



Subtypes identification on heart failure with preserved ejection fraction via network enhancement fusion using multi-omics data



Yongqing Wu^a, Huihui Wang^a, Zhi Li^b, Jinfang Cheng^c, Ruiling Fang^a, Hongyan Cao^{a,d,*}, Yuehua Cui^{e,*}

^a Division of Health Statistics, School of Public Health, Shanxi Medical University, Taiyuan, Shanxi 030001, PR China

^b Department of Hematology, Taiyuan Central Hospital of Shanxi Medical University, Taiyuan, Shanxi 030001, PR China

^c Department of Cardiology, Bethune Hospital, Shanxi Medical University, Taiyuan, Shanxi 030001, PR China

^d Shanxi Provincial Key Laboratory of Major Disease Risk Assessment, Taiyuan, Shanxi 030001, PR China

^e Department of Statistics and Probability, Michigan State University, East Lansing, MI 48824, USA

ARTICLE INFO

Article history:

Received 2 November 2020

Received in revised form 3 March 2021

Accepted 6 March 2021

Available online 17 March 2021

Keywords:

ne-SNF

HFpEF

Multi-omics data integration

Subtypes identification

Biomarkers

ABSTRACT

Heart failure with preserved ejection fraction (HFpEF) is associated with multiple etiologic and pathophysiologic factors. HFpEF leads to significant cardiovascular morbidity and mortality. There are various reasons that fail to identify effective therapeutic interventions for HFpEF, primarily due to its clinical heterogeneity causing significant difficulties in determining physiologic and prognostic implications for this syndrome. Thus, identifying clinical subtypes using multi-omics data has great implications for efficient treatment and prognosis of HFpEF patients. Here we proposed to integrate mRNA, DNA methylation and microRNA (miRNA) expression data of HFpEF with a similarity network fusion (SNF) method following a network enhancement (ne-SNF) denoising technique to form a fused network. A spectral clustering method was then used to obtain clusters of patient subtypes. Experiments on HFpEF datasets demonstrated that ne-SNF significantly outperforms single data subtype analysis and other integrated methods. The identified subgroups were shown to have statistically significant differences in survival. Two HFpEF subtypes were defined: a high-risk group (16.8%) and a low-risk group (83.2%). The 5-year mortality rates were 63.3% and 33.0% for the high- and low-risk group, respectively. After adjusting for the effects of clinical covariates, HFpEF patients in the high-risk group were 2.43 times more likely to die than the low-risk group. A total of 157 differentially expressed (DE) mRNAs, 2199 abnormal methylations and 121 DE miRNAs were identified between two subtypes. They were also enriched in many HFpEF-related biological processes or pathways. The ne-SNF method provides a novel pipeline for subtype identification in integrated analysis of multi-omics data.

© 2021 The Authors. Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Heart failure (HF) is the leading cause of hospitalization and death around the world. Approximately half of HF patients have a preserved left ventricular ejection fraction (HFpEF) and its prevalence is rising at an alarming rate of 1% per year relative to HF with reduced ejection fraction (HFrEF) [1,2]. HFpEF patients are highly heterogeneous and have poor prognosis. No effective treatment has been found to treat HFpEF patients [3]. Compared to HFrEF, HFpEF patients are more likely to be older female, hypertensive,

anemic, and having atrial fibrillation. The annual mortality rate for HFpEF patients is close to 8%; and the prevalence and 5-year mortality rate of HFpEF patients over the age of 70 are close to 50% [4]. With an aging population worldwide, the growing challenge of HFpEF requires urgent attention to speed up disease characterization and to identify subtypes of patients for more targeted clinical diagnosis and treatment of HFpEF patients.

To date, most approaches to identify the subtypes of HFpEF are mainly focused on phenomapping of HFpEF [5] using latent variable model and hierarchical clustering methods, based on phenotypic features such as demographic, physical and clinical characteristics, laboratory, electrocardiographic and echocardiographic data. Although the targeted therapy for the identified HFpEF phenogroups seems promising, the current therapy is limited to diuretic drugs and the treatment of co-morbidities [6]. The subtyping based on phenotypic features is still insufficient to

* Corresponding authors at: Department of Statistics and Probability, Michigan State University, East Lansing, MI 48824, USA (Y. Cui) & Division of Health Statistics, School of Public Health, Shanxi Medical University, Taiyuan, Shanxi 030001, PR China (H. Cao).

E-mail addresses: caohy@sxmu.edu.cn (H. Cao), cuiy@msu.edu (Y. Cui).

identify HFpEF subgroups with unique underlying pathophysiology and to assess the response to personalized therapy [7]. As the heritability of cardiovascular diseases (CVD) is at 40–50% in the general population [8] several molecular subtype studies were developed for CVD [9]. The subtypes based on common molecular features lead to novel classification of HFpEF, and shed light on mechanisms and plasma biomarkers involved in HFpEF [6]. With the radical advances in sequencing and computer technology, we are now able to extract high throughput multi-omics data without limit. Integrative genomic study is being increasingly emphasized [10]. It is promising that multi-omics data integration can capture associations and potential causal relationships between different data types [11] and further form a comprehensive understanding of the underlying biological processes [10].

Rappoport and Shamir [12] did a comprehensive review of multi-omic data integration methods in general. Specifically, the authors divided joint clustering of multiple data types into four categories: 1) Early integration, which simply combines multiple omics data matrices to form a single matrix and is the most simple one to apply. However, it has a few discernible limitations in practice as described in the paper; 2) Late integration, a method in which each omics data type is clustered separately using single-omic algorithm, then different clusters are integrated. One noticeable drawback is that one can lose signals that are weak in individual omic data type; 3) Methods based on statistical modeling (e.g. iCluster [13]), which assume a probabilistic distribution of the data and are powerful but they run slower and are sensitive to feature selections; and 4) Similarity-based methods (e.g. similarity network fusion (SNF) [14]). The similarity-based methods construct similarity matrices for each omic data type separately. These similarities are then integrated, with the advantage of incorporating diverse omics data types. SNF uses message-passing theory to propagate information through interactions between samples, thus effectively reduces noises and strengthens similarities present in one or more networks. However, it is acknowledged that in the fused network, the unavoidable noise features generated by the limitations of measurement technology and inherent natural variation, may dilute clustering signals and lead to some spurious associations between samples [15]. To alleviate this disadvantage, it is essential to further denoise the fused networks to achieve efficient and precise subtyping.

Recently, Wang et al. [16] proposed a network enhancement (NE) approach to denoise weighted biological networks. NE uses a doubly stochastic matrix operator to induce sparsity. Therefore, it can remove weak network edges and enhance true connectivity, which leads to better performance in downstream analysis. Given the promise of NE, in this work, we incorporated the NE strategy to denoise the SNF fused network (ne-SNF), thus strengthening strong similarities between patients and removing weak edges mainly caused by noises in the fused network. The proposed ne-SNF successfully improved the performance of subtype identification compared to SNF and other current state-of-the-art integration methods. Specifically, using multi-omics data collected by Framingham Heart Study (FHS), we aimed to discover molecular subtypes of HFpEF patients, and identify biomarkers and key pathways. Subsequent biological analysis of the key molecular features and pathways were shown to hold potential prognostic value and biological significance.

2. Materials and methods

2.1. Data

2.1.1. Ethics statement

The Framingham Heart Study (FHS) data set used for this work contain fully deidentified individuals obtained through dbGAP

(<http://dbgap.ncbi.nlm.nih.gov>). This is a secondary data analysis and the authors had no role in collecting the patient data. An IRB approval was awarded before accessing the data set.

2.1.2. Diagnosis of HFpEF patients in Framingham Heart study

The FHS data in this study included clinical data, survival data and multi-omics data, downloaded from dbGAP (study accession: phs000007, <http://dbgap.ncbi.nlm.nih.gov>). The cohort includes participants from Framingham, MA, to undergo biennial examinations to investigate CVD and its risk factors since 1948 [17]. Offspring (and their spouses) and adult grandchildren of the original cohort participants were recruited into the second- and third-generation cohorts in 1971 and 2002 [18].

According to the “2016 ESC guidelines for diagnosis and treatment of acute and chronic heart failure”, we selected HFpEF participants with the following criteria: (1) signs and symptoms of HF; (2) left ventricular ejection fraction >50%; (3) B-type natriuretic peptide >35 pg/ml and/or N-terminal-pro hormone B-type natriuretic peptide >125 pg/ml; and (4) showing related structural heart failure (left ventricular hypertrophy/left atrial enlargement) and/or diastolic dysfunction. Patients with valvular vitium of the heart and hypertrophic cardiomyopathy were removed. Out of a total of 175 HFpEF patients, 125 participants from the offspring cohort that took place between 1971 and 2011 (at examination 8) who had mRNA expression, DNA methylation and miRNA expression profiling were included in this study. Survival time was from the time of admission of HFpEF diagnosis to the time of last follow-up (2011) or time of death from cardiovascular disease. The outcome was death from cardiovascular disease or survival. The survival time ranged from 0 to 29.6 years.

