

## Article

# Identifying *OGN* as a Biomarker Covering Multiple Pathogenic Pathways for Diagnosing Heart Failure: From Machine Learning to Mechanism Interpretation

Yihao Zhu <sup>1,2</sup>, Bin Chen <sup>3,\*</sup> and Yao Zu <sup>1,2,4,\*</sup> 

- <sup>1</sup> International Research Center for Marine Biosciences, Ministry of Science and Technology, Shanghai Ocean University, Shanghai 201306, China
- <sup>2</sup> Key Laboratory of Exploration and Utilization of Aquatic Genetic Resources, Ministry of Education, Shanghai Ocean University, Shanghai 201306, China
- <sup>3</sup> Department of Cardiology, Shanghai Sixth People's Hospital Affiliated to Shanghai Jiao Tong University School of Medicine (Lin-gang), Shanghai 201306, China
- <sup>4</sup> Marine Biomedical Science and Technology Innovation Platform of Lin-gang Special Area, Shanghai 201306, China
- \* Correspondence: taihe-chenbin@163.com (B.C.); yzu@shou.edu.cn (Y.Z.)

**Abstract:** Background: The pathophysiologic heterogeneity of heart failure (HF) necessitates a more detailed identification of diagnostic biomarkers that can reflect its diverse pathogenic pathways. Methods: We conducted weighted gene and multiscale embedded gene co-expression network analysis on differentially expressed genes obtained from HF and non-HF specimens. We employed a machine learning integration framework and protein–protein interaction network to identify diagnostic biomarkers. Additionally, we integrated gene set variation analysis, gene set enrichment analysis (GSEA), and transcription factor (TF)-target analysis to unravel the biomarker-dominant pathways. Leveraging single-sample GSEA and molecular docking, we predicted immune cells and therapeutic drugs related to biomarkers. Quantitative polymerase chain reaction validated the expressions of biomarkers in the plasma of HF patients. A two-sample Mendelian randomization analysis was implemented to investigate the causal impact of biomarkers on HF. Results: We first identified *COL14A1*, *OGN*, *MFAP4*, and *SFRP4* as candidate biomarkers with robust diagnostic performance. We revealed that regulating biomarkers in HF pathogenesis involves TFs (*BNC2*, *MEOX2*) and pathways (cell adhesion molecules, chemokine signaling pathway, cytokine–cytokine receptor interaction, oxidative phosphorylation). Moreover, we observed the elevated infiltration of effector memory CD4<sup>+</sup> T cells in HF, which was highly related to biomarkers and could impact immune pathways. Captopril, aldosterone antagonist, cyclophosphamide, estradiol, tolazoline, and genistein were predicted as therapeutic drugs alleviating HF via interactions with biomarkers. In vitro study confirmed the up-regulation of *OGN* as a plasma biomarker of HF. Mendelian randomization analysis suggested that genetic predisposition toward higher plasma *OGN* promoted the risk of HF. Conclusions: We propose *OGN* as a diagnostic biomarker for HF, which may advance our understanding of the diagnosis and pathogenesis of HF.

**Keywords:** heart failure; diagnostic biomarker; pathogenic mechanism; bioinformatics



**Citation:** Zhu, Y.; Chen, B.; Zu, Y. Identifying *OGN* as a Biomarker Covering Multiple Pathogenic Pathways for Diagnosing Heart Failure: From Machine Learning to Mechanism Interpretation.

*Biomolecules* **2024**, *14*, 179. <https://doi.org/10.3390/biom14020179>

Academic Editors: Cornelia Deeg, Jelena Skuljec and Thomas R. Caulfield

Received: 21 November 2023

Revised: 14 January 2024

Accepted: 23 January 2024

Published: 2 February 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Heart failure (HF) is a heterogeneous and life-threatening clinical syndrome encompassing symptoms such as breathlessness, excessive fatigue, and swollen ankles [1]. HF has been considered a global pandemic estimated to affect more than 64 million individuals worldwide. Its prevalence has continuously increased over the past decades, posing a great threat to public health [2,3]. Traditionally, HF was defined as a pathophysiological condition in which cardiac structure or function abnormality results in increased intra-cardiac

pressures or reduced cardiac output at rest or during activities [4]. However, the traditional definition only mentioned the non-specific symptoms or signs in clinical practice but did not include objective indicators to support HF diagnosis [4,5]. In 2021, a universal consensus on the definition of HF was proposed, namely a clinical syndrome with symptoms and/or signs caused by structural and/or functional cardiac abnormalities and corroborated by elevated natriuretic peptide levels and/or objective evidence of pulmonary or systemic congestion by diagnostic modalities [4], which was conceptually comprehensive and clinically practical with at least one objective detection indicator included. Notably, the elevated levels of natriuretic peptides, including the B-type natriuretic peptide (BNP) and N-terminal pro-B-type natriuretic peptide (NT-proBNP), were enrolled in this newly proposed definition of HF, which suggested the pivotal role of objective detective signs in HF recognition.

Biomarker-guided diagnosis and management have gained popularity in clinical application [5], and the natriuretic peptide has become a golden standard widely used in HF diagnosis or prognostic evaluation [6]. Nevertheless, the elevated concentrations of natriuretic peptides do not always meet the commands of diagnosing HF in clinics because it might be affected by numerous factors, such as renal failure, pulmonary embolism, and obesity [7]. Previously, evidence has shown that a deterioration in renal function significantly elevates the concentrations and associations of natriuretic peptides, indicating that the clinical interpretation of increased natriuretic peptides should consider renal function [8]. In the condition of acute pulmonary embolism, the concentrations of natriuretic peptides were also identified to increase as well as predict adverse prognostic outcomes [9]. Additionally, it has been demonstrated that obese patients with HF show a state of decreased levels of natriuretic peptides. Thus, lower thresholds of natriuretic peptide levels should be made to meet the diagnosis of obese individuals with HF [10]. Therefore, the application of natriuretic peptides remains limited in HF because of its vulnerability to other pathophysiological conditions, indicating that exploring novel biomarkers that could provide additional diagnostic utility is urgently needed [11]. Given the complex pathophysiology of HF, which ranges from cardiac dysfunction to extra-cardiac alteration, emerging HF biomarkers that tackle different pathophysiological processes could significantly benefit the risk stratification and management of patients with HF [12]. In the past decades, two novel identified HF biomarkers, Galectin-3 (Gal-3) and tumorigenesis-2 (sST2), have yet to prove their utility in the clinics [13]. Moreover, growing evidence suggested that the multi-biomarker panel strategy, combining novel biomarkers with well-established ones, showed both strong diagnostic and prognostic values compared to a single-biomarker approach [5,14]. Hence, we hypothesized investigating newer specific diagnostic biomarkers together with elucidating their regulated pathogenic pathways involved in HF, which could provide candidates for the multi-biomarker strategy, thus deepening our understanding of molecular alterations in the progression of HF.

Given the explosion of biomedical big data, there is growing research enthusiasm in utilizing bioinformatics approaches for mining high-throughput RNA sequencing (RNA-seq) or single-cell RNA-seq (scRNA-seq) data, which make us better understand the altered molecular features in HF at tissue or cell levels [15]. With these rapidly developed sequencing techniques and computational analysis approaches, novel biomarkers and mechanisms could be rapidly identified rather than speckled through traditional hypothetical-driven research [5,15], bringing new insight into understanding the physiopathology of HF and developing the diagnosis or clinical management. Although several bioinformatics studies have already tackled exploring the underlying biomarkers and molecular mechanisms of HF [16–19], apparent limitations existed mainly due to the lack of comprehensive multi-datasets analysis from both tissue and single-cell views, rational speculation of the biomarkers-regulated pathways, and external experimental validation of biomarkers on testing patients' samples. Considering these shortcomings in previous research, we intended to mine novel candidate biomarkers of HF as well as explicitly interpret their mechanisms (including upstream and downstream ways) that participated in HF via the

in silico analysis on multiple RNA-seq datasets (including bulk RNA-seq and scRNA-seq) and in vitro validation on plasma collected from HF patients and control cases.

In this study, multiple transcriptional datasets derived from patients with HF and control, including four expression profiles (GSE141910, GSE57338, GSE42955, GSE135055) and one sc-RNA seq dataset (GSE121893), were acquired from the Gene Expression Omnibus (GEO) database. Training the cohort (GSE141910) for major analysis and test cohorts (GSE57338 and GSE42955) for validation were then processed and obtained, respectively. GSE135055 served as an external validation cohort. First, the differential expression genes (DEGs) between the HF and control were identified. Second, weighted gene co-expression network analysis (WGCNA) and multiscale embedded gene co-expression network analysis (MEGENA) complementing each other via different algorithms were performed on the DEG expression profiles to filter key gene modules. Subsequently, Disease Ontology (DO), Gene Ontology (GO), and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses were conducted on the overlapping genes between WGCNA-identified and MEGNA-identified modules, and genes significantly enriched in HF-related terms were retained. Fourth, the retained genes were subjected to a machine learning-based integration workflow and protein–protein interaction (PPI) network selection. Fifth, hub genes performed with robust diagnostic in the training and test cohorts were generated. Hub genes' expression and correlation patterns were evaluated in the training and test cohorts. Sixth, sc-RNA seq analysis was conducted on GSE121893 to uncover the cell-specific expression patterns of hub genes. Seventh, together with gene set variation analysis (GSVA), gene set enrichment analysis (GSEA), transcription factor (TF) prediction, and single-sample gene set enrichment analysis (ssGSEA), we systemically identified the pathways co-regulated by hub genes involved in HF. Then, we used the ssGSEA approach to quantify the infiltration levels of immune cells in HF to screen crucial immune cell types highly correlated to the hub genes. Eighth, the therapeutic drugs targeting hub genes and TFs were verified using molecular docking. Finally, quantitative reverse transcription PCR (RT-qPCR) was utilized to explore the relative mRNA expression levels of hub genes in plasma collected from patients with HF, thereby developing plasma biomarkers with HF. Based on large-scale genome-wide association studies (GWASs), a two-sample Mendelian randomization (MR) analysis was further performed to investigate the causal effect of hub genes on the risk of HF. The overall study flow chart is shown in Figure S1.

