a diagnostic model for pre-decompensated heart failure and completed nomogram evaluation and GSEA of individual genes. We also validated the diagnostic model on external datasets. Figure 1 depicts the study flowchart.

Materials and methods

Microarray data

We choose GSE38783 as the VC data set (9). The trial involved 12 healthy subjects and inflated a pressure cuff around the

dominant arm to increase arm venous pressure to approximately 30 mmHg above baseline for 75 min. Endothelial cell samples were procured both pre- and post-intervention through an angiocatheter assisted by an intravascular guidewire. For the exploration of heart failure, the GSE57338 dataset was chosen, representing the most comprehensive dataset available in this domain. It comprises left ventricular myocardial specimens from 95 patients diagnosed with ischemic heart failure, alongside 136 control samples from healthy individuals (10). To corroborate our findings, the GSE16499 dataset served as the validation cohort, encompassing left ventricular myocardial samples from 15 patients with ischemic heart failure and an equivalent number of healthy controls (11).

Data processing and differentially expressed gene screening

We used R software (R version 4.2.0) to complete data preprocessing. We remove probes corresponding to multiple molecules. For multiple probes corresponding to the same molecule, only the probe with the largest signal value is retained. Finally, we remove the batch effect of the data and convert the probe ID into a gene symbol according to the annotation file of the platform. We used the limma package for differential analysis and selected genes with p-value <0.05 and $|\log2(FC)| \geq 1$ as differential genes (12). We use the ggplot2 package to complete the drawing of the picture.

Immune infiltration analysis

Endothelial cell activation will produce cytokines. Therefore, we used the CIBERSORT package to evaluate whether venous pressure has an impact on the endothelial immune environment (13). Bar charts are used to visualize the proportion of each type of immune cell in different samples. The differences in cell distribution between VC and normal groups were compared by t test, and the cutoff value was set at $p < 0.05$.

Protein–protein interaction network construction

In order to understand the process of VC affecting HF and discover the interactions between protein-coding genes, we established a PPI network using the STRING database (14). Parameter selection default settings. Subsequently, we imported the results into Cytoscape3.6.1 to complete visualization and subsequent analysis (15).

Functional enrichment analysis

Utilizing the Protein-Protein Interaction (PPI) network, we identified crucial genes implicated in Heart Failure (HF) that are

FIGURE 1
The study flowchart.
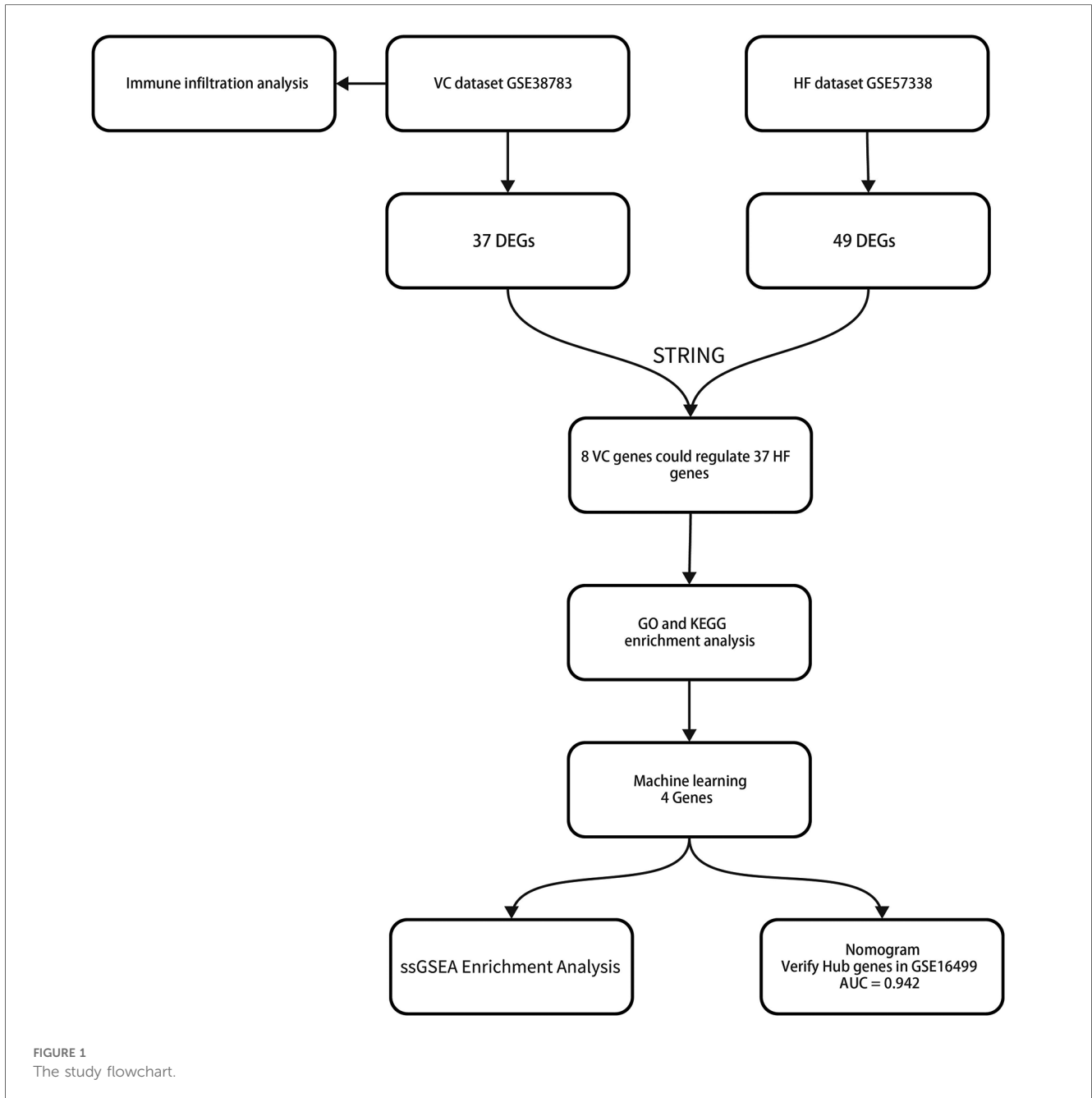
influenced by Venous Congestion (VC). The "clusterProfiler" package facilitated the execution of Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses, enabling the comprehensive visualization and elucidation of these key genes' roles and pathways (16–18). In addition, we also performed Reactome pathway enrichment analysis on key genes.

Machine learning

To refine the selection of diagnostic markers for the pre-decompensation stage of heart failure, our study employs a triad of machine learning algorithms: Least Absolute Shrinkage and Selection Operator (LASSO), Random Forest (RF), and Support Vector Machine-Recursive Feature Elimination (SVM-RFE) (19–21). The LASSO algorithm was implemented using the "glmnet" package, with ten-fold cross-validation employed to identify the most significant genes. The "randomForest" package was utilized to execute the RF algorithm, prioritizing genes based on their importance scores. Similarly, the SVM-RFE algorithm was conducted using the "e1071" package, selecting genes that contribute to the highest classification accuracy. After completing the calculation, we selected the intersection of the three as the diagnostic gene for the pre-decompensation stage of heart failure.

Nomogram construction and validation of diagnostic model

For the construction of a diagnostic nomogram for the precompensation phase of heart failure, our analysis leveraged the "rms" package, which facilitated the visualization of the contributory weight of each candidate gene as individual points on the nomogram, with "Total Points" reflecting the cumulative score derived from all selected genes (22). Subsequently, boxplots were generated to illustrate the expression profiles of these genes, and Receiver Operating Characteristic (ROC) curves were plotted to assess their diagnostic efficacy. The diagnostic utility was quantified using the Area Under the Curve (AUC), with values exceeding 0.7 deemed indicative of substantial diagnostic merit. To validate our findings, we conducted an analysis on both individual and combined gene sets within the GSE16499 dataset. The discriminatory power of our diagnostic model was further appraised through the ROC curve analysis, ensuring a robust evaluation of its predictive capacity.

ssGSEA enrichment analysis

To elucidate the functional roles of genes identified in the pre-decompensated stage of heart failure, our study employed the "clusterProfiler" package to conduct single-gene Gene Set Enrichment Analysis (GSEA) (23). This analysis was pivotal in uncovering the biological processes and pathways through which these genes may contribute to the transition from chronic to acute decompensated heart failure.