2.1.3. Omics data preprocessing

Three types of omics data were obtained, including 17,874 mRNA expression probes, 443,207 methylation CpG sites, and 416 miRNAs. We converted 17,874 mRNA expression probes to 17,358 gene level expressions by taking the mean value of multiple probes as the gene level expression, following the annotation from the Affymetrix GeneChip Human Exon 1.0 ST platform. Using the annotation from the Illumina Human Methylation 450 BeadChip platform, we mapped 443,207 CpG sites to 27,604 genes. We then took the mean beta value of multiple CpG sites in a gene as the gene level methylation signal. For miRNA, we filtered out biological features which had >30% of missing values across patients, leaving 212 miRNA features. There was no missing data in mRNA expression. Missing values in both methylation and miRNA data were imputed with the K nearest neighbor (KNN) imputation method [19].

2.2. Statistical method

2.2.1. Similarity network fusion (SNF)

SNF [14] is a computational method to integrate multi-omics data by calculating and fusing the similarity networks of patients. To make the work self-contained, here we briefly introduce the SNF algorithm which is implemented through the following three steps:

Step 1, generates similarity networks from each omics data type separately. Suppose we have n samples and m data sources. A patient similarity network is represented as a graph $G = (V, E)$. The vertices V correspond to patients and the edges E represent the correlations between two nodes (i.e., patients). Edge weights are represented by an $n \times n$ similarity matrix W . $W(i, j)$ indicates the similarity between patients x_i and x_j and is computed as follows,

$$W(i, j) = \exp\left(-\frac{\rho^2(x_i, x_j)}{\mu \epsilon_{ij}}\right)$$

where $\rho(x_i, x_j)$ is the Euclidean distance between patients x_i and x_j ; μ is a hyperparameter that can be empirically set in which we set μ as 0.5; $\varepsilon_{i,j} = \frac{\text{mean}(\rho(x_i, N_i)) + \text{mean}(\rho(x_j, N_j)) + \rho(x_i, x_j)}{3}$, where $\text{mean}(\rho(x_i, N_i))$ is the averaged value of the distances between x_i and each of its neighbors.

To compute the fused similarity matrix from multiple data sources, Wang et al. [14] defined a full and sparse kernel P given as,

$$P(i, j) = \begin{cases} \frac{W(i, j)}{2 \sum_{k \neq i} W(i, k)}, & j \neq i \\ 1/2, & j = i \end{cases}$$

so that $\sum_j P(i, j) = 1$ holds.

Step 2, a local kernel matrix S using K nearest neighbors (KNN) is introduced to further reduce noise between patients, where S only retains the similarities between samples and their nearest neighbors, while the similarities between samples far away from each other are discarded (set as 0). i.e.,

$$S(i, j) = \begin{cases} \frac{W(i, j)}{\sum_{k \in N_i} W(i, k)}, & j \in N_i \\ 0, & \text{otherwise} \end{cases}$$

Step 3, similarity networks are iteratively updated as described below. Specifically, let

$$P^v = S^v \times \left(\frac{\sum_{k \neq v} P^{(k)}}{m-1} \right) \times (S^{(v)})^T$$

$$P_{\text{final}} = \frac{\sum_{v=1}^m P^v}{m}, \quad v = 1, 2, \dots, m$$

where P^v represents the similarity matrix derived from the v th data type; S^v represents the local affinity which contains the nearest neighbors' information; P_{final} represents the fused network; and m is the number of different data types. As a result, the similarity networks from different sources are more similar to realize the network fusion.

2.2.2. Denoising SNF with network enhancement (ne-SNF)

To improve the signal-to-noise ratio of the fused network P_{final} based on SNF, we incorporated the NE strategy [16] to denoise SNF, termed as ne-SNF. Given the fused similarity network P_{final} initially obtained with SNF, NE constructed a symmetric and doubly stochastic matrix (DSM) $T \in \mathbb{R}^{n \times n}$ with the following two steps:

$$S_{ne}(i, j) = \frac{P_{\text{final}}(i, j)}{\sum_{k \in N_i} P_{\text{final}}(i, k)} * I\{j \in N_i\}$$

$$T_{ij} = \sum_{k=1}^n \frac{S_{ne}(i, k) S_{ne}(j, k)}{\sum_{v=1}^n S_{ne}(v, k)}$$

where $I\{\hat{A}\cdot\}$ is an indicator function, N_i represents a set of x_i 's neighbors including x_i in $P_{\text{final}}(i, j)$, T is a symmetric DSM which encodes the local structures of the original network. The further diffusion process using T was defined as follows [16]:

$$\Phi_{t+1} = \alpha T \Phi_t T + (1 - \alpha) T$$

where t is the iteration step and α is a regularization parameter. The initial value Φ_0 can be set as the SNF fused similarity matrix P_{final} . Φ_t remains as a symmetric DSM at each iteration t and further converges to a non-trivial equilibrium symmetric DSM network.

NE defines a diffusion process that uses random walks of length three or less and a form of regularized information flow to denoise

the input network [16]. Following Wang et al. [16] after NE, 1) if the original eigenvalues are either 0 or 1, the process preserves these eigenvalues; 2) Let eigen-pair (λ_0, v_0) denote the eigen-pair of the initial DSM (T_0). Then the final converged graph has eigen-pair $(f_\alpha(\lambda_0), v_0)$, where $f_\alpha(\lambda) = \frac{(1-\alpha)\lambda}{1-\alpha\lambda^2}$. Because $\frac{(1-\alpha)\lambda_0}{1-\alpha\lambda_0^2} \leq \lambda_0$, the NE process always decreases the eigenvalues; and 3) While all eigenvalues are reduced, function f_α denoises the input by aggressively down-weighting smaller eigenvalues more than it does on larger eigenvalues. Thus, the constructed undirected network can effectively improve the similarity between related nodes, leading to better subtyping performance.

2.2.3. Spectral clustering

Given the ne-SNF denoised similarity network $\Phi_{\text{ne-SNF}}$, we used spectral clustering [20] to obtain network clusters. The spectral clustering can effectively capture the global structure of a graph [21] with the following processes:

First, each sample x_i is associated with a label indicator $y_i \in \{0, 1\}^c$, where $y_i(k) = 1$ if patient x_i belongs to the k th cluster, otherwise $y_i(k) = 0$. A partition matrix $Y = (y_1^T; y_2^T; \dots; y_n^T)$ is obtained to represent a clustering scheme. The spectral clustering based on similarity matrix $\Phi_{\text{ne-SNF}}$ aims to minimize the objective function as follows,

$$\min_{Q \in \mathbb{R}^{n \times c}} \text{Trace}(Q^T L^+ Q), \quad \text{s.t. } Q^T Q = I$$

where $Q = Y(Y^T Y)^{-\frac{1}{2}}$ and $L^+ = I - D^{-\frac{1}{2}} \Phi_{\text{ne-SNF}} D^{-\frac{1}{2}}$ where D is a diagonal matrix whose diagonal element is the sum of the row elements of $\Phi_{\text{ne-SNF}}$.

2.2.4. Evaluation of prognosis of HfPEF patients

We used survival analysis to evaluate the performance of the subtyping results and if the identified subgroups are clinically meaningful in terms of survival rate. The Kaplan-Meier survival curve represents the change in survival rate over time, and it is a monotone decreasing curve representing the cumulative survival rate [22]. The abscissa represents the survival time, and the ordinate represents the cumulative survival probability. When $t = 0$, the survival probability is 1, and the survival probability gradually decreases with the increase of time. The slope of decline indicates the speed of death rate, and the survival curve can intuitively measure the survival risk of different subgroups. In addition, the log-rank test was used to test the difference of survival curves.

Controlling for age, gender, smoking history, drinking, hypertension, hyperlipidemia and body weight index, we conducted Cox regression to explore the association between subtypes and HfPEF survival outcomes, and identified patients with high risk of prognosis. Meanwhile, the subtypes were compared in demographic and clinical features using Chi-squared test, two-sample t -test or Mann-Whitney U test.

We further explored the significantly differentially expressed (DE) mRNAs (DEmRNAs) and DE miRNAs (DEmiRNAs) between subtypes using the significance analysis of microarrays (SAM) tool [23] and used the R Limma package [24] to identify abnormal methylation genes. We presented the top 15 features of mRNA, DNA methylation and miRNA expression features that were most relevant to the final clustering results, and performed a series of biological functional analysis on these features using DAVID tools [25] including GO [26] enrichment analysis and KEGG [27] pathway analysis. miRTarBase (<http://mirtarbase.cuhk.edu.cn/php/search.php>) [28] was used to predict the target genes of DEmiRNAs.

3. Results

3.1. Characteristics of HFpEF in FHS

In this study, we used the data of all 125 HFpEF patients, who ranged from 51 to 88 years old, with an average age of 74.8 years. By the last follow-up, 57 patients survived and 68 died; the survival time ranged from 0 to 29.6 years. Their baseline characteristics were shown in Table 1.

3.2. Subtyping of HFpEF using ne-SNF

We carried out the ne-SNF method for the 125 patients with HFpEF using the mRNA expression (17,358 mRNAs), miRNA expression (212 sites) and DNA methylation (27,604 genes) data. Fig. 1 shows the flowchart of the ne-SNF analysis.