## 2. Materials and Methods

### 2.1. Gathering and Processing of Gene Expression Profiling Datasets

Three gene expression profiles (GSE141910, GSE57338, GSE42955, GSE135055) of left ventricles derived from HF patients were retrieved from the Gene Expression Omnibus (GEO) database (<https://www.ncbi.nlm.nih.gov/geo/>, accessed on 30 November 2022). GSE141910 was the training group in this study, which included 200 HF and 166 non-HF (control) samples [20]. We removed the batch effect of GSE57338 and GSE42955 and merged them into the test group using the “combat” algorithm of the R packages “sva” (version 3.44.0), containing 201 HF and 141 non-HF (control) samples [21,22]. The removal of batch effect on GSE57338 and GSE42955 was evaluated via the principal component analysis (PCA) plot, as shown in Figure S2. Furthermore, we selected GSE135055 (21 HF and 9 control samples) as an external validation group [23]. We then performed log transformation on the gene expression profiles and utilized the function “normalizeBetweenArrays” of the R package “limma” (version 3.52.2) to execute data normalization [24]. The detailed information on these enrolled datasets is summarized in Table 1. Moreover, the clinical characteristics included in the datasets can be found in Tables S1–S4.

**Table 1.** Information of HF-related datasets enrolled in study.

Accession	Sample Source	Sequencing Type	Control Samples	HF Samples	Dataset Usage
GSE141910	left ventricle	RNA-seq	166	200	Training dataset
GSE57338	left ventricle	Array	136	177	Testing dataset
GSE42955	left ventricle near the apex	Array	5	24	Testing dataset
GSE135055	left ventricle	RNA-seq	9	21	External validation dataset

## 2.2. Partial Least Squares Discriminant Analysis, DEGs Identification, and KEGG Pathway Enrichment Analysis

To evaluate the separation between control and HF groups, partial least squares discriminant analysis (PLS-DA) was performed using the R package “mixOmics” (version 6.20.0) [25]. The DEGs between the HF group and the control group (HF versus control) were identified by the R package “limma” [24]. The cut-off criteria of DEGs were set as  $|\log_2\text{Fold Change} (\log_2\text{FC})| > 1$  and adjusted  $p$ -value  $< 0.05$  [26]. Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis was then conducted on DEGs through the R package “clusterProfiler” (version 4.7.1) [27]. The enrichment KEGG pathway showed with  $q$  value  $< 0.05$  was considered statistically significant.

## 2.3. Weighted Gene Co-Expression Network Analysis on DEGs

The R package “WGCNA” (version 1.71) was utilized to construct a weighted gene co-expression network on DEGs, which was suitable for further screening functional gene modules highly associated with HF [28]. The soft threshold power ( $\beta = 1\sim 20$ ) was first determined to establish a scale-free topology. The weighted adjacency matrix was then generated and transformed into a topological overlap matrix (TOM). Moreover, the dissTOM was obtained for hierarchical clustering, and the dynamic tree-cutting method was adopted to screen various modules clustered on gene similarity. Here, the parameters “minModuleSize” and “MEDissThres” were set to 60 and 0.3, respectively. Subsequently, we related the recognized modules to two traits (control and HF). Genes in the module that displayed high relevance with HF were selected for further analysis.

## 2.4. Multiscale Embedded Gene Co-Expression Network Analysis on DEGs

The R package “MEGENA” (version 1.3.7) was performed to identify gene co-expression networks on DEGs by Planar Filtered Networks (PFNs) construction, thus dissecting into multiscale gene modular structures in HF [29]. The input parameters were set to default. First, we obtained a fast PFN via calculating, filtering, and clustering correlations between each gene pair based on the DEG expression profiles of HF. Multiscale clustering analysis (MCA) was then conducted on PFN, and sub-modules were identified from the connected components of the initial PFN as the parent modules. Next, we implemented multiscale hub analysis to detect key hubs of individual modules. The identified modules with large gene sizes were retained for later selection.

## 2.5. DO, GO, and KEGG Enrichment Analyses on Crucial Gene Modules

Taking the intersection between the significantly co-expressed gene modules identified by WGCNA and MEGENA, we obtained overlapping genes for functional and pathway annotation. Annotation of human diseases, biological functions, and signaling pathways could reflect whether the overlapping genes in these identified modules are directly enriched in HF-related terms [30], contributing to our understanding of these genes' role in HF. Accordingly, we systemically performed Human Disease Ontology (DO), Gene Ontology (GO), and KEGG enrichment analyses on these overlapping genes using the R packages “clusterProfiler” and “DOSE” (version 3.24.2) [27,31]. A  $q$  value  $< 0.05$  was considered the cut-off criterion for identifying significant enrichment terms.

### 2.6. Feature Selection via Machine Learning-Based Integration

The overlapping genes co-identified via WGCNA and MEGENA were subjected to a machine learning-based integration pipeline [32]. A total of 12 machine learning models for integration included least absolute shrinkage and selection operator (Lasso), Ridge, elastic network (Enet), stepwise multiple generalized linear model (Stepglm), support vector machine (SVM), generalized linear model by likelihood-based boosting (glmBoost), linear discriminant analysis (LDA), partial least squares regression for generalized linear models (plsRglm), Random Forest (RF), gradient boosting machine (GBM), eXtreme Gradient Boosting (XGBoost), and naïve Bayes. Herein, the machine learning-based integration pipeline applied one algorithm for feature selection and another for constructing an integrative classification prediction model under the 10-fold cross-validation (CV), generating 113 algorithm combinations in this study. The execution of the machine learning-based integration pipeline involves the following procedures. (1) We first performed Z-score transformation on the expression profiles of training and testing cohorts, which could enhance comparability between diverse cohorts and accelerate running speed. (2) In our integrative ML-based framework, four algorithms with feature selection capability (Lasso, RF, Stepglm, glmBoost) were initially conducted to narrow down the genes. Then, we adopted eight other algorithms (Ridge, Enet, SVM, LDA, plsRglm, GBM, XGBoost, NaiveBayes) to fit prediction models based on the genes selected via the four algorithms, respectively. In total, 113 model combinations were used to tune hyperparameters and fit models under a 10-fold CV framework. (3) All model combinations were subsequently tested in training and testing cohorts. For the performance evaluation of each model, the AUC score across these cohorts was computed. The model with the highest average AUC within the training and testing cohorts was regarded as optimal. Feature genes were subsequently obtained via the best-performing model.

### 2.7. Construction of Protein-Protein Interaction Network

Leveraging the feature genes generated from machine learning-based integration, a protein-protein interaction (PPI) network was constructed using the Search Tool of the Retrieval of Interacting Genes (STRING) database and then visualized by Cytoscape software (version 3.9.1) [33,34]. We further screened out the top 20 genes with higher connectives in the PPI network using five ranking algorithms of Cytoscape plug-in cytoHubba, respectively. The Degree, Maximum Neighborhood Component (MNC), Maximal Clique Centrality (MCC), Edge Percolated Component (EPC), and Density of Maximum Neighborhood Component (DMNC) algorithms were used to screen genes in the PPI network [35]. We noted the intersected genes co-selected in five algorithms as pivotal genes of HF for further selection and assessment.

### 2.8. Diagnostic Abilities, Expression Levels, and Correlation Pattern of Hub Genes

To investigate the diagnostic performances of pivotal genes in HF, we calculated the AUC score of receiver operating characteristic curve (ROC) analysis using the R package “pROC” (version 1.18.0) to measure the classification accuracy of the HF group and control group [36]. Genes performed with an AUC higher than 0.85 distinguishing HF from control in both the training and test cohorts were considered hub genes, potentially serving as biomarkers of HF diagnosis. We then utilized the R package “glmnet” (version 4.1-4) to construct logistic regression models to evaluate the joint diagnostic abilities of hub genes in the training and test cohorts. To compare the diagnostic abilities of our identified hub genes with clinical biomarkers, we retrieved a panel of current HF biomarkers, including myocardial stress/injury-related (BNP, NT-proBNP, TnI, TnI), neurohormone-related (MR-proADM, AVP, Peptide), myocardial remodeling-related (sST2, Gal-3, MMP, GDF-15), inflammatory-related (CRP, IL-6, TNF- $\alpha$ , PCT), and renal function injury-related (Cys C, NGAL) biomarkers. Comparison of the predictive power of hub genes and current biomarkers was then assessed via AUC score within the training and testing cohorts. Using the R package “ggpubr” (version 0.4.0), the significant over-expression pattern of hub genes

in HF compared to the control was shown in the violin plot and verified in the test cohort. Correlation analyses between hub genes in the training and test cohorts were performed and visualized using the R package “PerformanceAnalytics” (version 2.0.4).