Results

Identification of differentially expressed genes

Following a differential analysis, we discerned a total of 37 distinct genes within the VC dataset, comprising 36 up-regulated

and a single down-regulated gene. In the HF dataset, we ascertained a total of 49 distinct genes, inclusive of 28 up-regulated and 21 down-regulated genes. Notably, there exists no overlap between the differential genes identified in the two datasets. The illustrative outcomes are depicted in Figure 2.
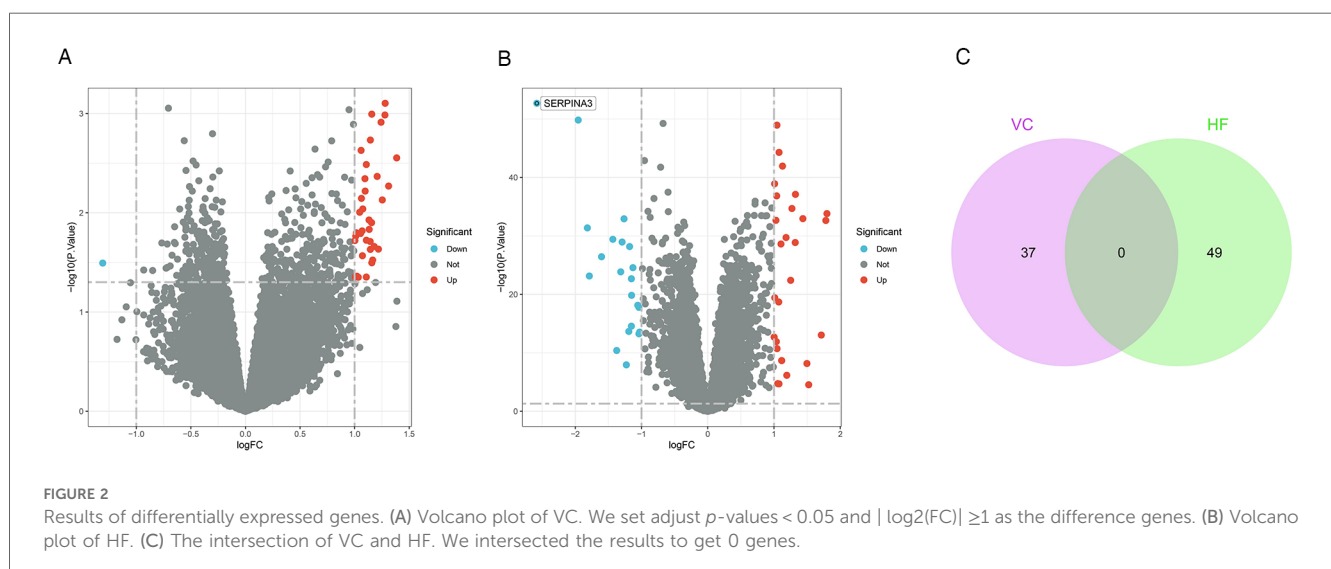
Immune cell infiltration analysis

We performed immune infiltration analysis on the VC data set through the CIBERSORT algorithm. Figure 3 clearly shows the different subpopulation contents within each sample. The results showed that 75 min of venous hypertension did not have a significant impact on the immune environment of endothelial cells.

PPI network construction and hub gene selection

Owing to the lack of overlap between the differential genes of VC and HF, we employed the STRING database to construct a Protein-Protein Interaction (PPI) network, thereby elucidating the interactions among the encoded proteins of VC and HF. Through this PPI network, we pinpointed 45 nodes and 75 interactions. The analysis implied that 8 genes from VC might interact with 37 genes from HF. The visualization results are shown in Figure 4.

Functional enrichment analysis

To evaluate the effects of VC on the biological processes and functions of HF, we performed GO analysis, KEGG enrichment analysis, and Reactome pathway enrichment analysis on 37 HF genes. Categories of GO analysis include biological processes (BP), cellular components (CC), and molecular functions (MF). The predominantly enriched BP terms encompassed those