We compared the result of ne-SNF with the ones obtained by using single data type as well as two other subtyping strategies (unsupervised multiple kernel learning (UMKL) [29] and SNF) (see Fig. 2). As expected, networks built with individual data type yielded quite different patterns of patient similarity, while the fused data clustered tightly and quite similarly. After the denoising step with ne-SNF, the clustering result clearly shows three distinct sub-clusters.

We further reported p-values using the log-rank test to evaluate the significance of the difference in survival profiles between subgroups (Fig. 3). Analysis with single data type did not lead to significantly different survival profiles except for the result using miRNA expression, while the fused networks had significant differences in survival between subtypes in most cases. Although the smallest p-value is observed with 3 clusters when using the miRNA data alone, the heatmap does not show a clear pattern of three clusters (see Fig. 2). Among the three fusion methods (UMKL, SNF and ne-SNF), ne-SNF achieved the smallest p-value ($P = 0.0059$) with 3 clusters.

We further explored the association between prognosis of HFpEF and the identified 3 subgroups based on ne-SNF. A total of 75 patients (60%) in group 1 had a 5-year mortality rate of 34.5%, 29 patients (23.2%) in group 2 with a 5-year mortality rate of 29.6%, and 21 patients (16.8%) in group 3 with a 5-year mortality

Table 1
Baseline characteristics of the study population (N = 125).

Item	Classification	n (%) / mean ± SD
Age, year		74.81 ± 8.35
Gender	Female	48(38.4)
	Male	77(61.6)
Smoking history	Yes	18(14.4)
	No	107(85.6)
Drinking	Yes	63(50.4)
	No	62(49.6)
Comorbidities		
	Hypertension	Yes No
Hyperlipidemia	Yes	78(62.4)
	No	47(37.6)
Diabetes	Yes	33(26.4)
	No	92(73.6)
Chronic kidney disease	Yes	19(15.2)
	No	106(84.8)
Vital signs and laboratory data		
Systolic blood pressure, mmHg		132.86 ± 20.37
Diastolic blood pressure, mmHg		66.75 ± 11.35
Body weight index, kg/m ²		30.36 ± 5.52
Serum creatinine, mg/dl		1.25 ± 0.91
Heart rate, bpm		63.33 ± 11.77
BNP, median (IQR), pg/ml		1036.5(947.4)
LVEF, %		59.82 ± 12.61

rate of 63.3%. The difference of survival conditions between group 1 and group 2 was not significant ($\chi^2 = 0.499, P = 0.48$). Therefore, we combined group 1 and group 2 patients into one group which was defined as the low-risk group, while group 3 patients were defined as the high-risk group. Table 2 shows the 5-year mortality of the two groups. The high-risk group comprised of fewer patients (21 patients, 16.8%), but had the higher 5-year mortality rate (63.3%). Those in the low-risk group had 5-year mortality rate of 33%. The survival curve and the first three principal components (PCs) of the two subtypes were shown in Fig. 4. It can be seen from the survival curve that the two subtypes based on ne-SNF had significant differences in clinical prognosis (left figure). The survival probability of the high-risk group was significantly lower than that of the low-risk group ($P = 0.0017$). Besides, the 3D scatter plot (right figure) based on the top three PCs can distinguish the two subtypes (as shown in different colors).

3.3. Association of prognosis with identified molecular subtypes

Controlling for age, gender, smoking history, drinking, hypertension, hyperlipidemia and body mass index, we conducted Cox regression to explore the association between the two subtypes and HFpEF survival outcomes. The results were shown in Table 3. Patients in the high-risk group were 2.43 times higher in risk of death than the low-risk group. The only significant covariate is drinking ($P = 0.04$) and it is a protective factor.

Clinical characteristic varied significantly by low- and high-risk groups. As shown in Table 4, hypertension, hyperlipidemia and chronic kidney disease were reported more frequently in the high-risk group.

3.4. Functional annotation of important features between HFpEF subtypes

Focusing on the low- and high-risk groups, we conducted a DE analysis for each omics data type. The low-risk group was used as a control group. Based on an FDR q -value <0.05 threshold, a total of 157 DE mRNAs were identified, among which 95 were up-regulated and 62 were down-regulated; 121 DE miRNAs were identified, among which 10 were up-regulated and 111 were down-regulated. We also identified 2199 abnormal DNA methylation genes, among which 1627 were hypermethylated and 572 were hypomethylated following the Bonferroni adjusted criteria of $P_{adj} < 0.05$ and $|t| > 2$. A heatmap of different omics features between the low- and high-risk group is shown in Fig. 5.

To further demonstrate the biological significance of features associated with the two subtypes, we selected the top 15 significant features in mRNAs, methylations and miRNAs for further analysis. If these selected features are differentially expressed across the two subgroups, their biological implications could help us confirm that the subtypes are biologically meaningful in addition to their clinical implication. Firstly, we constructed a corresponding heatmap of the total 45 features between two subtypes (see Fig. 6). For the mRNA expression, in the low-risk group, *SLC25A24*, *MTERF1*, *SCLT1*, *USP1*, *ARHGAP18*, *SH3BGRL2*, *MTURN*, *NEIL3* showed low expression, while high-risk group exhibited completely opposite performance in the same gene set. In the methylation data, low-risk group showed hypermethylation in *IMP2*, *SPATA5L1*, while high-risk group showed hypermethylation in other 13 methylations. The heatmap of miRNAs showed decreased expression of *hsa-miR-233-3p*, *hsa-miR-126-5p*, *hsa-miR-454-3p*, *hsa-miR-590-5p*, *hsa-miR-186-5p-a1*, *hsa-miR-186-5p-a2*, *hsa-miR-19a-3p*, *hsa-miR-222-3p*, *hsa-miR-374b-5p*, *hsa-miR-144-5p*, *hsa-miR-16-5p*, *hsa-miR-106a-5p*, and *hsa-miR-17-5p*, which correlated with the decreased survival rate.

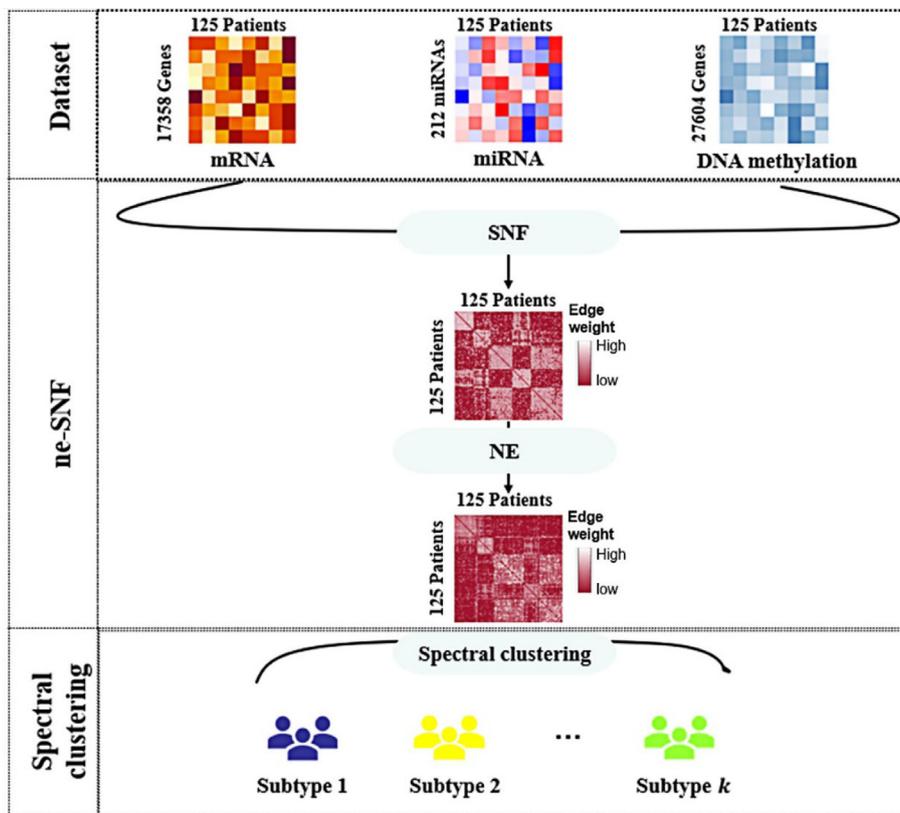


Fig. 1. Schematic representation of the pipeline for the proposed ne-SNF method.