### 2.9. Single-Cell RNA-seq Data Processing

To explore the specific cell expression patterns of hub genes in HF, we obtained one scRNA-seq dataset (GSE121893) derived from six patients with HF [37], including 1682 cardiomyocytes (CMs) and 2539 non-cardiomyocytes (NCMs). The R package “Seurat” (version 4.1.1) was implemented to process and analyze single-cell RNA-seq datasets [38]. First, we conducted quality control and normalization of the data. Cells with unique feature counts less than 200 and mitochondrial proportions larger than 5% were eliminated in further analysis. The “Log normalize” function was then employed to normalize the filtered data. Second, we obtained the top 2000 highly variable genes identified by the function “FindVariableFeatures” and scaled the data as the pre-processing procedure of subsequent dimensional reduction. Third, we performed principal component analysis (PCA) on the scaled data with the top 2000 highly variable genes as input. Combining the JackStraw and Elbow plots, we determined the optimal components (1 to 15) to cluster cells. Fourth, the Uniform Manifold Approximation and Projection (UMAP) algorithm was executed to display the cell clusters. We then assigned five cell types (CM, cardiomyocyte; EC, endothelial cell; FB, fibroblast; SMC, smooth muscle cell; MP, macrophage) to the identified clusters according to the expression distributions of gene markers. The markers of five cell populations were listed: CM (*TNNT2*, *MYH7*, *MYOM1*, *ACTN2*), EC (*VWF*, *ENG*, *RGCC*, *EMCN*), FB (*FN1*, *VIM*, *DCN*, *VCAN*), SMC (*ACTA2*, *MYLK*, *MYL6*, *MYL9*), and MP (*CCR2* negative, *CD74*, *ITGAM*, *MRC1*) [37]. We also used the R package “AUCell” (version 1.22.0) to calculate the AUCell scores of the specific gene markers for each cell to reflect the marker activities within the five cell populations [39]. The R package “Nebulosa” (version 1.6.0), a novel gene expression signal recovering approach on weighted kernel density estimation, was utilized to uncover the expression distributions of hub genes in the five annotated cell populations [40].

### 2.10. GSVA and GSEA of Hub Genes

Herein, we integrated gene set variation analysis (GSVA) and gene set enrichment analysis (GSEA) to investigate the underlying pathways of each hub gene in HF. The reference gene sets of GSEVA and GSEA were the KEGG subset of canonical pathways collected from the Human Molecular Signatures Database (MSigDB) (<http://www.gsea-msigdb.org/gsea/msigdb/collections.jsp>, accessed on 30 November 2022) [41]. We first performed GSVA on the expression profiles of HF leveraging the R package “GSVA” (version 1.44.2) [42]. Then, the R package “limma” was conducted to compare the GSVA scores of pathways between the low-expression and high-expression groups according to the medium value of the hub gene.  $p$  value  $< 0.05$  and  $|t| > 2$  were considered as cut-off criteria for identifying the significant activated ( $t > 2$ ) or suppressed ( $t < 2$ ) pathways [43]. Meanwhile, GSEA was conducted for each hub gene based on the expression profiles of HF using the R package “clusterProfiler” [27,43]. The GSEA pathways with  $|$ normalized enrichment score (NES) $| > 1$ ,  $p$  value  $< 0.05$ , and false discovery rate (FDR)  $< 0.25$  were regarded as statistically significant enrichment [44]. NES larger than 1 and lower than 1 indicated the activation and suppression of the enriched pathway, respectively. The statistically significant overlap between the GSEA and GSVA on each hub gene was identified as the common pathways (activated or suppressed) co-regulated by hub genes in HF. Moreover, the R package “CBNplot” (version 0.99.2) was employed to construct and visualize the gene regulatory network (GRN) via Bayesian network (BN) inference from expression profiles and enrichment analysis [45].

### 2.11. Expression Levels of Predicted TFs and Their Interaction with Hub Gene-Related Pathways

To gain insight into the upstream mechanism regulating hub genes, we acquired the TRRUST v2 database ([www.grnpedia.org/trrust](http://www.grnpedia.org/trrust), accessed on 23 March 2023) to search for underlying transcription factors (TFs) interacting with hub genes [46]. The top 10 predicted TFs with higher-ranked sentences in the database were retained. TFs that were up-regulated in HF may actively target hub genes. Next, we performed a correlation analysis between the key TFs and hub genes. Additionally, single-sample gene set enrichment analysis (ssGSEA) was employed using the R packages “GSVA” and “GSEABase” (version 1.58.0) to quantify the enrichment levels of the pathways related to hub genes in HF samples [42,47]. Subsequently, we conducted a correlation analysis between the key TF expressions and pathway enrichment levels.

### 2.12. Evaluation of Immune Cell Characteristics in HF via ssGSEA and Identifying Immune Cell Types Highly Correlated with Hub Genes

To explore whether the immune landscapes were altered in HF, we employed the ssGSEA algorithm to quantify the infiltration levels of 28 immune cell types based on gene expression profiles [48,49]. Using ssGSEA by interrogating the expression levels of reported specific genes of adaptive and innate immune cell types, we computed the infiltration levels of 28 immune cell types per sample. Next, we integrated Lasso and RF algorithms to obtain key immune cell types from 28 immune cell types [50]. Lasso and RF were constructed under 10-fold cross-validation, and the setting parameters of RF were 800 decision trees. Using the R package “glmnet”, a Lasso regression model was established under the optimal penalization coefficient ( $\lambda$ ), and immune cell types with regression coefficients larger than zero were retained [51]. The R package “randomForest” (version 4.6-7) was utilized to construct an RF classifier using 586 decision trees that was performed with a minimum classification error rate, and immune cell types with mean decrease Gini scores larger than five were screened [52]. Taking the intersection between Lasso and RF, we identified a final set of key immune cell types in HF. Subsequently, we assessed the correlation patterns between hub genes and key immune cell types that were significantly up-regulated in HF compared to control.

### 2.13. Molecular Docking of Predictive Drugs Targeting Hub Genes and TFs

First, candidate small-molecule agents with targets (hub genes or key TFs) were predicted using the Drug Signatures Database (DSigDB) [53]. Molecular docking, a computational approach for exploring the interaction between receptors and ligands, was executed to evaluate the binding affinities of the selected candidate drugs to their targets [54]. AutoDock Vina (version 1.1.2) was used to perform molecular docking analysis [55]. Second, the 2D structures of molecule ligands (candidate drugs) were downloaded from the PubChem database (<https://pubchem.ncbi.nlm.nih.gov/>, accessed on 26 March 2023) in SDF formats. The ChemBio3D was performed to minimize energy on molecule ligands and export their 3D structures. AutoDock Vina then converted the molecule ligands (mol2 format) into PDBQT format for further docking analysis. Third, the 3D structures of protein receptors (hub genes and key TFs) were passed through the Protein Data Bank (PDB) (<https://www.rcsb.org/>, accessed on 26 March 2023). We used the PyMOL (version 2.4.0) to remove water molecules and ligands from proteins. The protein receptors (PDB format) were imported into the AutoDock Vina and exported in the PDBQT format after polar hydrogenation and docking site setting. Herein, the parameters of the docking site, including X-Y-Z coordinates and grid size, were adjusted to include the active pocket that completely binds molecule ligands. Finally, AutoDock Vina was executed to dock the molecule ligand and protein receptors 20 times. Confident docking results with the lowest binding affinity and Root Mean Squared Error (RMSE) lower than 2 Å were retained and subsequently visualized by PyMOL.

#### 2.14. Quantitative Reverse Transcription PCR

To investigate whether the hub genes could serve as plasma biomarkers of HF, we collected the whole blood specimens stored in ethylenediaminetetraacetic acid (EDTA) tubes, derived from three HF patients and three non-HF controls in the Shanghai Sixth People's Hospital Affiliated to the Shanghai Jiao Tong University School of Medicine (Lingang). The detailed information of patients, including gender, age, disease diagnosis, HF grade, HF type, ejection fraction (EF), and NT-proBNP concentration, are shown in Table S5. The types of HF of the three patients include congestive heart failure and ischemic heart failure. Whole blood was first centrifuged at 3500 relative centrifugal force (RCF) for 10 min at room temperature, and plasma was separated. RNA was extracted from the collected plasma using a BIOG cfRNA Easy Kit (Changzhou Baidai, Changzhou, China). Subsequently, we utilized the PrimeScript RT Reagent Kit with a gDNA Eraser Kit (Takara, Kusatsu, Japan) to reverse transcribe RNA into cDNA. Quantitative reverse transcription PCR (qRT-PCR) analysis was performed on cDNA using the AceQ Universal SYBR qPCR Master Mix (Vazyme, Nanjing, China). The relative expression of hub genes was calculated by the 2-DDCt method adjusted to  $\beta$ -actin. The primer sequences of four hub genes and  $\beta$ -actin are shown in Table S6.