FIGURE 2
Results of differentially expressed genes. **(A)** Volcano plot of VC. We set adjust p-values < 0.05 and | log2(FC)| ≥1 as the difference genes. **(B)** Volcano plot of HF. **(C)** The intersection of VC and HF. We intersected the results to get 0 genes.
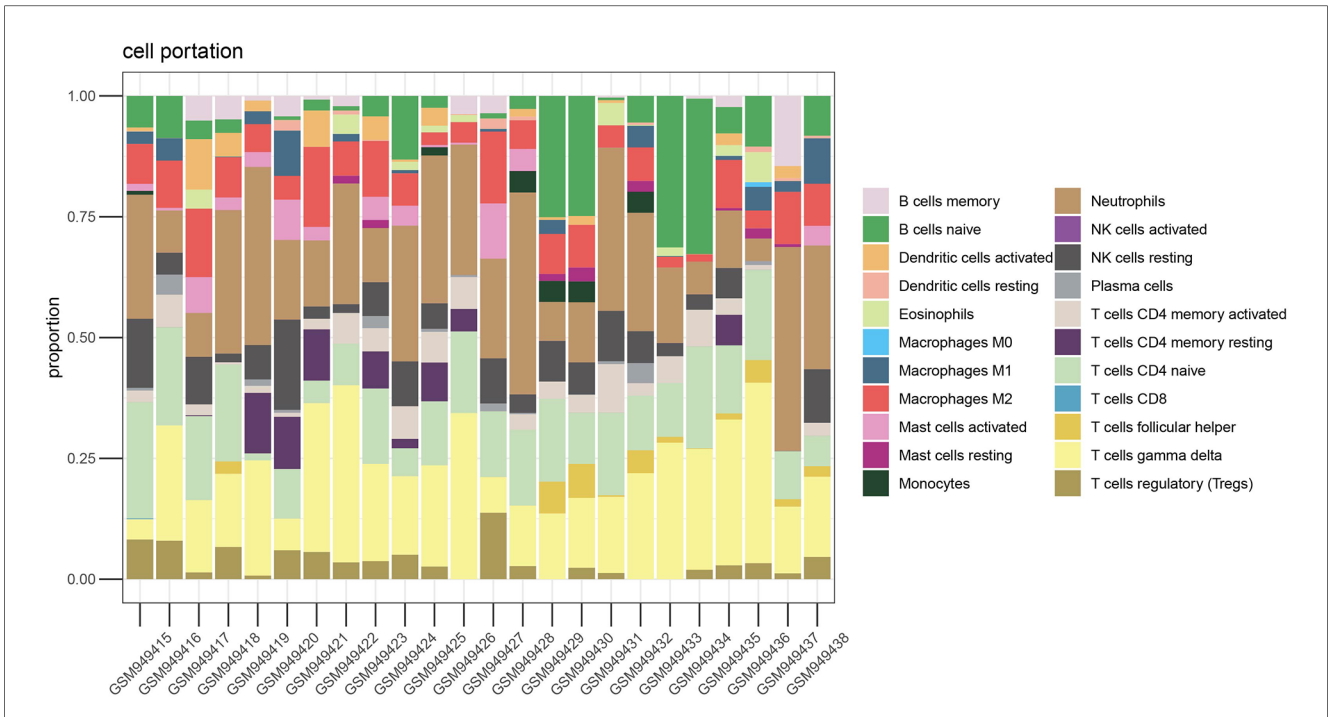
FIGURE 3
Immune infiltration analysis before and after venous hypertension treatment. The proportion of 22 kinds of immune cells in different samples visualized from the bar plot.
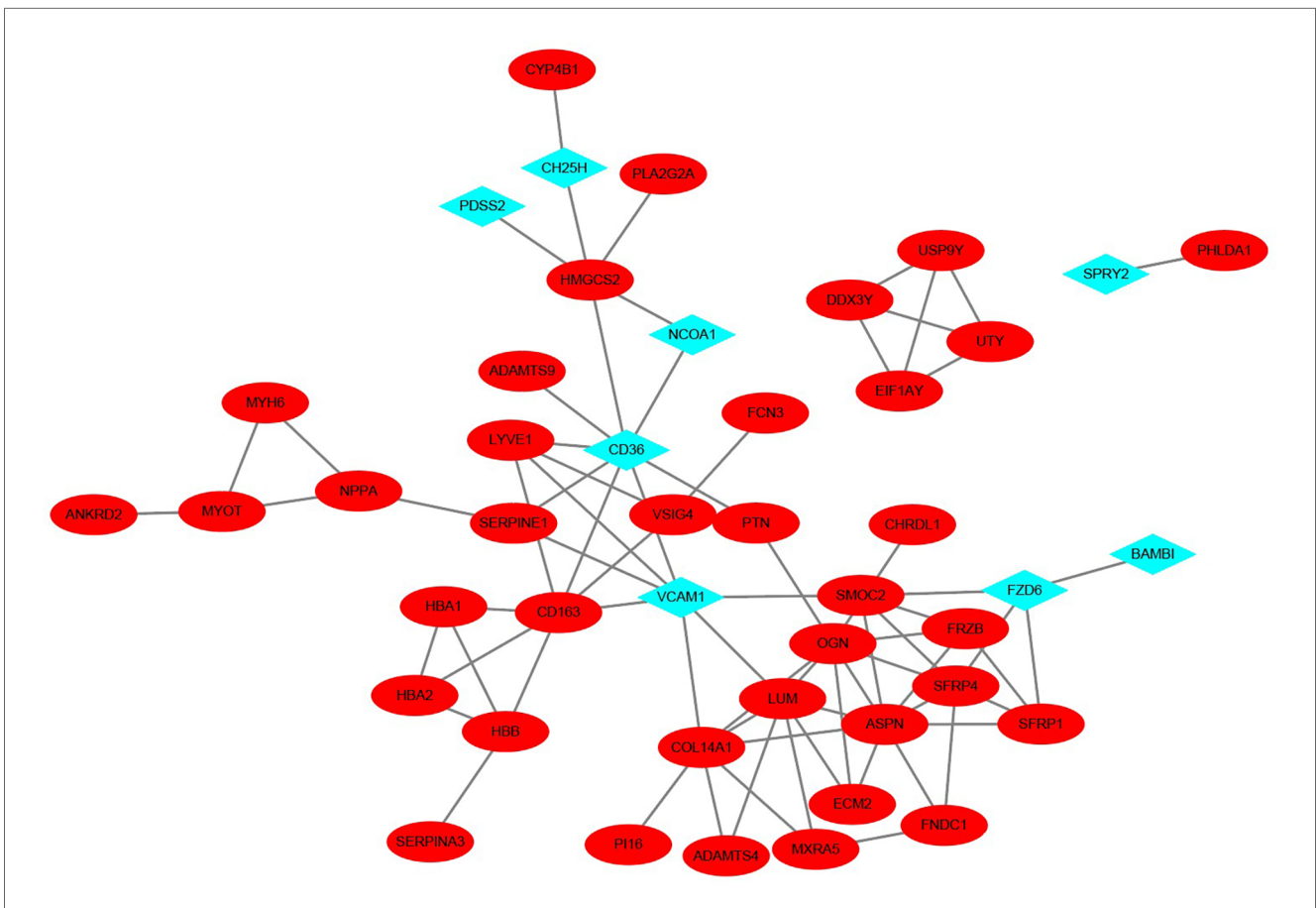

FIGURE 4
PPI network of VC and HF. The blue diamond-shaped node represents the VC gene, and the red oval-shaped node represents the HF gene.

related to oxygen transport, gas transport, extracellular matrix organization, extracellular structure organization, and the organization of external encapsulating structures. The chiefly enriched CC terms included the collagen-containing extracellular matrix, the haptoglobin-hemoglobin complex, the hemoglobin complex, blood microparticles, and the lumen of endocytic vesicles. The MF terms that were most enriched involved haptoglobin binding, oxygen binding, oxygen carrier activity, collagen binding, and the extracellular matrix structural constituent conferring compression resistance. In the KEGG pathway enrichment analysis, these 37 genes were primarily enriched in pathways associated with African trypanosomiasis, Malaria, the Wnt signaling pathway, and the Complement and coagulation cascades. Reactome pathway enrichment analysis showed that these 37 HF genes were mainly related to oxygen transport and exchange, oxidative stress, and cell structure maintenance. The enrichment results were shown in Figure 5.

Identification of Hub genes via machine learning

We employed three machine learning algorithms—LASSO, RF, and SVM-RFE—to identify diagnostic genes for the pre-decompensation stage of heart failure. The LASSO algorithm revealed 15 potential genes, while the RF algorithm prioritized genes based on their calculated significance, selecting the top six genes with the utmost importance as candidates. The SVM-RFE algorithm indicated optimal accuracy with 29 genes; hence, we adopted the initial 29 genes from the SVM-RFE output as candidate genes. By intersecting the findings from all three algorithms, we pinpointed four diagnostic genes for the pre-decompensation stage of heart failure: SMOC2, OGN, FCN3, and SERPINA3. These results are displayed in Figure 6.
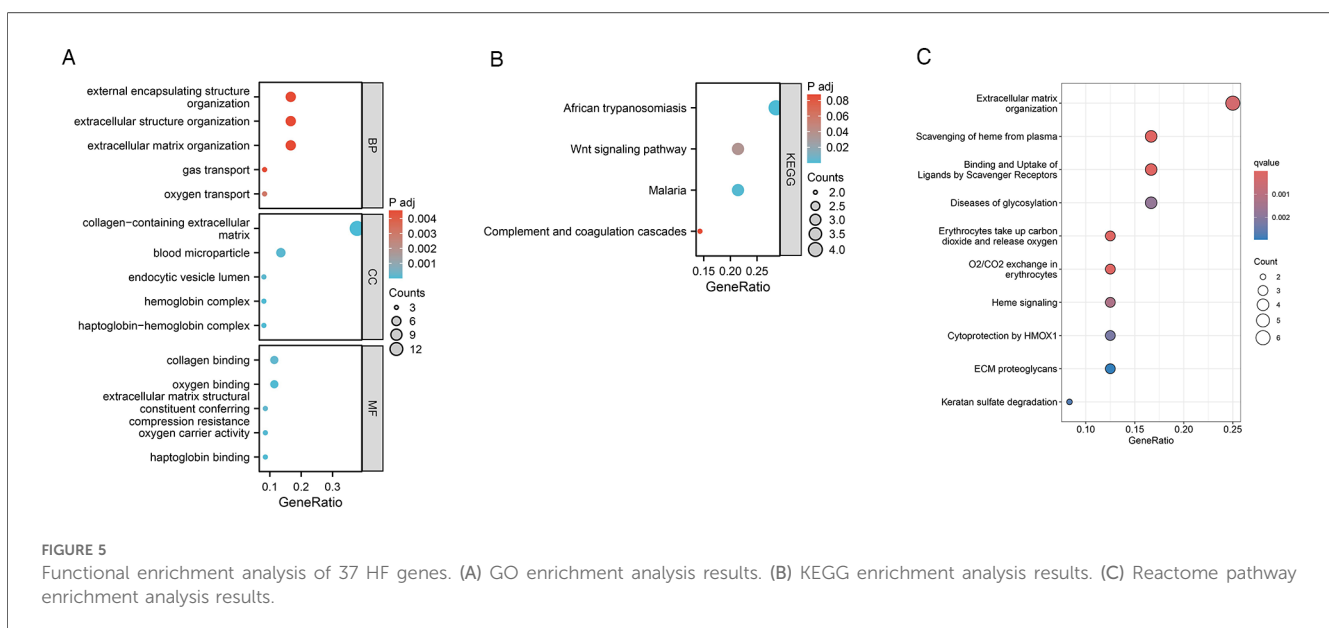
Diagnostic value assessment

We devised a nomogram anchored on four pivotal hub genes, accompanied by the formulation of a Receiver Operating Characteristic (ROC) curve, to meticulously gauge the diagnostic specificity and sensitivity conferred by each gene and the nomogram as a whole. Furthermore, we illustrated the expression disparities of these hub genes within the heart failure (HF) dataset via a box plot. Subsequent validation of the hub gene on the external dataset GSE16499 demonstrated robust diagnostic efficacy, with the Area Under the Curve (AUC) for each gene surpassing 0.9 and the collective genomic AUC reaching 0.942, underscoring significant diagnostic potential. The visual representation of these findings was encapsulated in Figure 7.