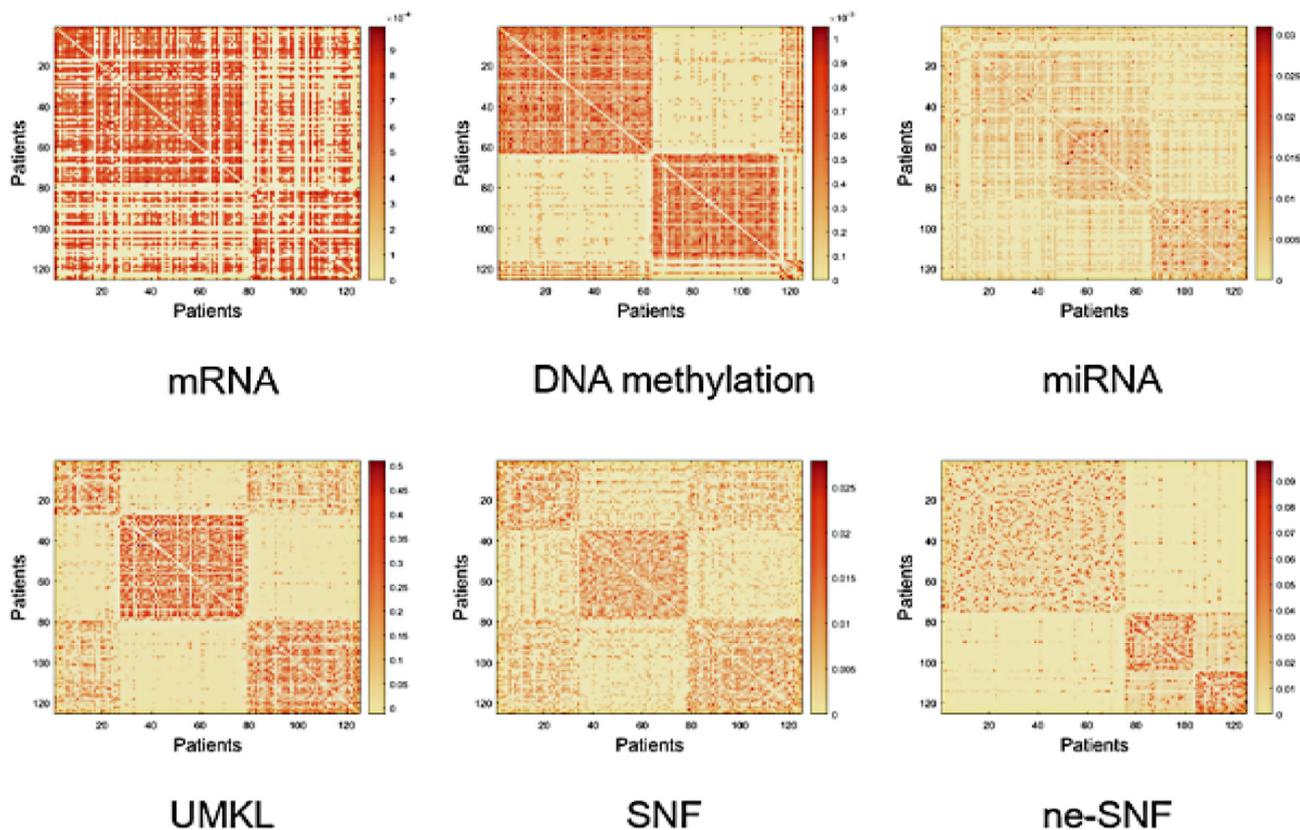


Fig. 2. Heatmaps of similarity matrix derived from single data type (mRNA, DNA methylation and miRNA) and using different methods (UMKL, SNF and ne-SNF). In each similarity matrix, the color shade represents the degree of patient-patient similarity: the darker the color, the higher the similarity between two individuals. Patients with high similarities define a subtype group, called a subgroup.

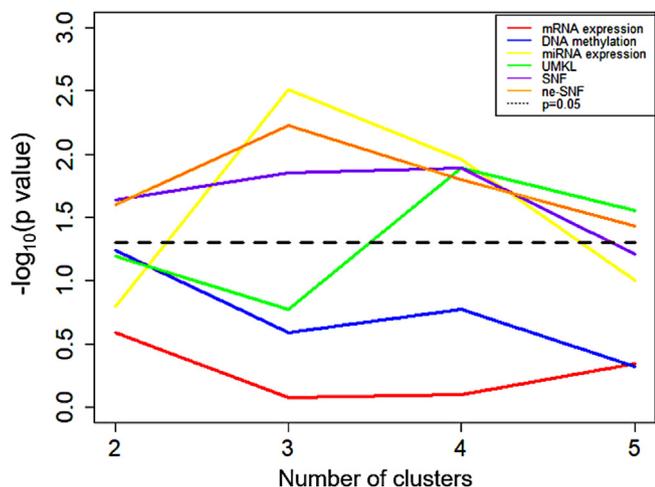


Fig. 3. Comparison of $-\log_{10}(p\text{-value})$ based on single data type, UMKL, SNF and ne-SNF under different number of clusters (2,3,4 and 5).

Second, focusing on each single feature, we performed survival analysis with a log-rank test, and found that >1/3 features showed good partition ability ($P < 0.05$), including three mRNAs (*EIF4A1*, *SH3BGRL2*, *MTERF1*), seven methylations (*ZNF689*, *IMPG2*, *SPATA5L1*, *LRR1Q1*, *ATP6V1G2*, *PYDC2*, *RAC3*) and six miRNAs (*hsa-miR-186-5p-a2*, *hsa-miR-186-5p-a1*, *hsa-miR-19a-3p*, *hsa-miR-17-5p*, *hsa-miR-106a-5p*, *hsa-miR-374b-5p*). Fig. 7 shows the Kaplan-Meier survival curves of the top 3 most significant features in mRNAs, methylations and miRNAs. It can be seen that these identified biomarkers had direct clinical prognostic value.

Table 2

The 5-year mortality of low-risk group and high-risk group.

Item	Low-risk group	High-risk group
Total, n (%)	104 (83.2)	21 (16.8)
5-year mortality (%)	33.0	63.3
χ^2		9.835
P-value		0.002

We further merged the identified mRNAs, methylation genes and targeted genes of miRNAs into a core set, to determine the functional relevance of the selected features, then performed the GO enrichment analysis and KEGG pathway analysis. The genes targeted by miRNAs were predicted using miRTarBase (an experimentally validated miRNA-target interaction database). The final core gene set contained 241 genes including 15 original mRNAs, 15 methylation genes and 211 miRNA targeted genes. The result showed that the core gene set was enriched in 17 GO biological processes, with the Bonferroni-adjusted p-value <0.05. Fig. 8 depicted all significant GO pathways. For KEGG pathway analysis, 45 pathways showed statistical significance (Bonferroni-adjusted p-value <0.05). Fig. 9 showed 17 KEGG pathways with the most significant enrichment.

4. Discussion

4.1. Classification of HFpEF based on ne-SNF

HFpEF is a highly heterogenous disease with no effective treatment so far. Identifying high-risk group with the aid of molecular data can provide meaningful clinical support for effective therapeutic treatment of HFpEF patients. Shah [30] did a comprehensive

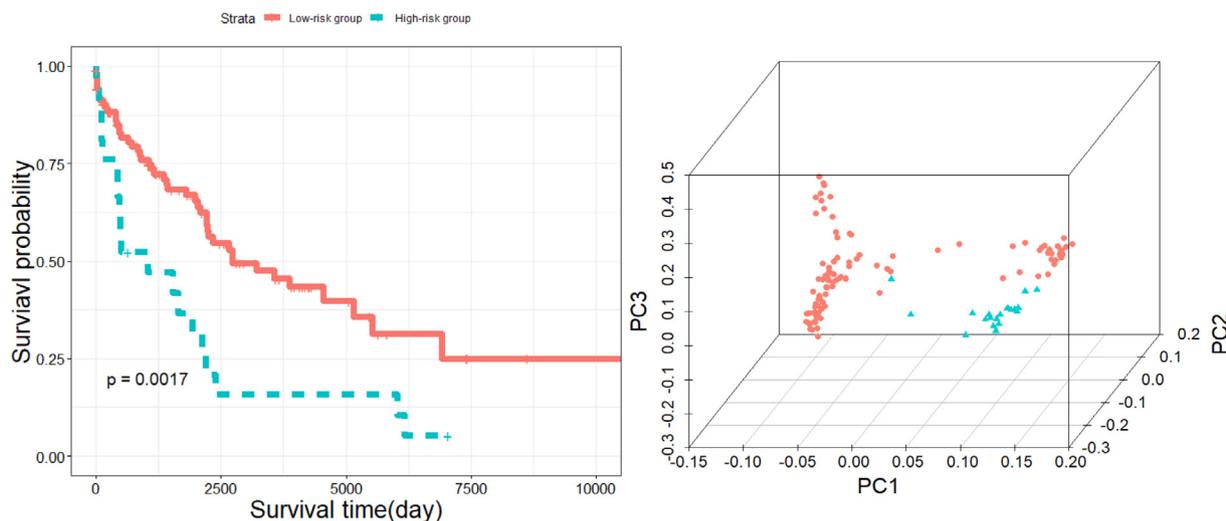


Fig. 4. Kaplan-Meier survival curves for the low- and high-risk groups (left), and 3D scatter plots of the first three PCs (right).

Table 3

List of Cox regression results of 125 HFpEF patients.