#### 2.15. Two-Sample Mendel Randomization Analysis

We sought to explore the causal effect of plasma OGN (exposure) on HF (outcome) using a two-sample Mendel Randomization (MR) analysis. Herein, we used the R package "TwoSampleMR" (version 0.5.7) to analyze. The exposure data (plasma OGN; ID: PRJEB15197) from 35,559 Icelanders and outcome data (HF, ID: ebi-a-GCST009541) from 47,309 Europeans with HF reported were retrieved from the deCODE genetics (<https://www.decode.com/summarydata/>, accessed on 27 December 2023) and the IEU Open GWAS data source (<https://gwas.mrcieu.ac.uk/>, accessed on 27 December 2023), respectively. Single-Nucleotide Polymorphisms (SNPs) used as instrumental variables (IVs) in MR analysis should be subject to three assumptions [56]: (1) SNPs are closely related to exposure; (2) SNPs are independent of cofounders of exposure and outcome; (3) SNPs only affect the outcome through exposure. Accordingly, the function "extract\_instruments" was utilized to retain the SNPs that were significantly associated with OGN ( $p$ -value  $< 5 \times 10^{-6}$ ) but not with HF ( $p$ -value  $> 0.05$ ) [57]. SNPs with linkage disequilibrium (LD)  $R^2$  larger than 0.01 within a cropping range of 5000 Kb were excluded [57]. Next, the function "harmonise\_data" was leveraged to harmonize the exposure and outcome data. Two-sample MR analysis was subsequently performed via the function "mr". Five default methods were adopted: MR Egger, Weighted median, Inverse variance weighted (IVW), Simple mode, and Weighted mode. The R package "forestploter" (version 1.1.1) [58] was employed to generate a forest plot showing the Wald ratio-estimated odds ratio (OR) of OGN on HF. The IVW approach, with the highest statistical power [56,57], was selected to illustrate the association between the level of plasma OGN and the risk of AMI. We employed heterogeneity, horizontal pleiotropy, and leave-one-out tests as sensitivity analysis as well as serving as a reliability evaluation of MR analysis.

#### 2.16. Statistical Analysis

Statistical analysis was performed using R programming (version 4.2.2). A Wilcoxon rank-sum test was used to analyze the two groups' differential expression levels of hub genes. We performed Pearson's correlation analysis on expression levels of hub genes. Spearman's correlation analysis was conducted to explore the relationship between infiltrations of immune cell types and expressions of hub genes. Spearman's correlation analysis was also utilized to assess the interaction between expressions of TFs and ssGSEA scores of pathways. The results of qRT-PCR were presented as mean  $\pm$  standard error (SD), and an unpaired t-test was carried out to compare the mean expressions of hub genes between HF and control.  $p$  value  $< 0.05$  was considered statistically significant.

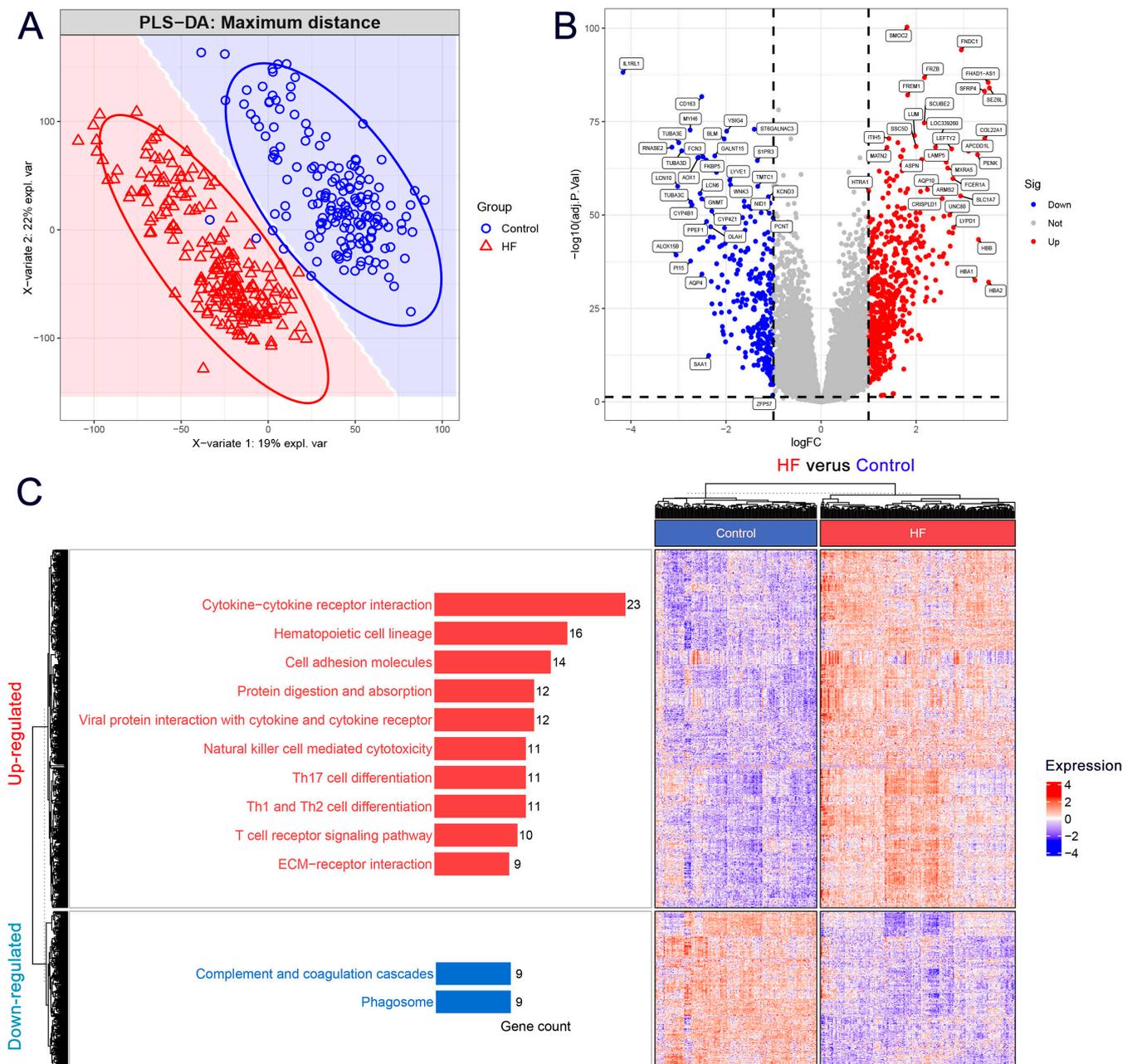
### 3. Results

#### 3.1. DEGs Mainly Participated in the Extracellular Matrix and Immune-Related Pathways in HF

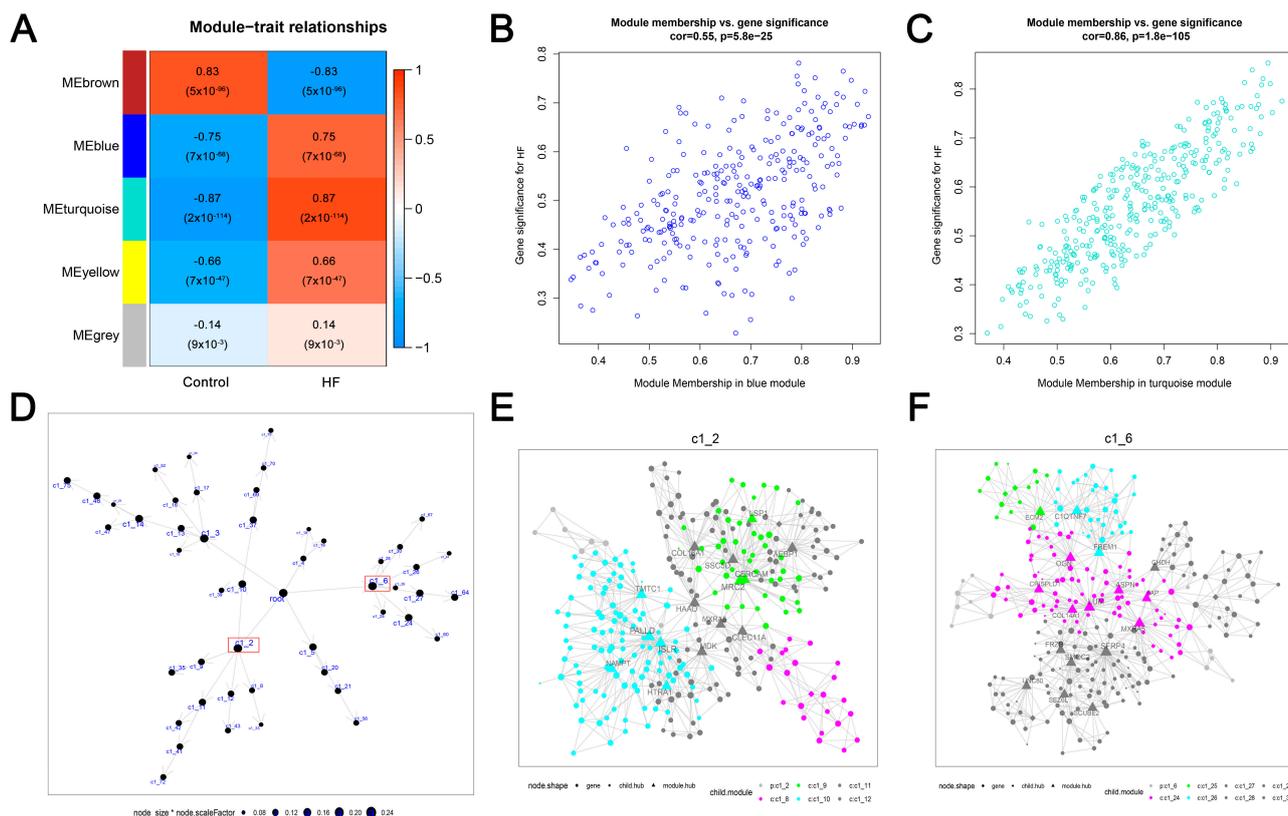
To investigate the differences between HF and control groups, we performed PLS-DA on the expression profile of the training cohort. Figure 1A shows a distinct sample separation between HF and control groups, which could be used for DEG analysis. We then identified 926 DEGs (HF group versus control group) with  $\log_2FC$  larger than 1 and adjusted  $p$  value lower than 0.05, including 648 up-regulated DEGs and 278 down-regulated DEGs (Figure 1B). Detailed information on these 926 DEGs can be found in Table S7. KEGG enrichment analysis was then performed to understand better the potential functions of these selected DEGs. The significantly enriched pathways of up-regulated and down-regulated DEGs are shown in Figure 1C (left panel), respectively. The up-regulated DEGs were mainly involved in cytokine–cytokine receptor interaction, cell adhesion molecules, the T cell receptor signaling pathway, and ECM–receptor interaction. Interestingly, these enrichment pathways were largely related to the extracellular matrix (ECM) or immune-related functions, suggesting their crucial roles positively participated in HF. The down-regulated DEGs mainly participated in complement and coagulation cascades and the phagosome, which indicated that coagulation functions were abnormally altered in HF. We also observed the distinct expression patterns of DEGs in the clustering heatmap in Figure 1C (right panel), which illustrates that the DEGs could clearly distinguish the samples as HF from the control.