Single-sample gene set enrichment analysis

We performed ssGSEA on SMOC2, OGN, FCN3, and SERPINA3, respectively. The analysis revealed that these genes contribute variably to the biosynthesis and metabolism of amino acids, the production of steroid compounds, and the synthesis of nucleotide sugars. Comprehensive results and their visual representations are detailed in Figure 8.

Discussion

Venous congestion induces endothelial and neurohormonal activation, signifying a decline in cardiac function. This activation is not only a consequence but also a catalyst for further hemodynamic decline in heart failure, often precipitating the transition from chronic to acute decompensated heart failure. Consequently, venous congestion emerges as a crucial hemodynamic indicator for predicting rehospitalization and

FIGURE 5
Functional enrichment analysis of 37 HF genes. **(A)** GO enrichment analysis results. **(B)** KEGG enrichment analysis results. **(C)** Reactome pathway enrichment analysis results.

FIGURE 6

Machine learning in screening candidate diagnostic biomarkers for pre-decompensation stage of HF. **(A)** In the Lasso model for biomarker screening, we plotted LASSO coefficient profiles of the selected genes, determining the optimal lambda at the point where the model's error was minimal. Each line in the left graph represents a gene, while the right graph's vertical lines indicate the model's error, pinpointing the best gene selection count at 15. **(B)** In the RF model, we focused on the relative importance of the candidate genes, specifically highlighting the top six genes as determined by the random forest's importance calculation. **(C)** The SVM-RFE model's screening involved utilizing the algorithm to identify the most accurate genes, demonstrated by the plot where the number of genes selected is balanced against their predictive accuracy. Based on the highest accuracy and lowest error obtained from the curves, we selected the top 29 genes as candidate genes. **(D)** A Venn diagram consolidates the outcomes, indicating four genes consistently identified across the three analytical methods.

mortality post-discharge in patients with acute decompensated heart failure (24). However, it's pivotal to recognize that venous congestion manifests in the initial phase of decompensation. Deciphering its impact on cardiac function and identifying early diagnostic markers are essential for timely modulation of therapeutic strategies.

In this study, we first obtained the differential genes of VC and HF through differential analysis. Regrettably, there was no overlap between the differential genes for VC and HF, suggesting that the pathogenesis of heart failure during stable periods devoid of congestion might not be directly linked to VC. However, utilizing the Protein-Protein Interaction (PPI) network, we discovered that subsequent to the onset of VC, eight differential genes in the vascular endothelium could potentially influence 37 genes associated with HF. Therefore, VC may affect HF through signal transduction between molecules. These 37 genes might be pivotal in the transition of heart failure patients from a stable condition to the acute decompensated stage.