Item	Coefficient (SE)	Wald	P	HR	95%CI
Subtypes*	0.890 (0.282)	3.153	0.002	2.432	(1.400,4.227)
Age	0.021 (0.016)	1.313	0.189	1.021	(0.990,1.053)
Gender	-0.239 (0.280)	-0.856	0.392	0.787	(0.455,1.361)
Smoking history	0.308 (0.331)	0.933	0.351	1.361	(0.712,2.602)
Drinking*	-0.562 (0.286)	-1.966	0.040	0.570	(0.325,0.998)
Hypertension	0.189 (0.394)	0.481	0.630	1.209	(0.599,2.614)
Hyperlipidemia	-0.108 (0.331)	-0.326	0.745	0.898	(0.469,1.718)
Body mass index	-0.037 (0.025)	-1.489	0.136	0.964	(0.919,1.012)

*Showing statistical significance at the 0.05 significance level.

Table 4
Clinical and laboratory characteristics stratified by subtypes.

Characteristic	Low-risk group	High-risk group	χ^2/t	p-value
Age, years	75.08 ± 8.59	73.52 ± 7.37	0.773	0.441
Female, n (%)	42(40.4)	6(28.6)	1.031	0.310
Comorbidities, n (%)				
Hypertension*	72(68.9)	19(90.9)	4.422	0.035
Hyperlipidemia*	61(58.3)	17(81.8)	4.291	0.038
Diabetes	28(26.9)	5(23.8)	0.087	0.768
Chronic kidney disease*	12(11.7)	7(31.8)	4.263	0.039
Vital signs and laboratory data				
Systolic blood pressure, mmHg	133.90 ± 20.35	127.6 ± 20.65	1.278	0.204
Diastolic blood pressure, mmHg	66.59 ± 11.71	67.57 ± 9.92	-0.360	0.720
Body weight index, kg/m ²	30.43 ± 5.49	30.05 ± 5.91	0.282	0.778
Serum creatinine, mg/dl	1.26 ± 0.98	1.19 ± 0.41	0.313	0.755
Heart rate, bpm	63.68 ± 11.78	61.57 ± 12.13	0.746	0.457
BNP, median (IQR), pg/ml	1003.5(979.9)	1074.99(965.3)	-0.746	0.456
LVEF, %	59.79 ± 12.62	59.96 ± 12.87	-0.058	0.954

Categorical variables are presented as counts and percentages, continuous variables are presented as mean ± SD.

*Showing statistical significance at the 0.05 significance level.

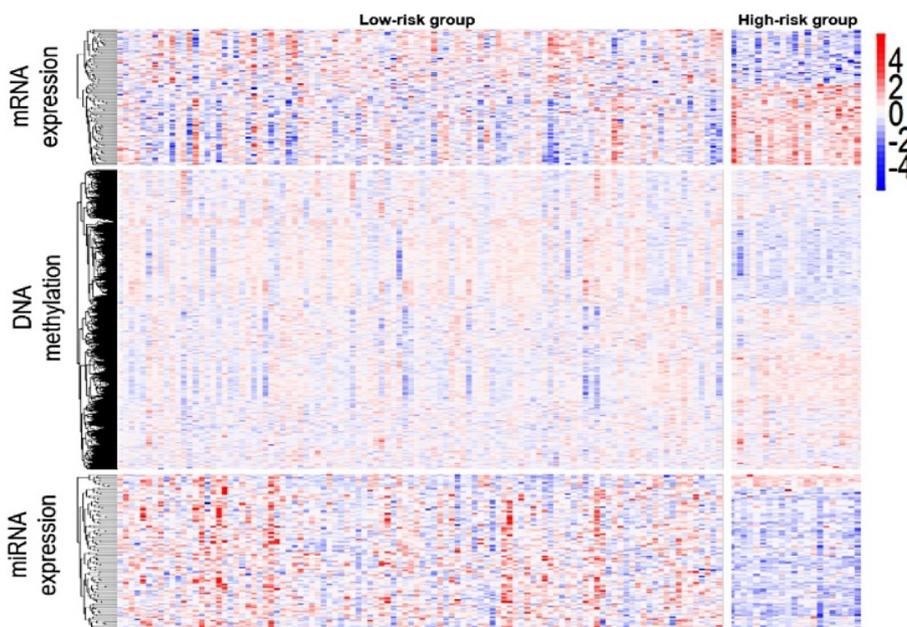


Fig. 5. The heatmap of DEMRNAs, abnormal methylations and DEMiRNAs between the low- and high-risk group. Each row represents an individual feature and each column represents a patient. Red and blue color represents relatively high and low expression respectively, with the intensity of the color representing the magnitude of high/low expression. The heatmap indicates that the high- and low-risk groups of HFpEF patients are highly heterogeneous among the three data types. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

review on precision medicine for HFpEF. While the reviewed works mostly focus on clinical data from a deep phenotyping perspective, there has been a large shift on precision medicine to omics-based data driven studies. In this work, we conducted subtype identification of HFpEF patients using multi-omics data. We proposed an ne-SNF method to improve SNF for subtype identification. HFpEF patients were further classified into two groups (low-risk and high-risk group) which had dramatic difference in survival rate. The high-risk group consisted of a smaller proportion of the study population (16.8%); however, it had a higher 5-year mortality rate at 63.3% compared to the low-risk group with a 5-year mortality rate of 33.0%. Furthermore, after adjusting for the covariates' effects, HFpEF patients in the high-risk group were 2.43 times higher in death risk compared to the low-risk group. The high-risk population and molecular markers identified in this analysis can potentially help for the development of novel targeted therapies.

Our proposed ne-SNF method took the advantages of both SNF and network denoising with the network enhancing technique [16]. Given the SNF fused similarity network, we used NE as a pre-processing step for spectral clustering, thus effectively denoising the fused similarity network, and further improving the clustering performance. The ne-SNF method generated meaningful disease groups compared to the ones obtained with single data type and other integrated methods for HFpEF subtyping; thus, provides a novel pipeline for disease subtype identification in integrated analysis of multi-omics data.

4.2. HFpEF was associated with chronic comorbidities

The subtypes obtained with the ne-SNF demonstrated both biological and clinical relevance for HFpEF patients. The identified high-risk group had high 5-year mortality, and was commonly associated with hypertension, hyperlipidemia and chronic kidney

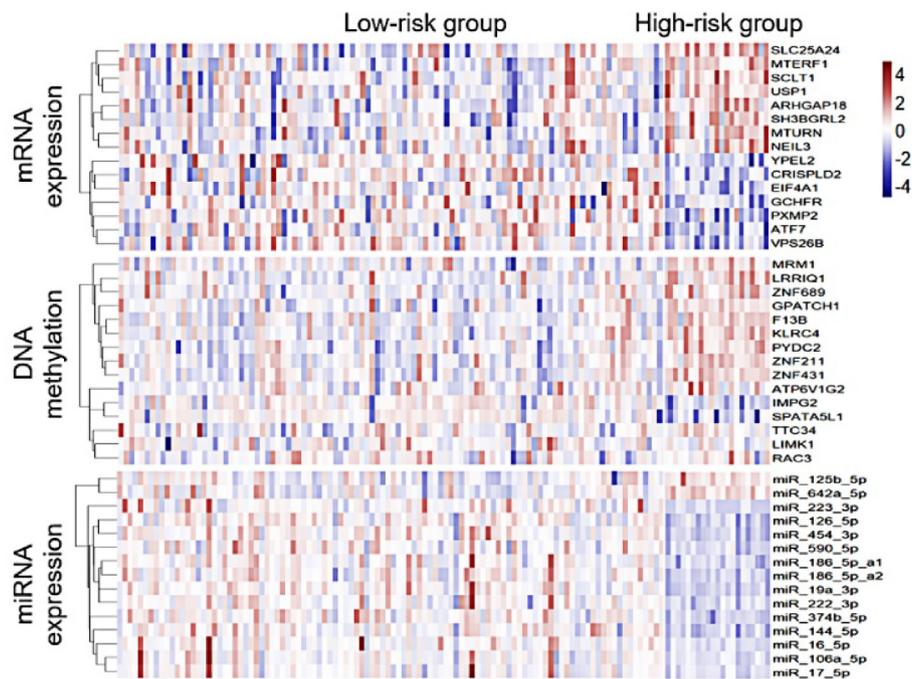


Fig. 6. The heatmap of the selected top15 mRNAs, methylation genes and miRNAs that are associated with the two risk subgroups.

disease. This result illustrated that HFpEF along with chronic diseases might be the main cause of patients' death. Individuals with hypertension were at high risk of HFpEF, and multimorbidity is common in HFpEF [31]. HFpEF patients had a high rate of adverse outcome with a high rate of hyperlipidemia and chronic kidney disease [5]. This implies that one should make more effort in treating HFpEF patients with chronic diseases.

4.3. Role of molecular biomarkers in the pathophysiology of HFpEF

Our study discovered 157 DE mRNAs, 121 DE miRNAs and 2199 abnormal methylations for HFpEF. Among DE mRNAs, 95 were up-regulated and 62 were down-regulated. Among DE miRNAs, 10 were up-regulated and 111 were down-regulated. Among 2199 abnormal methylations, 1627 were hypermethylated and 572 were hypomethylated.