#### 3.2. Two Crucial Modules Strongly Related to HF Were Identified by WGCNA on DEGs

To understand the gene expression network in HF, we performed WGCNA to establish a weighted gene co-expression network based on the expression profiles of the identified 926 DEGs. First, a similarity matrix was generated by calculating gene pairwise correlation. We then evaluated soft threshold power ( $\beta$ ) between 1 and 20 to construct a scale-free network. Second, the similarity matrix was converted into an adjacent matrix via the optimal  $\beta$  achieving 2 (Figure S3A). Furthermore, a negative correlation ( $R^2 = 0.83$ , slope = 1.6) between  $\log_{10}(k)$  and  $\log_{10}(p(k))$  was observed in Figure S3B, indicating that the transformed adjacent matrix was close to a scale-free network for further analysis. Third, the topological overlap matrix (TOM) was generated using the adjacent matrix for dynamic tree-clustering on genes (Figure S3C), thus identifying gene modules with HF. Genes recognized in each WGCNA-identified module can be found in Table S8. The correlation between identified gene modules and two traits (HF and control) was calculated. As shown in Figure 2A, the ME turquoise and ME blue modules showed significantly positive correlations to HF ( $R > 0.7$ ,  $p < 0.05$ ). We also observed the significantly distinct distribution of gene significance (GS) across modules ( $p < 0.05$ , Figure S3D). Fourth, we performed intramodular analysis and found significantly positive correlations between module membership (MM) and gene significance (GS) in the ME turquoise ( $R = 0.55$ ,  $p < 0.05$ ) (Figure 2B) and ME blue modules ( $R = 0.86$ ,  $p < 0.05$ ) (Figure 2C). Finally, we considered the genes in the ME turquoise module ( $n = 356$ ) and ME blue module ( $n = 298$ ) as HF-related genes for later analysis.



**Figure 1.** Identification of DEGs of HF from the training cohort with  $n = 366$  (200 HF and 166 non-HF controls) and KEGG enrichment analysis. (A) The plot of PLS-DA scores using two components shows the distinct separation between the HF and control groups. (B) The Volcano plot of DEGs with a significant change in  $|\log_2\text{FC}| > 1$  and  $\text{FDR} < 0.05$ . The red and blue dots represent up-regulated and down-regulated DEGs, respectively. (C) The clustering heatmap of DEGs among the HF group and control group. Left panel annotation: The significant KEGG enrichment pathways of up- and down-regulated DEGs (all  $q < 0.05$ ).



**Figure 2.** Crucial modules identified WGCNA and MEGENA from the training cohort with  $n = 366$  (200 HF and 166 non-HF controls). **(A)** The heatmap showed the relationship between five WGCNA-identified modules and two traits (HF and control). Each cell includes the correlation coefficient and  $p$  value. **(B)** Scatterplots of high correlations between GS versus MM for HF in blue module from WGCNA. **(C)** Scatterplots of high correlations between GS versus MM for HF in turquoise modules from WGCNA. **(D)** The gene co-expression network of HF via MEGENA. Node stands for distinct modules, and the sizes of node and label are proportional to gene numbers. The two modules with the largest genes in the network are shown: c1\_2 and c1\_6 (highlighted in red boxes). **(E)** Intergenic connectivity of genes in c1\_2 module. **(F)** Intergenic connectivity of genes in c1\_6 module.

### 3.3. Two Largest Crucial Modules of HF Were Identified by MEGENA on DEGs

To gain insight into a more biologically meaningful gene expression network in HF, we implemented MEGENA on the expression profiles of the 926 DEGs to establish a multiscale gene co-expression network. We first calculated and filtered the correlations between each gene pairwise of 926 DEGs. Subsequently, a fast PFN was calculated to filter and retain the significant correlation pairs. We then performed MCA on splitting PFN to identify sub-modules (summarized in Table S9). As shown in Figure 2D, 48 significant gene modules were identified (module significance  $p < 0.05$ ), of which the c1\_2 was the largest module of 288 genes, which was followed by the c1\_6 module containing 276 genes. Moreover, we displayed the subnetworks of the c1\_2 and c1\_6 modules. As shown in Figure 2E,F, the c1\_2 and c1\_6 modules comprised six and eight child modules, respectively.

### 3.4. A Large Set of Overlaps between Crucial Modules Showed Close Associations with HF-Related Biological Functions and Pathways

Comparing the two significantly positive gene modules of WGCNA with the two significantly large gene modules of MEGENA, we obtained two large sets of intersections between these identified modules. As shown in Figure 3A, 248 genes were overlapping between the ME turquoise of WGCNA and the c1\_6 module of MEGENA. We also observed that 234 genes intersected between ME blue of WGCNA and the c1\_2 module

of MEGENA. Enrichment analysis contributes to understanding whether the preservation of these modules between WGNCA and MEGENA could comprehensively represent the biological significance of HF. Subsequently, we conducted DO, GO, and KEGG enrichment analyses on the two sets of overlapping genes co-identified by WCGNA and MEGENA. DO enrichment analysis showed that the 234 genes largely participated in hematopoietic system disease, nutrition disease, autosomal recessive disease, overnutrition, and obesity (Figure 3B). GO enrichment analyses demonstrated that the 234 genes were highly associated with ECM, such as extracellular matrix structural constituent, collagen-containing extracellular matrix, and extracellular structure organization (Figure 3B). Furthermore, KEGG enrichment analysis revealed that 234 genes primarily participated in morphine addiction, Cushing syndrome, and phagosome (Figure 3B). DO enrichment analysis suggested that 248 genes were enriched in cardiac abnormality-related disorders, such as congestive heart failure, congenital heart disease, heart septal defect, and atrial heart septal defect (Figure 3C). It was also shown that 248 genes were mainly related to ECM-related functions, such as extracellular matrix structural constituent, collagen-containing extracellular matrix, and extracellular matrix organization. Notably, several pathways have been proven significant to HF progression, such as the cGMP-PKG signaling pathway [59], the cAMP signaling pathway [60], the Hippo signaling pathway [61], adrenergic signaling in cardiomyocytes [62], and the TGF-beta signaling pathway [63], were observed in the KEGG enrichment of 248 genes (Figure 3C). Taken together, we noticed that compared with 234 genes, 248 genes were enriched in more biological functions or pathways associated with HF, which was used as a gene set for further feature selection.