We explored the biological functions and signaling pathways involved in key genes through GO analysis, KEGG enrichment analysis, and Reactome pathway enrichment analysis. The results show that the mechanism by which VC causes acute decompensation of heart failure may be related to oxygen transport and oxygen binding. Heart failure is characterized by exercise intolerance, which is closely related to structural and functional abnormalities in oxygen transport (25). When heart function decreases, the heart cannot pump blood effectively and meet the body's demand for oxygen (26). It can be seen from the PPI network that after the occurrence of VC, VCAM1, CD36 and NCOA1 of endothelial cells change, affecting the HAB1, HAB2, and HABB genes of HF, which may further reduce oxygen transport and oxygen binding and other functions. Hemoglobin plays a key physiological role in transporting oxygen in the human body. However, VC may increase levels of haptoglobin, an acute-phase serum protein that binds very strongly and irreversibly to hemoglobin and is endocytosed by macrophages (27). At the same time, CD36 and VCAM1 may induce an increase in CD163 levels after endothelial activation, degrade the hemoglobin-haptoglobin complex, and produce an inflammatory response (28). The enrichment analyses further suggest that VC may influence collagen binding and the organization of the myocardial extracellular matrix. This matrix serves as a sophisticated, evolving scaffold that upholds tissue structure and function (29). Remodeling of the extracellular matrix can hinder cardiac filling, exacerbating heart failure through an inflammatory response (30). Moreover, this remodeling process can alter the collagen content,
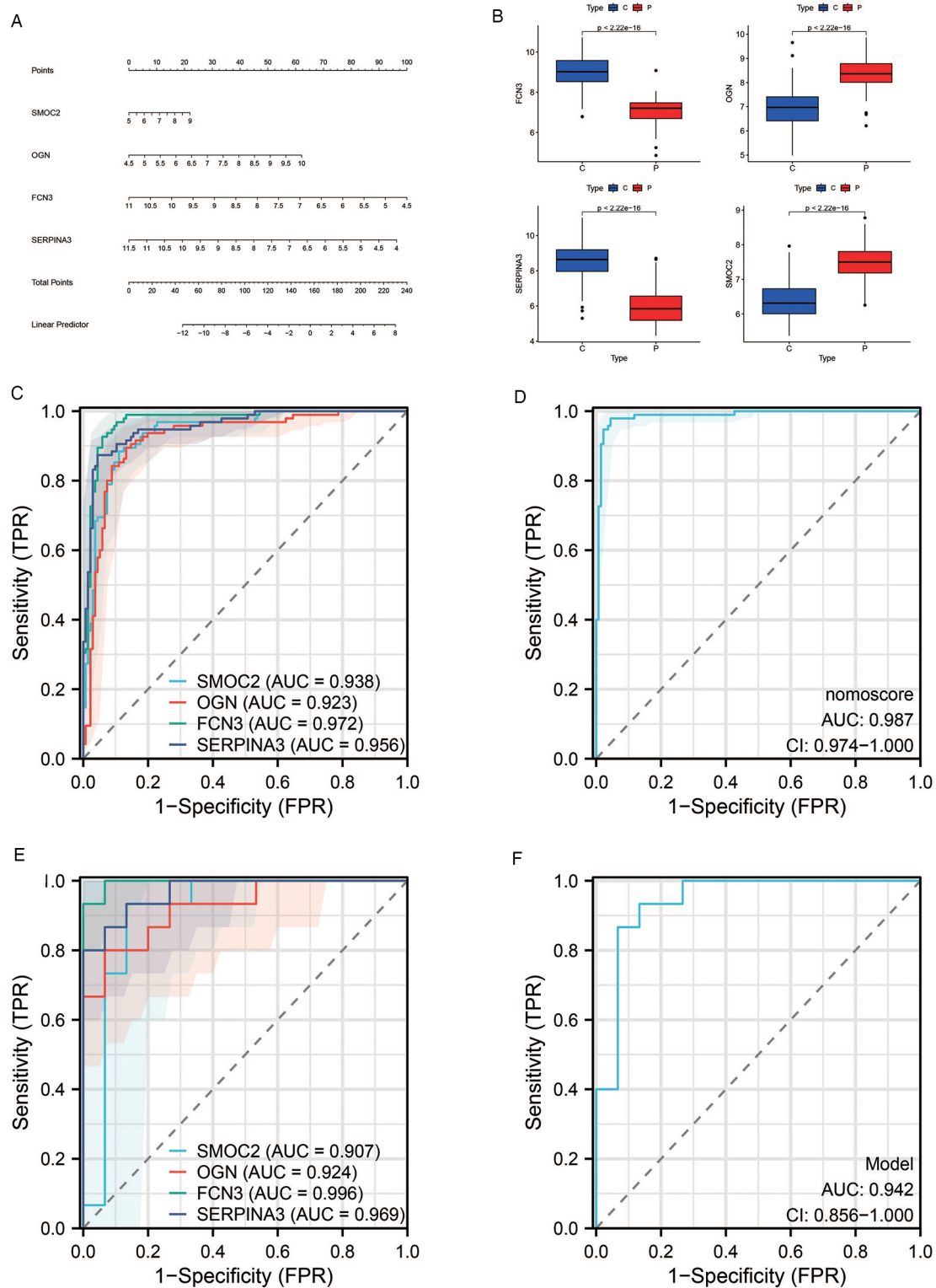
FIGURE 7
Results of diagnostic value assessment. (A) The illustrative nomogram for diagnosing pre-decompensation stage of HF. (B) Differential expression of hub genes in HF patients relative to normal controls in dataset GSE57338. (C) The ROC curve for individual candidate genes within the GSE57338 dataset. (D) The ROC curve for the nomogram highlighting their diagnostic value for pre-decompensation stage of HF. (E) The ROC curves for each candidate gene within the GSE16499 dataset. (F) The ROC curve representing the diagnostic efficacy of the 4-gene model in dataset GSE16499.
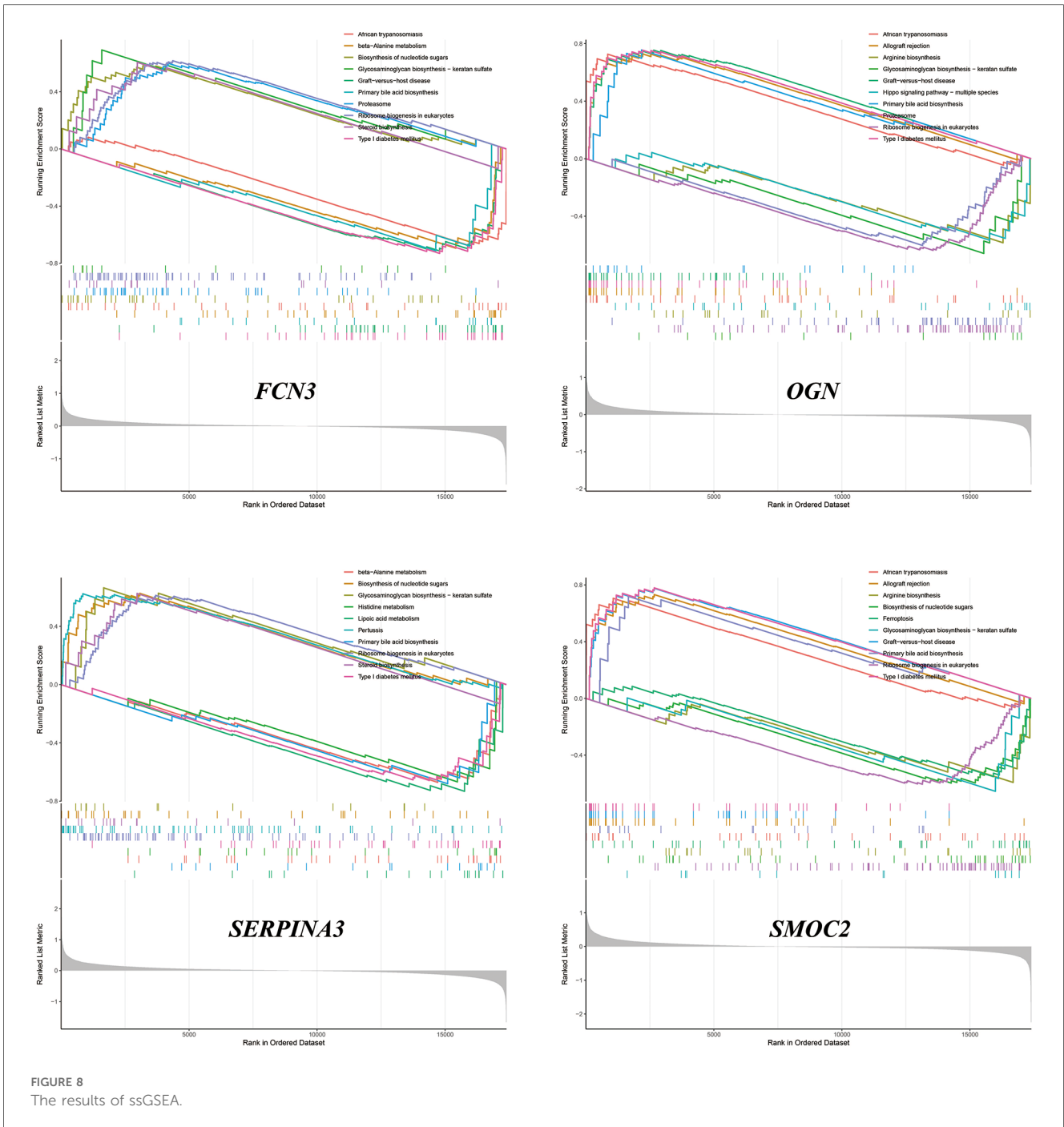
FIGURE 8
The results of ssGSEA.

impacting the heart chamber's stiffness. Variations in the quantity and quality of collagen can significantly affect myocardial stiffness. Persistent elevation in wall stress, exceeding the heart's adaptive capacity, may result in a disproportionate reduction in ventricular wall thickness relative to the ventricular volume, culminating in diastolic and systolic dysfunction (31). It should be noted that in this study, the VC data set we selected is a model of healthy people simulating venous congestion through upper limb compression. Compared with this model, the pathological process of heart failure involves pulmonary circulation congestion, and subsequent systemic circulation congestion and visceral congestion.

The pathological mechanism and consequences are more complex than simple upper limb venous congestion (32). There are significant differences in hemodynamics and neuroendocrine hormones between patients with heart failure and healthy individuals. Hemodynamically, in patients with heart failure, the heart's ability to pump blood is reduced, resulting in reduced cardiac output, which can cause insufficient blood supply to organs and tissues throughout the body. In heart failure, in order to maintain adequate cardiac output, the heart will increase heart rate, myocardial contractility, and blood vessels will constrict to increase blood pressure. These changes may help maintain blood

pressure and organ perfusion in the early stages, but in the long term they will increase heart failure. burden, leading to further damage to the heart and deterioration of function (33, 34). In terms of neuroendocrine hormones, the renin-angiotensin-aldosterone system (RAAS) and sympathetic nervous system are activated in patients with heart failure. Activation of these systems leads to vasoconstriction and water and sodium retention, which may help increase cardiac output in the short term, but long-term activation leads to increased cardiac load and myocardial remodeling, thereby exacerbating the deterioration of cardiac function and progression of heart failure. Serum levels of norepinephrine and aldosterone are often significantly increased in patients with heart failure and further increase as the disease progresses (35, 36). In addition, medications taken by patients with heart failure and the disease itself can have a dynamic impact on venous congestion. Nevertheless, this model also has certain advantages, because compared with the complex pathological mechanism of heart failure, this model only involves the impact of elevated venous pressure on endothelial cells, allowing us to better understand the local changes caused by venous congestion.