In DE mRNAs, heart failure patients had elevated myocardia *NEIL3* expression. *NEIL3* regulated LDL-C and HDL-C levels and was associated with traditional cardiovascular risk factors [32,33]. The expression level of *MMP13* was significantly increased while the expression of *VEGF* was significantly decreased in the high-risk group. Significant changes in gene expression in the two genes associated with HF were recently reported in a study using Dahl salt sensitive rat [34]. In addition, gene *LUM* coding for lumican was up-regulated in the high-risk group. Lumican is an extracellular matrix localized proteoglycan associated with inflammatory conditions known to bind collagen [35].

For the identified DE methylated genes, some studies have shown that DNA methylation can regulate the expression of genes related to cardiovascular disease and further affect the occurrence and development of cardiovascular disease. In our study, we found that *MYBPC3* had higher methylation levels in the high-risk group, which caused gene *MYBPC3* silencing; and *MYBPC3* ablation caused defective diastolic relaxation and may affect an individual's susceptibility to develop HFpEF [36].

miRNAs have been shown to be attractive therapeutic targets. Inhibiting an miRNA or stimulating its activity can potentially influence many gene expressions. This could lead to significant therapeutic effects compared to standard drug treatments [37]. In our analysis, *hsa-miR-106a-5p*, *hsa-miR-193b-3p*, *hsa-miR-193b-5p*, *hsa-miR-191-5p* and *hsa-miR-660-5p* showed significantly differential expression between the low- and high-risk group. Existing literature revealed that these miRNAs were associated with several mechanisms of potential HF [38,39].

In addition, some HFpEF-related genes identified in this study were found to be tumor-related. For example, gene *ARHGAP18* was shown to be overexpressed in highly migratory triple-negative breast cancer cells; gene *CLIC4* was shown to be a diagnostic biomarker for different types of epithelial ovarian cancer [40,41]. Low-grade inflammation and increased oxidative stress were the main cause of cardiovascular disease, including coronary artery disease and heart failure [42]. While inflammatory responses accompanied by oxidative stress may cause cancer [43], the cross phenomenon between HF and cancer may share common pathological mechanism, which needs to be further studied.

Some of the DE genes were enriched in 17 GO biological process terms. These terms have been shown to be correlated with HF. For example, intrinsic apoptotic signaling pathway responds to DNA damage and shows positive regulation of apoptotic process and replicative senescence participated in the full life-cycle of HF [44,45]. In addition, transmembrane receptor protein tyrosine kinase is involved in myocardial damage in HF [46]. For KEGG analysis, a total of 45 pathways were identified. In particular, MAPK signaling pathway showed statistical significance. MAPK signaling pathway induces proliferation response and apoptotic response which contribute to the augmented activity in HF, consequently promoting inflammation and renin-angiotensin system activity in regions with key cardiovascular regulations [47]. HIF-1 signaling pathway was shown to participate in the process of myocardial fibrosis, thereby impairing cardiac diastolic function; Ras signaling pathway played a role in the pathogenesis of diastolic dysfunction

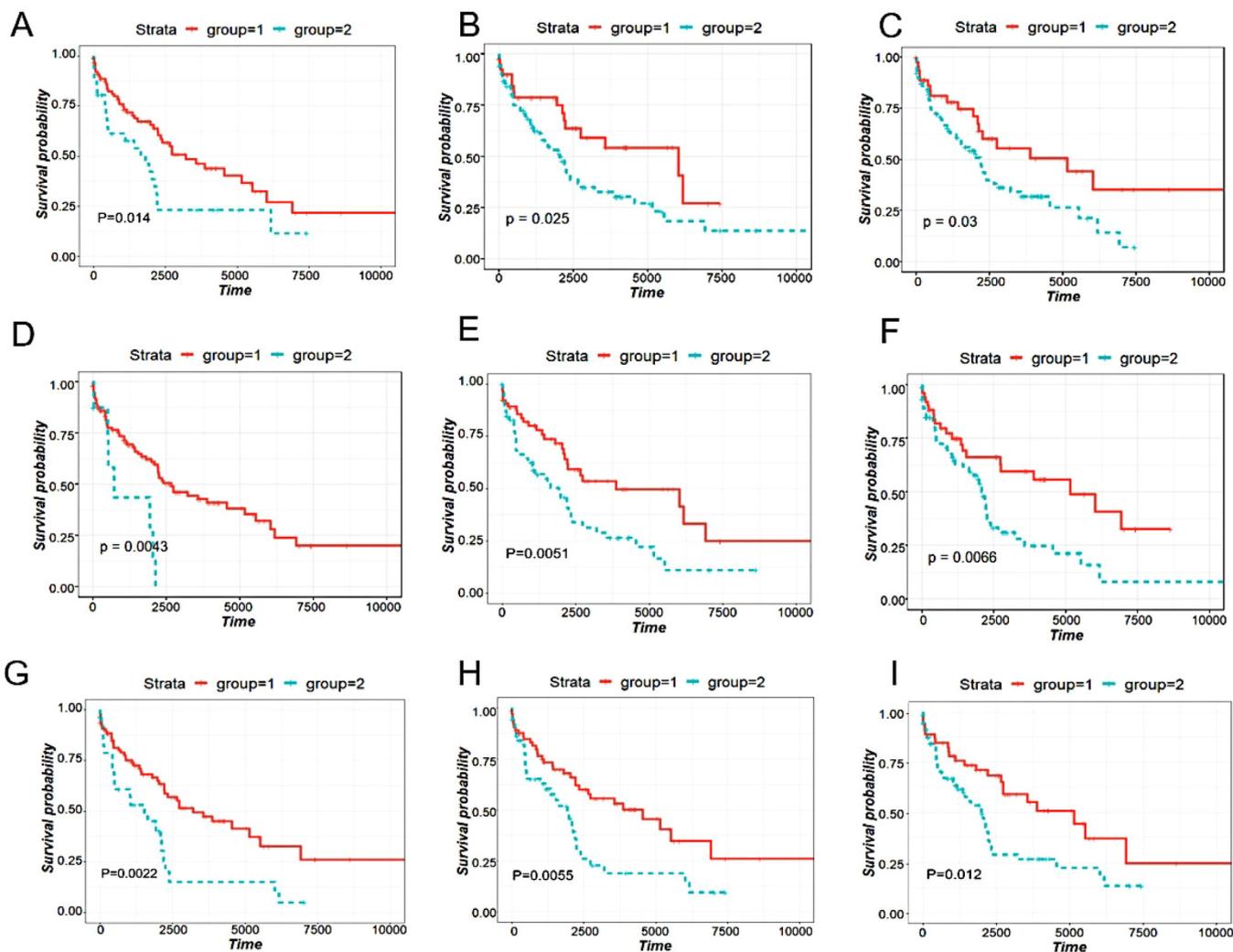


Fig. 7. Plots of survival curves of the top 3 most significant features in mRNAs, methylations and miRNAs: the top 3 most significant mRNAs, namely, *EIF4A1*, *SH3BGL2* and *MTERF1* (A-C), the top 3 most significant methylations, namely, *IMP2*, *PYDC2* and *ATP6V1G2* (D-F), and the top 3 most significant miRNAs, namely, *hsa-miR-19a-3p*, *hsa-miR-186-5p-a1*, *hsa-miR-186-5p-a2* (G-I).

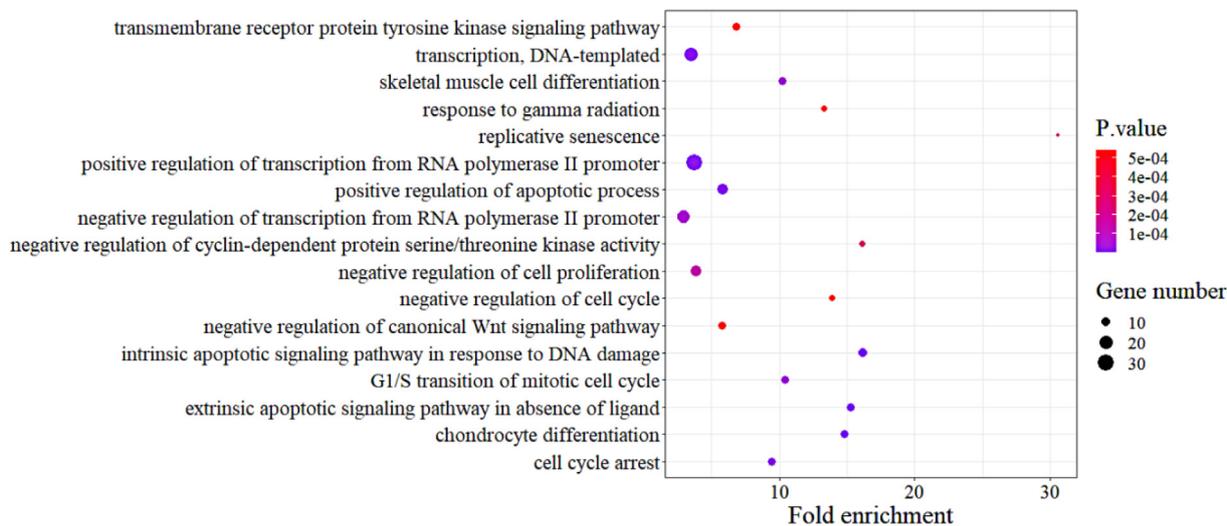


Fig. 8. GO biological process enrichment analysis of 241 core genes.