### 3.5. Machine Learning-Based Integration and PPI Network Analysis Screened 10 Pivotal Genes of HF

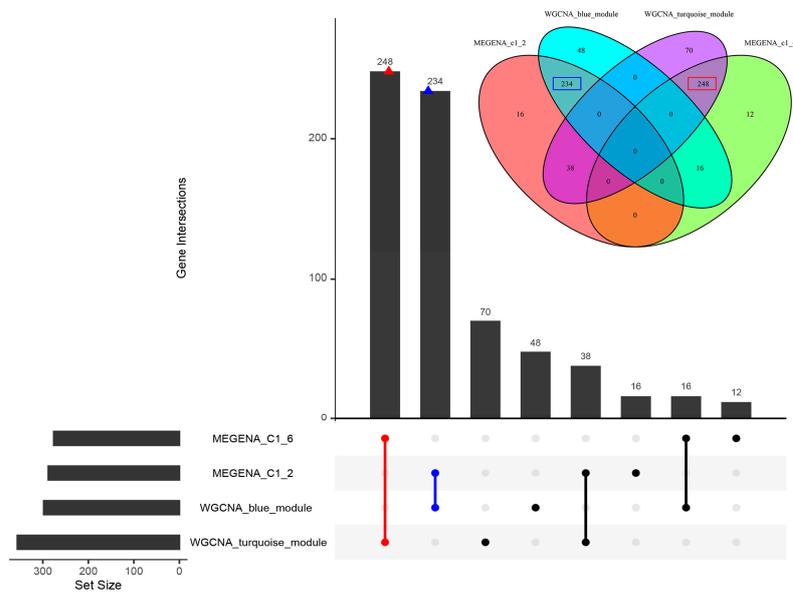
The expression profiles of 248 genes were subjected to the machine learning-based integration pipeline to filter pivotal genes of HF. Herein, we developed 113 predictive classification models fitted on the training and test cohorts via the CV framework. As shown in Figure 4A, the Enet ( $\alpha = 0.1$ ) with the highest mean AUC scores (0.971) performed within the training and test cohorts was considered the optimal classification model. Therefore, we obtained a set of 134 feature genes via the Enet ( $\alpha = 0.1$ ). To investigate the gene functional association, we subsequently constructed a 134 genes-dominant PPI network (Figure 4B). The genes showed interactions were retained in the network, and non-interactive genes were eliminated. Using the cytoHubba, we carried out the Degree, MNC, MCC, EPC, and DMNC algorithms to identify the network's top 20 genes with higher connectives (Figure S4). Finally, we obtained an interactive genes-dominant PPI network and identified 10 pivotal genes (*KIT*, *SLC6A1*, *SLC6A4*, *COL14A1*, *MME*, *OGN*, *SFRP4*, *CD1C*, *MFAP4*, *HTR2B*) overlapping between the five algorithms (Figure 4C).

### 3.6. *COL14A1*, *OGN*, *MFAP4*, and *SFRP4* Were Hub Genes as Candidate Diagnostic Biomarkers of HF

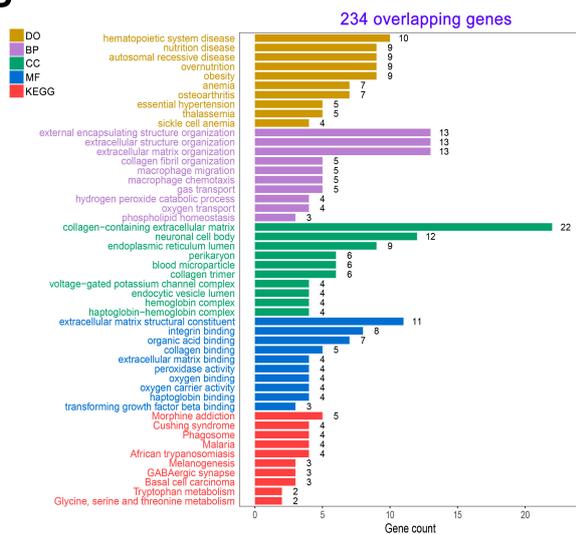
To assess the diagnostic abilities of 10 pivotal genes, we conducted ROC curve analysis in the training and test cohorts. The results demonstrated that *COL14A1*, *OGN*, *MFAP4*, and *SFRP4* achieved AUC scores all larger than 0.85 in the training cohort (Figures 5A and S5A) and test cohort (Figures 5D and S5B), suggesting the high sensitivity and specificity that distinguish HF from control. Accordingly, *COL14A1*, *OGN*, *MFAP4*, and *SFRP4* were considered hub genes of HF that could serve as underlying diagnostic biomarkers of HF. Additionally, combining four hub genes largely elevates the diagnostic performances within the training cohort (AUC = 0.988, Figure S5C) and test cohort (AUC = 0.935, Figure S5D). The diagnostic accuracy, correlations, and expression patterns of *COL14A1*, *OGN*, *MFAP4*, and *SFRP4* were further assessed within two datasets (GSE57338 and GSE42955) from the test cohort (Figure S6). Despite the different sample sources from the two datasets, robust diagnostic powers and elevated expressions of these genes were again verified, particularly *OGN* (all AUC > 0.9, Figure S6C,G). We subsequently compared the diagnostic performance of these four hub genes with current HF biomarkers, including myocardial stress/injury-related (model

1; BNP, NT-proBNP, TnT, TnI), neurohormone-related (model 2; MR-proADM, AVP, Peptide), myocardial remodeling-related (model 3; sST2, Gal-3, MMP, GDF-15), inflammatory-related (model 4; CRP, IL-6, TNF- $\alpha$ , PCT), and renal function injury-related (model 5; Cys C, NGAL) biomarkers. Importantly, the AUC scores of the four genes' combinations ranked first within the training cohort and test cohort (Figure S7), suggesting the superior prediction ability of our identified hub genes. The significant up-regulation patterns of four hub genes in HF were observed in the training cohort ( $p < 0.001$ , Figure 5B) and test cohort ( $p < 0.001$ , Figure 5E). Moreover, correlation analysis further revealed the close relationships between the expression levels of four hub genes. As shown in Figure 5C, *COL14A1* was significantly closely related to *OGN*, *SFRP4*, and *MEF4* ( $R > 0.6$ ,  $p < 0.001$ ). Additionally, *SFRP4* highly correlated to *MEF4* ( $R = 0.62$ ,  $p < 0.001$ ). The close expression correlation patterns of four hub genes were consistently verified in the test cohort (Figure 5F).

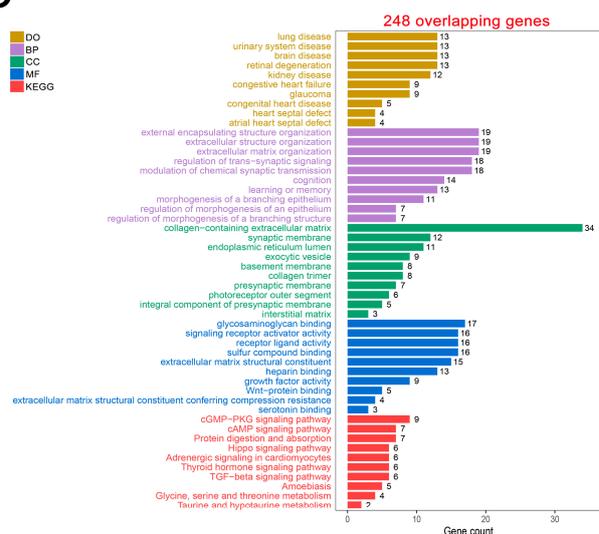
A



B



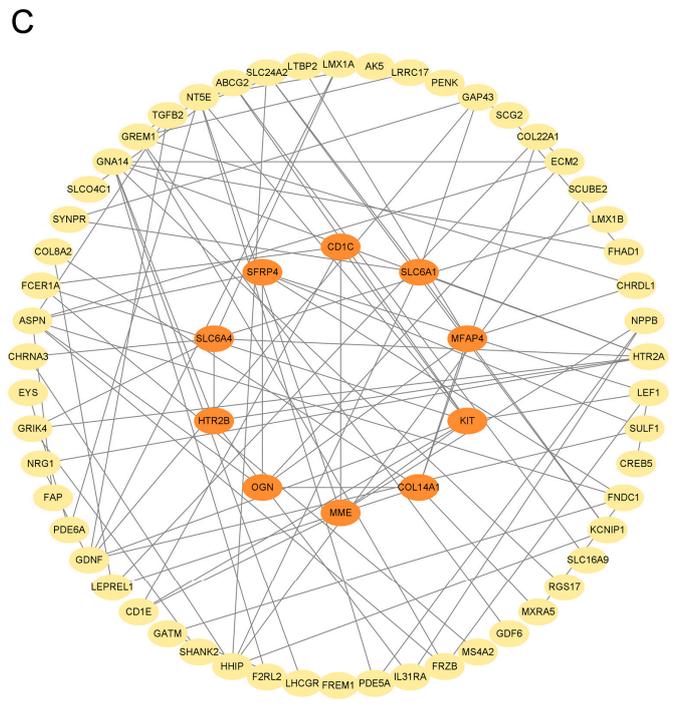
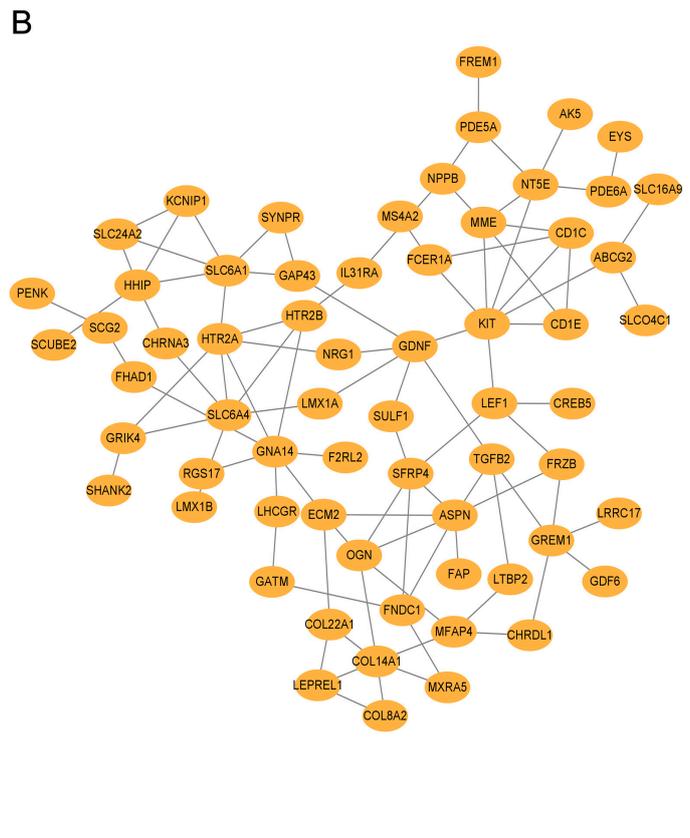
C