The emergence of venous congestion typically precedes the clinical signs of decompensation by weeks, often culminating in overt clinical deterioration that necessitates medical intervention. To address this, we employed a machine learning algorithm to identify four diagnostic genes associated with ADHF triggered by VC, namely SMOC2, OGN, FCN3, and SERPINA3. A nomogram was developed to enhance the practical utility of the diagnostic model. Potentially, clinicians might be able to detect these four indicators of heart failure patients and evaluate the stage of venous congestion through nomogram to adjust treatment plans in a timely manner, especially the use of diuretics. However, due to the lack of correction for indicators such as age, the accuracy of the nomogram still needs more clinical trials to verify and adjust. These four genes were subsequently validated against the external dataset GSE16499, where the AUC for each gene exceeded 0.9, and the composite genomic AUC was 0.942, underscoring their significant diagnostic potential. We completed the ssGSEA of these four genes, hoping to provide new ideas for our understanding of the molecular mechanism of ADHF after VC.

SPARC-related modular calcium-binding protein 2 (SMOC2), part of the SPARC family of matricellular proteins, is known to promote endothelial cell proliferation, migration, and angiogenesis (37). Research indicates that SMOC2 expression escalates in rats experiencing heart failure, and silencing SMOC2 can mitigate heart failure symptoms by modulating autophagy via the TGF-β1/Smad3 signaling pathway (38). Additionally, mouse studies have demonstrated that SMOC2 knockdown can ameliorate cardiac function impairment and cardiac fibrosis, potentially through the inhibition of the ILK/p38 signaling pathway (39). Furthermore, the overexpression of SMOC2 has been observed to enhance vascular smooth muscle cell proliferation, migration, and extracellular matrix breakdown, pointing to its significant role in cardiovascular pathophysiology (40).

Osteoglycin (OGN), a constituent of the small leucine-rich repeat proteoglycan (SLRP) family, plays a critical role in modulating inflammation and cardiac fibrosis (41). Research has revealed that OGN attenuation can suppress cardiac fibroblast proliferation and the epithelial/endothelial-mesenchymal transition, pivotal processes in cardiac remodeling. Moreover, OGN is known to facilitate apoptosis in cardiac fibroblasts, mediated through the regulation of the Wnt signaling pathway (42). Variations in OGN levels offer distinct insights into myocardial remodeling dynamics (43). Notably, an upsurge in OGN expression within infarct scars contributes to the proper maturation of collagen, thereby preventing cardiac rupture and counteracting adverse remodeling post-myocardial infarction (MI). Consequently, OGN holds promise as a valuable biomarker for ischemic heart failure (44).

Ficolins constitute a family of proteins characterized by their collagen-like and fibrinogen-like domains, playing a crucial role in triggering the complement lectin pathway (45). Ficolin-3 (FCN3), a principal molecule in this pathway, is pivotal for the activation of complement component 3 and has been linked to hypertension risk (46). Emerging evidence from studies on patients with acute congestive heart failure suggests an association with complement activation, positioning FCN3 could as a potential biomarker for this condition (47). However, the results of two clinical studies indicate that complement activation products have no prognostic value in acute heart failure (48, 49). Furthermore, the activation of the complement system and the consequent production of reactive oxygen species in the myocardium post-infarction are known to stimulate the upregulation of cytokines and chemokines. This cascade can precipitate adverse outcomes, such as left ventricular dilatation and contractility impairments (50).

Serpin Family A Member 3 (SERPINA3), also known as α-1-antichymotrypsin (AACT or ACT), is part of the serpin superfamily, which primarily functions to inhibit serine proteases. This protein's elevated levels are observed in conditions such as heart failure and various neurological disorders, including Alzheimer's disease and Creutzfeldt-Jakob disease (51). Research indicates that SERPINA3 acts as an endogenous gene responsive to the mineralocorticoid receptor and can be upregulated by aldosterone (52). Following VC, alterations in fluid retention and mineralocorticoid levels might disrupt SERPINA3 balance. Clinical evidence suggests that heightened SERPINA3 levels in individuals with new-onset or exacerbating heart failure correlate with increased mortality rates or unplanned cardiac readmissions, underscoring its prognostic value for heart failure (53). Furthermore, SERPINA3 has been proposed as a biomarker for assessing right ventricular myocardial function, particularly in the context of systemic congestion leading to right ventricular failure (54). Beyond its myocardial expression, elevated circulating levels of SERPINA3 have been linked to an increased cancer risk in heart failure patients, highlighting the necessity for more in-depth exploration of its clinical implications and molecular functions (55).

There is currently no gold standard for detecting congestion in established acute decompensated heart failure. This is usually assessed by signs and symptoms of heart failure, as well as

objective evidence of congestion (e.g., elevated BNP, pulmonary edema on chest x-ray, or pulmonary rales on auscultation) (56). There is also no uniform or readily available method for determining when effective dehydration occurs or when a patient has reached optimal volume status (57). Perhaps it would be an interesting direction to determine the congestion state of heart failure patients by detecting these protein markers, but clinical trials are still needed to further determine the sensitivity, effectiveness and specific thresholds.

The novel contributions of our investigation are manifold. Primarily, this study pioneers the bioinformatics exploration of the interplay between venous congestion and heart failure, dissecting the intricate mechanisms through which venous congestion may precipitate the transition from chronic to acute decompensated heart failure. Secondly, we have pinpointed four diagnostic genes —SMOC2, OGN, FCN3, and SERPINA3—associated with acute decompensated heart failure, employing three distinct machine learning techniques for their identification. Validation of these genes confirmed that their collective diagnostic model possesses substantial value, offering a new avenue to assess patient congestion status effectively. Furthermore, our analysis extends into the biological processes influenced by venous congestion in heart failure, providing fresh insights into its potential mechanisms and refining our understanding of the therapeutic implications of diuretics. In summary, our study provides a method to assess venous congestion in patients with heart failure who do not yet have physical signs of venous congestion (such as pulmonary rales). Theoretically, by monitoring these indicators, it can help clinicians to adjust the treatment plan of heart failure patients in a timely manner at an early stage to avoid the patient's transition from chronic stable heart failure to acute decompensated heart failure.

Despite its contributions, our study acknowledges certain limitations. Initially, the causal link between the observed elevation in mRNA levels and a corresponding increase in protein expression remains undetermined. Additionally, the possibility that endothelial activation post-venous congestion might trigger the secretion of other molecules influencing heart failure, such as exosomes, has not been fully explored. In addition, our study was based on the GEO dataset, and the sample size and selection bias of the GEO dataset may limit the applicability of our results. Of note, the venous congestion dataset used a venous congestion model based on healthy individuals (via arm compression) rather than actual HF patients, which may not accurately reflect the physiological status of HF patients. While we have developed a nomogram to facilitate the implementation of the diagnostic model, the precise values and their clinical relevance require validation through subsequent clinical trials.

Conclusion

We performed bioinformatics analysis of the GEO dataset to explore the potential molecular mechanisms by which venous congestion affects heart failure. Through three machine learning algorithms, LASSO, RF and SVM-RFE, we identified SMOC2, OGN, FCN3, and SERPINA3 as potential biomarkers for the evaluation of acute decompensated heart failure caused by venous congestion. More importantly, diagnostic models and nomogram tools based on these 4 genes were developed to help clinically assess the hyperemic status of patients with acute decompensated heart failure, which may become an interesting target for future in-depth research.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here https://www.ncbi.nlm.nih.gov/geo/, accession numbers GSE38783; GSE57338; GSE16499.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent from the patients/participants or patients/participants' legal guardian/next of kin was not required to participate in this study in accordance with the national legislation and the institutional requirements.

Author contributions

YY: Conceptualization, Formal Analysis, Investigation, Methodology, Resources, Validation, Visualization, Writing – original draft, Writing – review & editing. HY: Methodology, Validation, Writing – review & editing. QH: Investigation, Resources, Writing – review & editing. JS: Software, Supervision, Writing – review & editing. XL: Software, Supervision, Writing – review & editing. YX: Project administration, Writing – review & editing. YL: Funding acquisition, Project administration, Writing – review & editing.

Funding

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.

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.