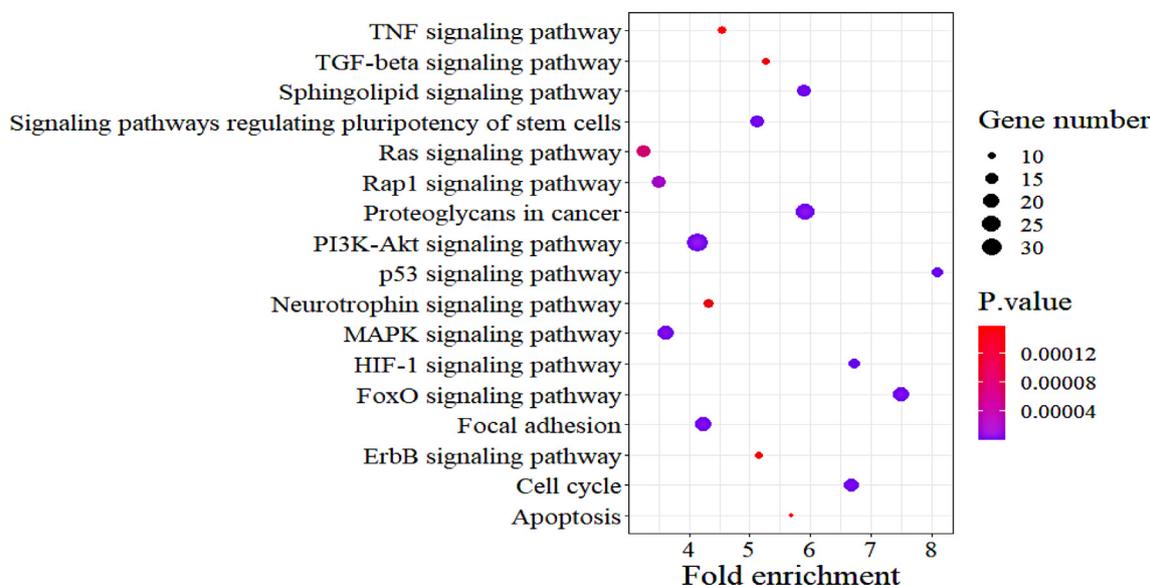


Fig. 9. KEGG enrichment analysis of 241 core genes.

in HFpEF [48,49]. In addition, Zhang et al. [50] revealed that genetic risk factors are involved in p53 signaling pathway of HFpEF based on gene set enrichment analysis of multi-omics data.

4.4. Can molecular biomarkers help direct future diagnosis and therapy for HFpEF?

Many studies used clinical data to phenotype HFpEF subgroups by machine learning and cluster-mapping approaches, and found that HFpEF patients at high-risk have hemodynamic overload, higher NT-proBNP, and renal dysfunction [51–53]. One of the limiting factors is that patients with HFpEF often have many overlapping characteristics, making it difficult to assign them to a pre-designated clinical subgroup. Therefore, we took a new approach that using multi-omics data to develop sub-clusters, then mapping the clusters back to clinical features and biological pathways. The recent emergence of genome-editing technologies has enabled a new paradigm in which the human genome sequence can be modified to achieve therapeutic effects. This includes adding therapeutic genes at specific sites in the genome and removing deleterious genes or genome sequences [54]. The identified biomarkers in this work will provide a basis for gene therapy. On the other hand, the regulatory pathways of HFpEF found in this study can be used as drug targets in clinical practice, so as to provide new ways for the treatment of HFpEF. For example, Nintedanib, a tyrosine kinase inhibitor, has a wide range of targets, including vascular endothelial growth factor receptor, PDGF, and fibroblast growth factor receptor [55]. Its effect on fibroblasts prevents ECM production and myofibroblast activation [56,57]. However, little study has been done on HFpEF. Alternative angiotensin inhibitors appeared safe in phase I studies, and this could represent a promising future additive therapy to target the angiotensin pathway [58]. It should be noted that the molecular features identified in this work need to be biologically validated before further clinical application.

Evidences from multiple clinical studies showed that BNP/NT-proBNP are the most reliable indicators for the diagnosis and treatment of HFpEF [59,60]. However, the role of clinical biomarkers in the diagnosis of HFpEF is currently limited. For example, HF biomarker-NT-proBNP was increased moderately in HFpEF and showed poor predictive value for HFpEF [39]. A study showed that biomarker panels of microRNAs demonstrated higher discriminative power for distinguishing HFpEF compared to the

result using plasma NT-proBNP levels alone [38]. Additionally, studies of HFpEF from the 1980s to the present have shifted from the perspective of mechanistic understanding to multi-organ focused disease study [61] including pulmonary arterial hypertension, obesity and metabolic diseases, vascular stiffness, chronological impotence, skeletal muscle disease, and renal insufficiency [62]. This has been blamed for the clinical failure of treatment trials [51]. At present, HFpEF therapy continues to improve symptoms, as well as focusing on treating comorbidities. The identification of novel molecular biomarkers may contribute to the understanding of the pathophysiology of HFpEF, help stratify HFpEF subgroups, and open a new arena for individualized therapy.

Like many other studies, there are some limitations in this work. First, we were unable to find a second HFpEF multi-omics data set to replicate the biological findings. Thus, the reproducible issue cannot be assessed in this analysis. Second, our results should be interpreted with caution. We only identified biomarkers associated with HFpEF, but could not establish causality. Further biological validations are needed to functionally validate the findings. Nevertheless, our proposed network enhanced similarity network fusion strategy (ne-SNF) for multi-omics data integration shows promising result in denoising the fused network and improving the performance of subtype identification. Given the heterogeneous nature of HFpEF, the identified subtypes and biomarkers will help for better design of clinical treatment and further lead to development of targeted molecular therapy.

CRediT authorship contribution statement

Yongqing Wu: Formal analysis, Methodology, Software, Writing - original draft. **Huihui Wang:** Formal analysis. **Zhi Li:** Visualization. **Jinfang Cheng:** Conceptualization. **Ruiling Fang:** Visualization. **Hongyan Cao:** Conceptualization, Supervision, Writing - review & editing. **Yuehua Cui:** Conceptualization, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We are grateful to the four anonymous reviewers for their insightful comments that helped us greatly improve our manuscript. The Framingham Heart Study was conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with Boston University (Contract No. N01-HC-25195, HHSN2682015000011, and 75N92019D00031). This manuscript was not prepared in collaboration with the investigators of the Framingham Heart Study and does not necessarily reflect the opinions or views of the Framingham Heart Study, Boston University, or the NHLBI.