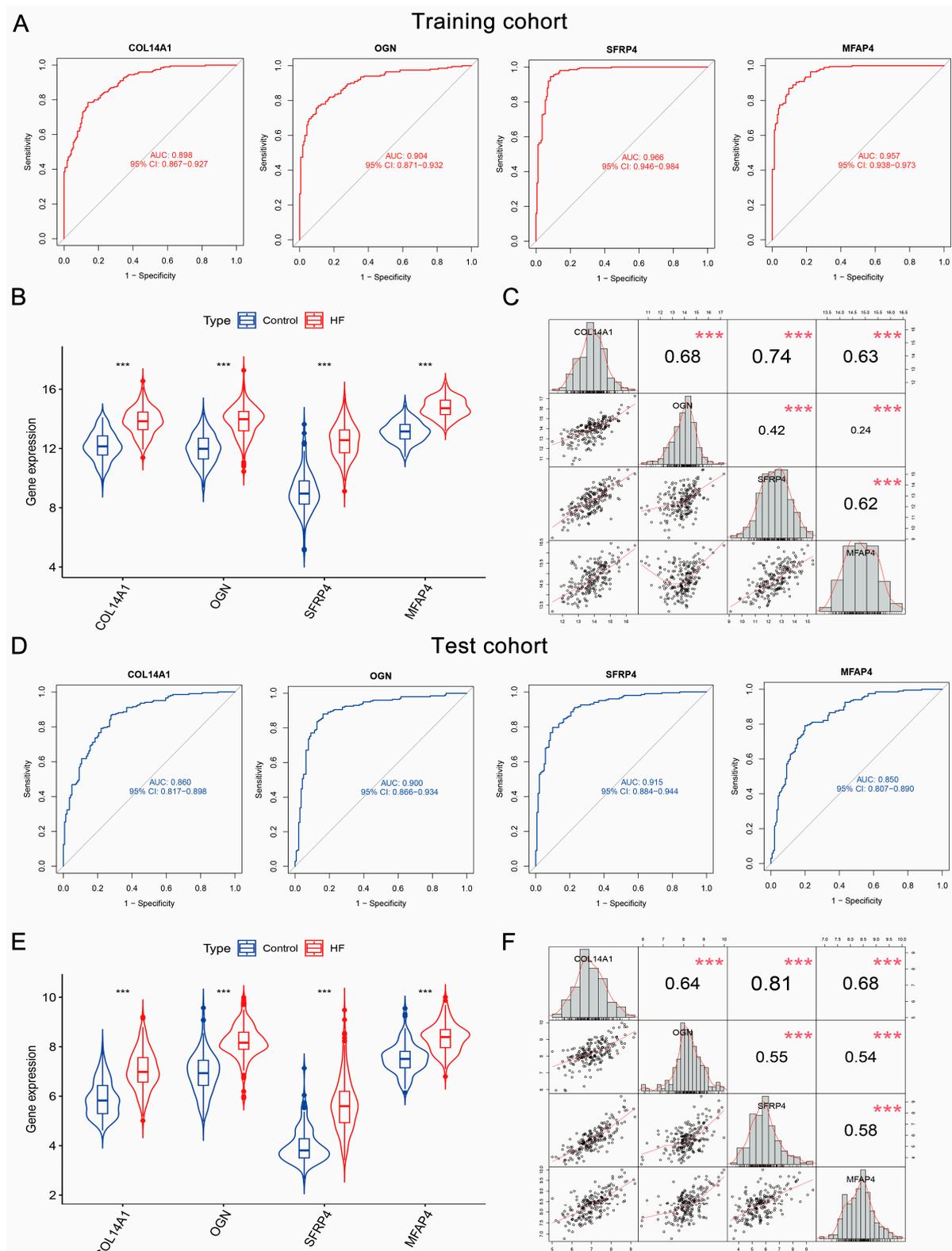
**Figure 3.** Overlap between crucial modules of WGCNA and MEGENA from the training cohort with  $n = 366$  (200 HF and 166 non-HF controls). (A) The Upset plot and Venn plot show two large sets of overlapping (indicated by red and blue triangles) between 4 selected modules, including 2 crucial modules of WGCNA and 2 crucial modules of MEGENA. (B) DO, GO, and KEGG enrichment analysis of shared genes between the blue module and c1\_2 module. (C) DO, GO, and KEGG enrichment analysis of shared genes between the turquoise module and c1\_6 module.

**A**

Model	1	2	3	4
Enet(alpha=0.1)	1.000	0.942	0.973	0.971
Enet(alpha=0.2)	1.000	0.933	0.971	0.968
Lasso+LDA	1.000	0.925	0.964	0.963
Ridge	1.000	0.906	0.978	0.962
Enet(alpha=0.3)	1.000	0.917	0.968	0.962
Enet(alpha=0.4)	1.000	0.917	0.965	0.961
Enet(alpha=0.5)	1.000	0.917	0.963	0.96
Enet(alpha=0.6)	1.000	0.917	0.962	0.959
LDA	1.000	0.933	0.940	0.958
RF+NaiveBayes	0.994	0.925	0.953	0.957
RF+Ridge	0.995	0.917	0.954	0.955
RF+LDA	0.996	0.917	0.952	0.955
RF+Enet(alpha=0.1)	0.996	0.917	0.952	0.955
RF+Enet(alpha=0.3)	0.997	0.917	0.951	0.955
RF+Enet(alpha=0.2)	0.997	0.917	0.951	0.955
glmBoost+LDA	1.000	0.992	0.964	0.952
Enet(alpha=0.7)	1.000	0.992	0.962	0.951
Enet(alpha=0.8)	1.000	0.992	0.961	0.951
Lasso+glmBoost	1.000	0.875	0.971	0.949
glmBoost	1.000	0.867	0.971	0.946
Enet(alpha=0.9)	1.000	0.867	0.962	0.943
glmBoost+Ridge	1.000	0.858	0.969	0.942
Stepglm[both]+glmBoost	0.998	0.875	0.948	0.94
Stepglm[backward]+glmBoost	0.998	0.875	0.948	0.94
Lasso	1.000	0.850	0.960	0.937
glmBoost+Enet(alpha=0.7)	1.000	0.858	0.945	0.934
glmBoost+Enet(alpha=0.9)	1.000	0.858	0.940	0.933
glmBoost+Lasso	1.000	0.858	0.939	0.932
glmBoost+Enet(alpha=0.8)	1.000	0.842	0.951	0.931
glmBoost+Enet(alpha=0.4)	1.000	0.842	0.950	0.93
glmBoost+Enet(alpha=0.3)	1.000	0.842	0.949	0.93
glmBoost+Enet(alpha=0.5)	1.000	0.842	0.945	0.93
glmBoost+Enet(alpha=0.1)	1.000	0.833	0.956	0.929
Stepglm[both]+NaiveBayes	0.994	0.850	0.945	0.929
Stepglm[backward]+NaiveBayes	0.994	0.850	0.945	0.929
glmBoost+Enet(alpha=0.6)	1.000	0.842	0.946	0.929
glmBoost+Enet(alpha=0.2)	1.000	0.833	0.954	0.929
Stepglm[both]+LDA	0.999	0.850	0.937	0.929
Stepglm[backward]+LDA	0.999	0.850	0.937	0.929
Stepglm[both]+Ridge	0.998	0.825	0.949	0.924
Stepglm[backward]+Ridge	0.998	0.825	0.949	0.924
Stepglm[both]+Enet(alpha=0.1)	1.000	0.825	0.931	0.919
Stepglm[both]+Enet(alpha=0.5)	1.000	0.825	0.929	0.918
Stepglm[backward]+Enet(alpha=0.1)	1.000	0.825	0.927	0.917
Stepglm[both]+Enet(alpha=0.7)	1.000	0.825	0.925	0.917
Stepglm[backward]+Enet(alpha=0.9)	1.000	0.825	0.923	0.916
Stepglm[both]+Enet(alpha=0.8)	1.000	0.825	0.923	0.916
Stepglm[both]+Enet(alpha=0.2)	1.000	0.825	0.923	0.916
Stepglm[backward]+Enet(alpha=0.5)	1.000	0.825	0.922	0.916
Stepglm[both]+Enet(alpha=0.4)	1.000	0.825	0.921	0.915
Stepglm[backward]+Enet(alpha=0.6)	1.000	0.825	0.921	0.915
Stepglm[backward]+Enet(alpha=0.8)	1.000	0.825	0.921	0.915
Stepglm[backward]+Enet(alpha=0.7)	1.000	0.825	0.918	0.914
Stepglm[backward]+Enet(alpha=0.4)	1.000	0.825	0.918	0.914
Stepglm[backward]+Lasso	1.000	0.825	0.920	0.915
Stepglm[both]+Enet(alpha=0.8)	1.000	0.825	0.919	0.915
Stepglm[backward]+Enet(alpha=0.7)	1.000	0.825	0.918	0.914
Stepglm[backward]+Enet(alpha=0.4)	1.000	0.825	0.918	0.914
Stepglm[backward]+Lasso	1.000	0.825	0.916	0.914
Stepglm[both]+Enet(alpha=0.9)	1.000	0.825	0.915	0.913
Stepglm[both]+Enet(alpha=0.3)	1.000	0.817	0.915	0.911
Stepglm[backward]+Enet(alpha=0.3)	1.000	0.817	0.915	0.911
glmBoost+GBM	1.000	0.500	0.515	0.672
Lasso+GBM	1.000	0.500	0.503	0.668
Lasso+Stepglm[both]	1.000	0.500	0.500	0.667
SVM	1.000	0.500	0.500	0.667
Lasso+SVM	1.000	0.500	0.500	0.667
Lasso+plsRglm	1.000	0.500	0.500	0.667
glmBoost+Stepglm[forward]	1.000	0.500	0.500	0.667
Lasso+Stepglm[forward]	1.000	0.500	0.500	0.667
Stepglm[forward]	1.000	0.500	0.500	0.667
Lasso+Stepglm[backward]	1.000	0.500	0.500	0.667
Stepglm[both]	1.000	0.500	0.500	0.667
Stepglm[backward]	1.000	0.500	0.500	0.667
glmBoost+Stepglm[both]	1.000	0.500	0.500	0.667
glmBoost+RF	1.000	0.500	0.500	0.667
Stepglm[both]+GBM	1.000	0.500	0.500	0.667
Stepglm[backward]+GBM	1.000	0.500	0.500	0.667
Stepglm[both]+RF	1.000	0.500	0.500	0.667
Stepglm[backward]+RF	1.000	0.500	0.500	0.667
XGBoost	1.000	0.500	0.500	0.667
Lasso+XGBoost	1.000	0.500	0.500	0.667
glmBoost+XGBoost	1.000	0.500	0.500	0.667
Stepglm[both]+XGBoost	1.000	0.500	0.500	0.667
plsRglm	1.000	0.500	0.500	0.667
glmBoost+plsRglm	1.000	0.500	0.500	0.667
Stepglm[backward]+XGBoost	1.000	0.500	0.500	0.667
Stepglm[both]+plsRglm	1.000	0.500	0.500	0.667
Stepglm[backward]+plsRglm	1.000	0.500	0.500	0.667
glmBoost+NaiveBayes	1.000	0.500	0.500	0.667
RF+GBM	1.000	0.500	0.500	0.667
Lasso+NaiveBayes	0.999	0.500	0.500	0.666
RF+XGBoost	0.999	0.500	0.500	0.666
glmBoost+SVM	0.998	0.500	0.500	0.666
RF+Stepglm[forward]	0.997	0.500	0.500	0.666
RF+plsRglm	0.997	0.500	0.500	0.666
Stepglm[both]+SVM	0.992	0.500	0.500	0.664
Stepglm[backward]+SVM	0.992	0.500	0.500	0.664
NaiveBayes	0.983	0.500	0.500	0.661
RF+SVM	0.969	0.500	0.500	0.656
Lasso+RF	1.000	0.479	0.461	0.647
RF	1.000	0.492	0.413	0.633
GBM	1.000	0.262	0.535	0.699