References

1. Yancy CW, Jessup M, Bozkurt B, Butler J, Jr CD, Colvin MM, et al. 2017 ACC/AHA/HFSA focused update of the 2013 ACCF/AHA guideline for the management of heart failure: a report of the American College of Cardiology/American Heart Association task force on clinical practice guidelines and the heart failure society of America. *Circulation*. (2017) 136(6):e137–61. doi: 10.1161/CIR.0000000000000509

2. Cotter G, Metra M, Milo-Cotter O, Dittrich HC, Gheorghiade M. Fluid overload in acute heart failure–re-distribution and other mechanisms beyond fluid accumulation. *Eur J Heart Fail*. (2008) 10(2):165–9. doi: 10.1016/j.ejheart.2008.01.007

3. Felker GM, Cotter G. Unraveling the pathophysiology of acute heart failure: an inflammatory proposal. *Am Heart J*. (2006) 151(4):765–7. doi: 10.1016/j.ahj.2005.07.004

4. Mullens W, Damman K, Harjola VP, Mebazaa A, Brunner-La Rocca HP, Martens P, et al. The use of diuretics in heart failure with congestion—a position statement from the heart failure association of the European Society of Cardiology. *Eur J Heart Fail*. (2019) 21(2):137–55. doi: 10.1002/ejhf.1369

5. Yu CM, Wang L, Chau E, Chan RH, Kong SL, Tang MO, et al. Intrathoracic impedance monitoring in patients with heart failure. *Circulation*. (2005) 112(6):841–8. doi: 10.1161/CIRCULATIONAHA.104.492207

6. Zhou J, Li YS, Chien S. Shear stress-initiated signaling and its regulation of endothelial function. *Arterioscler Thromb Vasc Biol*. (2014) 34(10):2191–8. doi: 10.1161/ATVBAHA.114.303422

7. Kawai M, Naruse K, Komatsu S, Kobayashi S, Nagino M, Nimura Y, et al. Mechanical stress-dependent secretion of interleukin 6 by endothelial cells after portal vein embolization: clinical and experimental studies. *J Hepatol*. (2002) 37(2):240–6. doi: 10.1016/S0168-8278(02)00171-X

8. Wang BW, Chang H, Lin S, Kuan P, Shyu KG. Induction of matrix metalloproteinases-14 and -2 by cyclical mechanical stretch is mediated by tumor necrosis factor-alpha in cultured human umbilical vein endothelial cells. *Cardiovasc Res*. (2003) 59(2):460–9. doi: 10.1016/S0008-6363(03)00428-0

9. Colombo PC, Onat D, Harxhi A, Demmer RT, Hayashi Y, Jelic S, et al. Peripheral venous congestion causes inflammation, neurohormonal, and endothelial cell activation. *Eur Heart J*. (2014) 35(7):448–54. doi: 10.1093/eurheartj/eht456

10. Liu Y, Morley M, Brandimarto J, Hannenhalli S, Hu Y, Ashley EA, et al. RNA-Seq identifies novel myocardial gene expression signatures of heart failure. *Genomics*. (2015) 105(2):83–9. doi: 10.1016/j.ygeno.2014.12.002

11. Kong SW, Hu YW, Ho JW, Ikeda S, Polster S, John R, et al. Heart failure-associated changes in RNA splicing of sarcomere genes. *Circ Cardiovasc Genet*. (2010) 3(2):138–46. doi: 10.1161/CIRCGENETICS.109.904698

12. Smyth GK. Limma: linear models for microarray data. In: Gentleman R, Carey VJ, Huber W, Irizarry RA, Dudoit S, editors. *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*. New York, NY: Springer New York (2005). p. 397–420.

13. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods*. (2015) 12(5):453–7. doi: 10.1038/nmeth.3337

14. Szklarczyk D, Gable AL, Nastou KC, Lyon D, Kirsch R, Pyysalo S, 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(D1):D605–12. doi: 10.1093/nar/gkaa1074

15. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. (2003) 13(11):2498–504. doi: 10.1101/gr.1239303

16. Yu G, Wang LG, Han Y, He QY. Clusterprofiler: an R package for comparing biological themes among gene clusters. *OMICS*. (2012) 16(5):284–7. doi: 10.1089/omi.2011.0118

17. The Gene Ontology Consortium. The gene ontology resource: 20 years and still GOing strong. *Nucleic Acids Res*. (2019) 47(D1):D330–8. doi: 10.1093/nar/gky1055

18. Kanehisa M, Goto S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*. (2000) 28(1):27–30. doi: 10.1093/nar/28.1.27

19. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat Softw*. (2010) 33(1):1–22. doi: 10.18637/jss.v033.i01

20. Petralia F, Wang P, Yang J, Tu Z. Integrative random forest for gene regulatory network inference. *Bioinformatics*. (2015) 31(12):i197–205. doi: 10.1093/bioinformatics/btv268

21. Huang S, Cai N, Pacheco PP, Narrandes S, Wang Y, Xu W. Applications of support vector machine (SVM) learning in cancer genomics. *Cancer Genomics Proteomics*. (2018) 15(1):41–51. doi: 10.21873/cgp.20063

22. Gou M, Qian N, Zhang Y, Wei L, Fan Q, Wang Z, et al. Construction of a nomogram to predict the survival of metastatic gastric cancer patients that received immunotherapy. *Front Immunol*. (2022) 13:950868. doi: 10.3389/fimmu.2022.950868

23. Ye L, Zhang T, Kang Z, Guo G, Sun Y, Lin K, et al. Tumor-Infiltrating immune cells as a marker for prognosis in colorectal cancer. *Front Immunol*. (2019) 10:2368. doi: 10.3389/fimmu.2019.02368

24. Damman K, van Deursen VM, Navis G, Voors AA, van Veldhuisen DJ, Hillege HL. Increased central venous pressure is associated with impaired renal function and mortality in a broad spectrum of patients with cardiovascular disease. *J Am Coll Cardiol*. (2009) 53(7):582–8. doi: 10.1016/j.jacc.2008.08.080

25. Poole DC, Hirai DM, Copp SW, Musch TI. Muscle oxygen transport and utilization in heart failure: implications for exercise (in)tolerance. *Am J Physiol Heart Circ Physiol*. (2012) 302(5):H1050–63. doi: 10.1152/ajpheart.00943.2011

26. Thorup L, Simonsen U, Grimm D, Hedegaard ER. Ivabradine: current and future treatment of heart failure. *Basic Clin Pharmacol Toxicol*. (2017) 121(2):89–97. doi: 10.1111/bcpt.12784

27. Nantasenamat C, Prachayasittikul V, Bulow L. Molecular modeling of the human hemoglobin-haptoglobin complex sheds light on the protective mechanisms of haptoglobin. *PLoS One*. (2013) 8(4):e62996. doi: 10.1371/journal.pone.0062996

28. Matuschik L, Riabov V, Schmuttermaier C, Sevastyanova T, Weiss C, Klüter H, et al. Hyperglycemia induces inflammatory response of human macrophages to CD163-mediated scavenging of hemoglobin-haptoglobin complexes. *Int J Mol Sci*. (2022) 23(3):1385. doi: 10.3390/ijms23031385

29. Valiente-Alandi I, Schafer AE, Blaxall BC. Extracellular matrix-mediated cellular communication in the heart. *J Mol Cell Cardiol*. (2016) 91:228–37. doi: 10.1016/j.yjmcc.2016.01.011

30. Perestrelo AR, Silva AC, Oliver-De La Cruz J, Martino F, Horváth V, Caluori G, et al. Multiscale analysis of extracellular matrix remodeling in the failing heart. *Circ Res*. (2021) 128(1):24–38. doi: 10.1161/CIRCRESAHA.120.317685

31. Brower GL, Gardner JD, Forman MF, Murray DB, Voloshenyuk T, Levick SP, et al. The relationship between myocardial extracellular matrix remodeling and ventricular function. *Eur J Cardiothorac Surg*. (2006) 30(4):604–10. doi: 10.1016/j.ejcts.2006.07.006