References

- Shah SJ, Kitzman DW, Borlaug BA, et al. Phenotype-specific treatment of heart failure with preserved ejection fraction: a multiorgan roadmap. *Circulation* 2016;134(1):73–90.
- Redfield MM. Heart failure with preserved ejection fraction. *N Engl J Med* 2016;375(19):1868–77.
- Shah SJ, Katz DH, Deo RC. Phenotypic spectrum of heart failure with preserved ejection fraction. *Heart Fail Clin* 2014;10(3):407–18.
- Zile MR, Brutsaert DL. New concepts in diastolic dysfunction and diastolic heart failure: part I—diagnosis, prognosis, and measurements of diastolic function. *Circulation* 2002;105(11):1387–93.
- Kao DP, Lewsey JD, Anand IS, et al. Characterization of subgroups of heart failure patients with preserved ejection fraction with possible implications for prognosis and treatment response. *Eur J Heart Fail* 2015;17(9):925–35.
- Rech M, Barandiarán Aizpurua A, van Empel V, et al. Pathophysiological understanding of HFpEF: microRNAs as part of the puzzle. *Cardiovasc Res* 2018;114(6):782–93.
- Magnussen C, Blankenberg S, et al. Biomarkers for heart failure: small molecules with high clinical relevance. *J Intern Med* 2018;283(6):530–43.
- Won HH, Natarajan P, Dobbyn A, et al. Disproportionate contributions of select genomic compartments and cell types to genetic risk for coronary artery disease. *PLoS Genet* 2015;11(10):e1005622.
- Leon-Mimila P, Wang J, Huertas-Vazquez A, et al. Relevance of multi-omics studies in cardiovascular diseases. *Front Cardiovasc Med* 2019;6–91.
- Zarayneh N, Ko E, Oh JH, et al. Integration of multi-omics data for integrative gene regulatory network inference. *Int J Data Min Bioinform* 2017;18(3):223–39.
- Lock EF, Hoadley KA, Marron JS, et al. Joint and individual variation explained (JIVE) for integrated analysis of multiple data types. *Ann Appl Stat* 2013;7(1):523–42.
- Rappoport N, Shamir R. Multi-omics and multi-view clustering algorithms: review and cancer benchmark. *Nucleic Acids Res* 2018;46(20):10546–62.
- Shen R, Olshen AB, Ladanyi M. Integrative clustering of multiple genomic data types using a joint latent variable model with application to breast and lung cancer subtype analysis. *Bioinformatics* 2009;25(22):2906–12.
- Wang B, Mezlini AM, Demir F, et al. Similarity network fusion for aggregating data types on a genomic scale. *Nat Methods* 2014;11(3):333–7.
- Ji J, Zhang A, Liu C, et al. Survey: functional module detection from protein-protein interaction networks. *IEEE Trans Knowl Data Eng* 2014;2:261–77.
- Wang B, Pourshafeie A, Zitnik M, et al. Network enhancement as a general method to denoise weighted biological networks. *Nat Commun* 2018;9(1):3108.
- Oppenheimer GM. Becoming the Framingham Study 1947–1950. *Am J Public Health* 2005;95(4):602–10.
- Joehanes R, Ying S, Huan T, et al. Gene expression signatures of coronary heart disease. *Arterioscler Thromb Vasc Biol* 2013;33(6):1418–26.
- Troyanskaya O, Cantor M, Sherlock G, et al. Missing value estimation methods for DNA microarrays. *Bioinformatics* 2001;17(6):520–5.
- Ng AY, Jordan MI, Weiss Y. “On spectral clustering: analysis and an algorithm,” in International Conference on Neural Information Processing Systems. Natural and Synthetic. (Vancouver, BC): 849–856.
- Luxburg UV. A tutorial on spectral clustering. *Stat Comput* 2007;17:395–416.
- Goel MK, Khanna P, Kishore J. Understanding survival analysis: Kaplan-Meier estimate. *Int J Ayurveda Res* 2010;1(4):274–8.
- Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A* 2001;98(9):5116–21.
- Smyth GK, Yifang Hu, Ritchie M, et al. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* 2004;3:Article3.
- Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 2009;4(1):44–57.
- Ashburner M, Ball CA, Blake JA, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* 2000;25(1):25–9.
- Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M. KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res* 2012;40:D109–14.
- Chou CH, Shrestha S, Yang CD, et al. miRTarBase update 2018: a resource for experimentally validated microRNA-target interactions. *Nucleic Acids Res* 2018;46(D1):D296–302.
- Mariette J, Villavialaneix N. Unsupervised multiple kernel learning for heterogeneous data integration. *Bioinformatics* 2018;34:1009–15.
- Shah SJ. Precision medicine for heart failure with preserved ejection fraction: an overview. *J Cardiovasc Transl Res* 2017;10(3):233–44.
- Dunlay SM, Roger VL, Redfield MM. Epidemiology of heart failure with preserved ejection fraction. *Nat Rev Cardiol* 2017;14(10):591–602.
- Olsen MB, Hildrestrand GA, Scheffler K, et al. NEIL3-dependent regulation of cardiac fibroblast proliferation prevents myocardial rupture. *Cell Rep* 2017;18(1):82–92.
- Skarpengland T, Holm S, Scheffler K, et al. Neil3-dependent base excision repair regulates lipid metabolism and prevents atherosclerosis in ApoE-deficient mice. *Sci Rep* 2016;6:28337.
- Yim J, Cho H, Rabkin SW. Gene expression and gene associations during the development of heart failure with preserved ejection fraction in the Dahl salt sensitive model of hypertension. *Clin Exp Hypertens* 2018;40(2):155–66.
- Das S, Frisk C, Eriksson MJ, et al. Transcriptomics of cardiac biopsies reveals differences in patients with or without diagnostic parameters for heart failure with preserved ejection fraction. *Sci Rep* 2019;9(1):3179.
- Pohlmann L, Kroger I, Vignier N, et al. Cardiac myosin-binding protein C is required for complete relaxation in intact myocytes. *Circ Res* 2007;101(9):928–38.
- Poller W, Dimmeler S, Heymans S, et al. Non-coding RNAs in cardiovascular diseases: diagnostic and therapeutic perspectives. *Eur Heart J* 2018;39(29):2704–16.
- Chen YT, Wong LL, Liew OW, et al. Heart Failure with Reduced Ejection Fraction (HFrEF) and Preserved Ejection Fraction (HFpEF): the diagnostic value of circulating microRNAs. *Cells* 2019;8(12):1651.
- Wong LL, Armugam A, Sepamaniam S, et al. Circulating microRNAs in heart failure with reduced and preserved left ventricular ejection fraction. *Eur J Heart Fail* 2015;17(4):393–404.
- Humphries B, Wang Z, Li Y, et al. ARHGAP18 downregulation by miR-200b suppresses metastasis of triple-negative breast cancer by enhancing activation of RhoA. *Cancer Res* 2017;77(15):4051–64.
- Singha B, Harper SL, Goldman AR, et al. CLIC1 and CLIC4 complement CA125 as a diagnostic biomarker panel for all subtypes of epithelial ovarian cancer. *Sci Rep* 2018;8(1):14725.
- Ndrepepa G. Myeloperoxidase-A bridge linking inflammation and oxidative stress with cardiovascular disease. *Clin Chim Acta* 2019;493:36–51.
- Khansari N, Shakiba Y, Mahmoudi M. Chronic inflammation and oxidative stress as a major cause of age-related diseases and cancer. *Recent Pat Inflamm Allergy Drug Discov* 2009;3(1):73–80.
- Henning RJ, Bourgeois M, Harbison RD. Poly (ADP-ribose) polymerase (PARP) and PARP inhibitors: mechanisms of action and role in cardiovascular disorders. *Cardiovasc Toxicol* 2018;18(6):493–506.
- Tsutsui H, Kinugawa S, Matsushima S. Oxidative stress and mitochondrial DNA damage in heart failure. *Circ J* 2008;72 Suppl A: A31–7.
- Batle M, Recarte-Pelz P, Roig E, et al. AXL receptor tyrosine kinase is increased in patients with heart failure. *Int J Cardiol* 2014;173(3):402–9.
- Wei SG, Yu Y, Weiss RM, et al. Endoplasmic reticulum stress increases brain MAPK signaling, inflammation and renin-angiotensin system activity and sympathetic nerve activity in heart failure. *Am J Physiol Heart Circ Physiol* 2016;311(4):H871–80.
- Warbrick I, Rabkin SW. Hypoxia-inducible factor 1-alpha (HIF-1 α) as a factor mediating the relationship between obesity and heart failure with preserved ejection fraction. *Obes Rev* 2019;20(5):701–12.
- Negi SI, Jeong EM, Shukrullah I, et al. Renin-angiotensin activation and oxidative stress in early heart failure with preserved ejection fraction. *Biomed Res Int.* 2015;2015:825027.
- Zhang W, Zhang H, Yao W, et al. Morphometric, hemodynamic, and multi-omics analyses in heart failure rats with preserved ejection fraction. *Int J Mol Sci* 2020;21(9):3362.
- Shah SJ, Katz DH, Selvaraj S, et al. Phenomapping for novel classification of heart failure with preserved ejection fraction. *Circulation* 2015;131(3):269–79.
- Cohen JB, Schrauben SJ, Zhao L, et al. Clinical phenogroups in heart failure with preserved ejection fraction: detailed phenotypes, prognosis, and response to spironolactone. *JACC Heart Fail* 2020;8(3):172–84.
- Segar MW, Patel KV, Ayers C, et al. Phenomapping of patients with heart failure with preserved ejection fraction using machine learning-based unsupervised cluster analysis. *Eur J Heart Fail* 2020;22(1):148–58.
- Maeder ML, Gersbach CA. Genome-editing technologies for gene and cell therapy. *Mol Ther* 2016;24(3):430–46.
- Hilberg F, Roth GJ, Krssak M, et al. BIBF 1120: triple angiokinase inhibitor with sustained receptor blockade and good antitumor efficacy. *Cancer Res* 2008;68(12):4774–82.
- Hostettler KE, Zhong J, Papakonstantinou E, et al. Anti-fibrotic effects of nintedanib in lung fibroblasts derived from patients with idiopathic pulmonary fibrosis. *Respir Res* 2014;15(1):157.
- Rangarajan S, Kurundkar A, Kurundkar D, et al. Novel mechanisms for the anti-fibrotic action of nintedanib. *Am J Respir Cell Mol Biol* 2016;54(1):51–9.
- Kanefendi F, Thuß U, Becka M, et al. Pharmacokinetics, safety, and tolerability of the novel chymase inhibitor BAY 1142524 in healthy male volunteers. *Clin Pharmacol Drug Dev* 2019;8(4):467–79.

- [59] Troughton RW, Richards AM. Outpatient monitoring and treatment of chronic heart failure guided by amino-terminal pro-B-type natriuretic peptide measurement. *Am J Cardiol* 2008;101(3A):72–5.
- [60] Pfisterer M, Buser P, Rickli H, et al. BNP-guided vs symptom-guided heart failure therapy: The Trial of Intensified vs Standard Medical Therapy in Elderly Patients with Congestive Heart Failure (TIME-CHF) randomized trial. *JAMA* 2009;301(4):383–92.
- [61] Sharma K, Kass DA. Heart failure with preserved ejection fraction: mechanisms, clinical features, and therapies. *Circ Res* 2014;115(1):79–96.
- [62] Shah SJ, Borlaug BA, Kitzman DW, et al. Research priorities for heart failure with preserved ejection fraction: national heart, lung, and blood institute working group summary. *Circulation* 2020;141(12):1001–26.