32. Fudim M, Kaye DM, Borlaug BA, Shah SJ, Rich S, Kapur NK, et al. Venous tone and stressed blood volume in heart failure: JACC review topic of the week. *J Am Coll Cardiol*. (2022) 79(18):1858–69. doi: 10.1016/j.jacc.2022.02.050

33. Galati G, Germanova O, Iozzo RV, Buraschi S, Shchukin YV, Germanov A, et al. Hemodynamic arterial changes in heart failure: a proposed new paradigm of "heart and vessels failure". *Minerva Cardiol Angiol*. (2022) 70(3):310–20. doi: 10.23736/S2724-5683.21.05786-0

34. Miller WL. Fluid volume overload and congestion in heart failure: time to reconsider pathophysiology and how volume is assessed. *Circ Heart Fail*. (2016) 9(8):e002922. doi: 10.1161/CIRCHEARTFAILURE.115.002922

35. Hartupee J, Mann DL. Neurohormonal activation in heart failure with reduced ejection fraction. *Nat Rev Cardiol*. (2017) 14(1):30–8. doi: 10.1038/nrcardio.2016.163

36. Colombo PC, Doran AC, Onat D, Wong KY, Ahmad M, Sabbah HN, et al. Venous congestion, endothelial and neurohormonal activation in acute decompensated heart failure: cause or effect? *Curr Heart Fail Rep*. (2015) 12(3):215–22. doi: 10.1007/s11897-015-0254-8

37. Rocnik EF, Liu P, Sato K, Walsh K, Vaziri C. The novel SPARC family member SMOC-2 potentiates angiogenic growth factor activity. *J Biol Chem*. (2006) 281(32):22855–64. doi: 10.1074/jbc.M513463200

38. Ren Y, Wu Y, He W, Tian Y, Zhao X. SMOC2 plays a role in heart failure via regulating TGF-β1/Smad3 pathway-mediated autophagy. *Open Med (Wars)*. (2023) 18(1):20230752. doi: 10.1515/med-2023-0752

39. Rui H, Zhao F, Yuhua L, Hong J. Suppression of SMOC2 alleviates myocardial fibrosis via the ILK/p38 pathway. *Front Cardiovasc Med*. (2023) 9:951704. doi: 10.3389/fcvm.2022.951704

40. Wang X, Wang M, Zhou Z, Zou X, Song G, Zhang Q, et al. SMOC2 promoted vascular smooth muscle cell proliferation, migration, and extracellular matrix degradation by activating BMP/TGF-β1 signaling pathway. *J Clin Biochem Nutr*. (2023) 73(2):116–23. doi: 10.3164/jcbn.22-100

41. Rienks M, Papageorgiou A, Wouters K, Verhesen W, Leeuwen RV, Carai P, et al. A novel 72-kDa leukocyte-derived osteoglycin enhances the activation of toll-like receptor 4 and exacerbates cardiac inflammation during viral myocarditis. *Cell Mol Life Sci*. (2017) 74(8):1511–25. doi: 10.1007/s00018-016-2423-7

42. Fang Y, Chang Z, Xu Z, Hu J, Zhou H, Yu S, et al. Osteoglycin silencing exerts inhibitory effects on myocardial fibrosis and epithelial/endothelial-mesenchymal transformation in a mouse model of myocarditis. *BioFactors*. (2020) 46(6):1018–30. doi: 10.1002/biof.1683

43. Motiwala SR, Szymonifka J, Belcher A, Weiner RB, Baggish AL, Gaggin HK, et al. Measurement of novel biomarkers to predict chronic heart failure outcomes and left ventricular remodeling. *J Cardiovasc Transl Res*. (2014) 7(2):250–61. doi: 10.1007/s12265-013-9522-8

44. Van Aelst LN, Voss S, Carai P, Van Leeuwen R, Vanhoutte D, Sanders-van Wijk S, et al. Osteoglycin prevents cardiac dilatation and dysfunction after myocardial infarction through infarct collagen strengthening. *Circ Res*. (2015) 116(3):425–36. doi: 10.1161/CIRCRESAHA.116.304599

45. Garred P, Honoré C, Ma YJ, Munthe-Fog L, Hummelshøj T. MBL2, FCN1, FCN2 and FCN3-the genes behind the initiation of the lectin pathway of complement. *Mol Immunol*. (2009) 46(14):2737–44. doi: 10.1016/j.molimm.2009.05. 005

46. Lu J, Li M, Zhang R, Hu C, Wang C, Jiang F, et al. A common genetic variant of FCN3/CD164L2 is associated with essential hypertension in a Chinese population. *Clin Exp Hypertens*. (2012) 34(5):377–82. doi: 10.3109/10641963.2012.665538

47. Suffritti C, Tobaldini E, Schiavon R, Strada S, Maggioni L, Mehta S, et al. Complement and contact system activation in acute congestive heart failure patients. *Clin Exp Immunol*. (2017) 190(2):251–7. doi: 10.1111/cei.13011

48. Trendelenburg M, Stallone F, Pershyna K, Eisenhut T, Twerenbold R, Wildi K, et al. Complement activation products in acute heart failure: potential role in pathophysiology, responses to treatment and impacts on long-term survival. *Eur Heart J Acute Cardiovasc Care*. (2018) 7(4):348–57. doi: 10.1177/2048872617694674

49. Holt MF, Michelsen AE, Shahini N, Bjørkelund E, Bendz CH, Massey RJ, et al. The alternative complement pathway is activated without a corresponding terminal pathway activation in patients with heart failure. *Front Immunol*. (2021) 12:800978. doi: 10.3389/fimmu.2021.800978

50. Frangogiannis NG. The inflammatory response in myocardial injury, repair, and remodelling. *Nat Rev Cardiol*. (2014) 11(5):255–65. doi: 10.1038/nrcardio.2014.28

51. de Mezer M, Rogaliński J, Przewoźny S, Chojnicki M, Niepolski L, Sobieska M, et al. SERPINA3: stimulator or inhibitor of pathological changes. *Biomedicines*. (2023) 11(1):156. doi: 10.3390/biomedicines11010156

52. Gravez B, Tarjus A, Jimenez-Canino R, El Moghrabi S, Messaoudi S, Alvarez de la Rosa D, et al. The diuretic torasemide does not prevent aldosterone-mediated mineralocorticoid receptor activation in cardiomyocytes. *PLoS One*. (2013) 8(9): e73737. doi: 10.1371/journal.pone.0073737

53. Delrue L, Vanderheyden M, Beles M, Paolisso P, Di Gioia G, Dierckx R, et al. Circulating SERPINA3 improves prognostic stratification in patients with a *de novo* or worsened heart failure. *ESC Heart Fail*. (2021) 8(6):4780–90. doi: 10.1002/ehf2.13659

54. di Salvo TG, Yang KC, Brittain E, Absi T, Maltais S, Hemnes A. Right ventricular myocardial biomarkers in human heart failure. *J Card Fail*. (2015) 21(5):398–411. doi: 10.1016/j.cardfail.2015.02.005

55. Meijers WC, Maglione M, Bakker SJL, Oberhuber R, Kieneker LM, de Jong S, et al. Heart failure stimulates tumor growth by circulating factors. *Circulation*. (2018) 138(7):678–91. doi: 10.1161/CIRCULATIONAHA.117.030816

56. Núñez J, de la Espriella R, Rossignol P, Voors AA, Mullens W, Metra M, et al. Congestion in heart failure: a circulating biomarker-based perspective. A review from the biomarkers working group of the heart failure association, European society of cardiology. *Eur J Heart Fail*. (2022) 24(10):1751–66. doi: 10.1002/ejhf.2664

57. Lala A, Hamo CE, Bozkurt B, Fiuzat M, Blumer V, Bukhoff D, et al. Standardized definitions for evaluation of acute decompensated heart failure therapies. *JACC Heart Fail*. (2024) 12(1):1–15. doi: 10.1016/j.jchf.2023.09.030