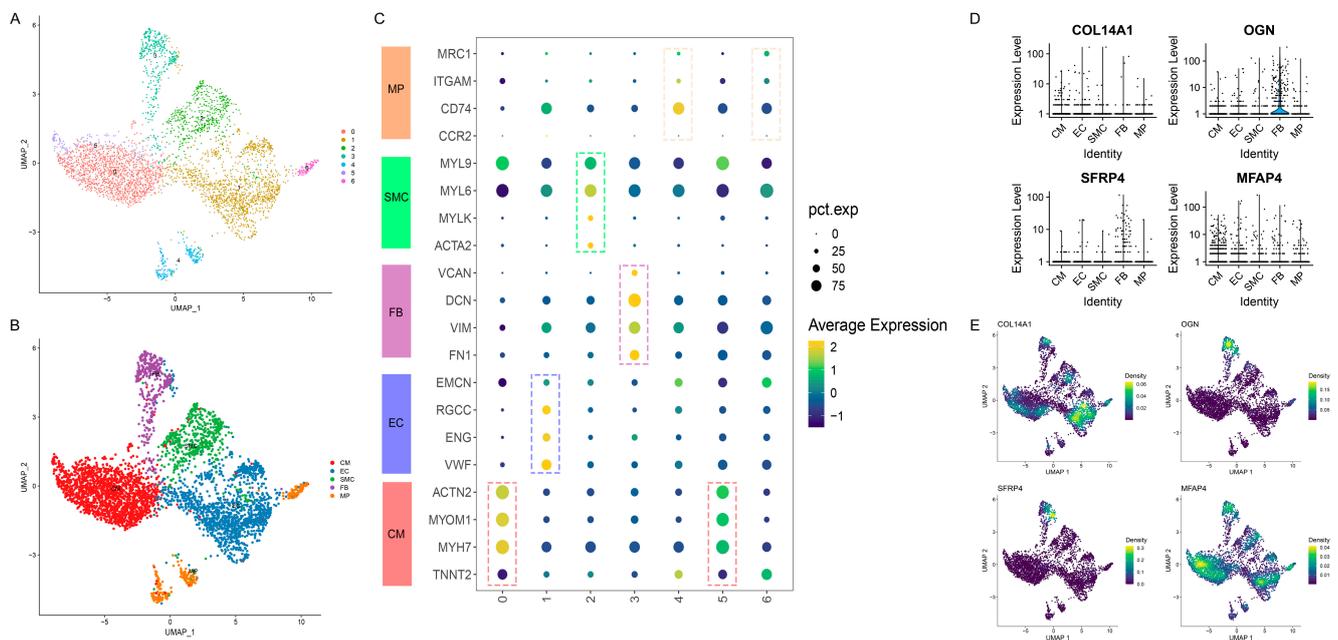
**Figure 4.** Pivotal genes were generated from the machine learning-based integration and PPI network. **(A)** A total of 113 machine learning-based prediction models performed feature selection on 210 genes in the training and test cohorts, and the AUC score of each model was then calculated. **(B)** PPI network of 134 feature genes. **(C)** The pivotal genes in the PPI network: Orange indicates 10 pivotal genes.



**Figure 5.** ROC curve analysis, expression levels, and correlations of four hub genes. (A) ROC curves of four hub genes in the training cohort. (B) Expression levels of four hub genes in HF compared to the control group in the training cohort. (C) Correlation heatmap of four hub genes expressed in HF in the training cohort. Red line indicates the fitted curve. (D) ROC curves of four hub genes in the test cohort. (E) Expression levels of four hub genes in HF compared to the control group in the test cohort. (F) Correlation heatmap of four hub genes expressed in HF in the test cohort. Red line (Figure 5C,F) indicates the fitted curve. \*\*\*,  $p < 0.001$ .

### 3.7. Distinct Cell-Specific Expression Patterns of Hub Genes in HF via scRNA-seq Analysis

To further explore the cell-specific expression of four hub genes in HF, we first conducted scRNA-seq analysis on GSE121893 to identify and characterize the cell populations in HF. The pre-processing procedure of scRNA-seq analysis, including quality control, data normalization, and dimensional reduction, is shown in Figure S8. Subsequently, the cells derived from six patients with HF were clustered into seven distinct clusters via the UMAP algorithm (Figure 6A). We then classified the seven clusters into five main cell types in HF (Figure 6B), including CMs, ECs, FBs, SMCs, and MPs, leveraging their specific molecular features (Figure 6C). Interestingly, we annotated clusters 4 and 6 as MPs, whereas distinct expression patterns of these clusters were observed. Cluster 4 is reflected with a higher *CD74*, *ITGAM*, and *MRC1* expression. However, their relatively lower expression was shown in cluster 6, suggesting that two cardiac MP populations are separately activated in the early stage and late stage. As depicted in Figure S9, we also evaluated the AUCell scores of the molecular features in the cell population, which notably differentiated five cell types. Next, we investigated the cell-specific expression patterns of four hub genes. The violin plots demonstrated that *OGN* and *SFRP4* were highly expressed in FBs (Figure 6D). However, the expression patterns of *COL14A1* and *MFAP4* showed no apparent differences among the five cell groups. Accordingly, the Nebulosa approach was implemented to recover the expression sparsity of hub genes from noisy signals in UMAP space. As shown in Figure 6E, we observed the clear expression patterns of four hub genes in five cell populations, in which *OGN* and *SFRP4* were mainly enriched in FBs, *COL14A1* in ECs, and *MFAP4* in CMs. Collectively, we uncovered the specific cellular compositions and expression patterns of hub genes in the context of HF.



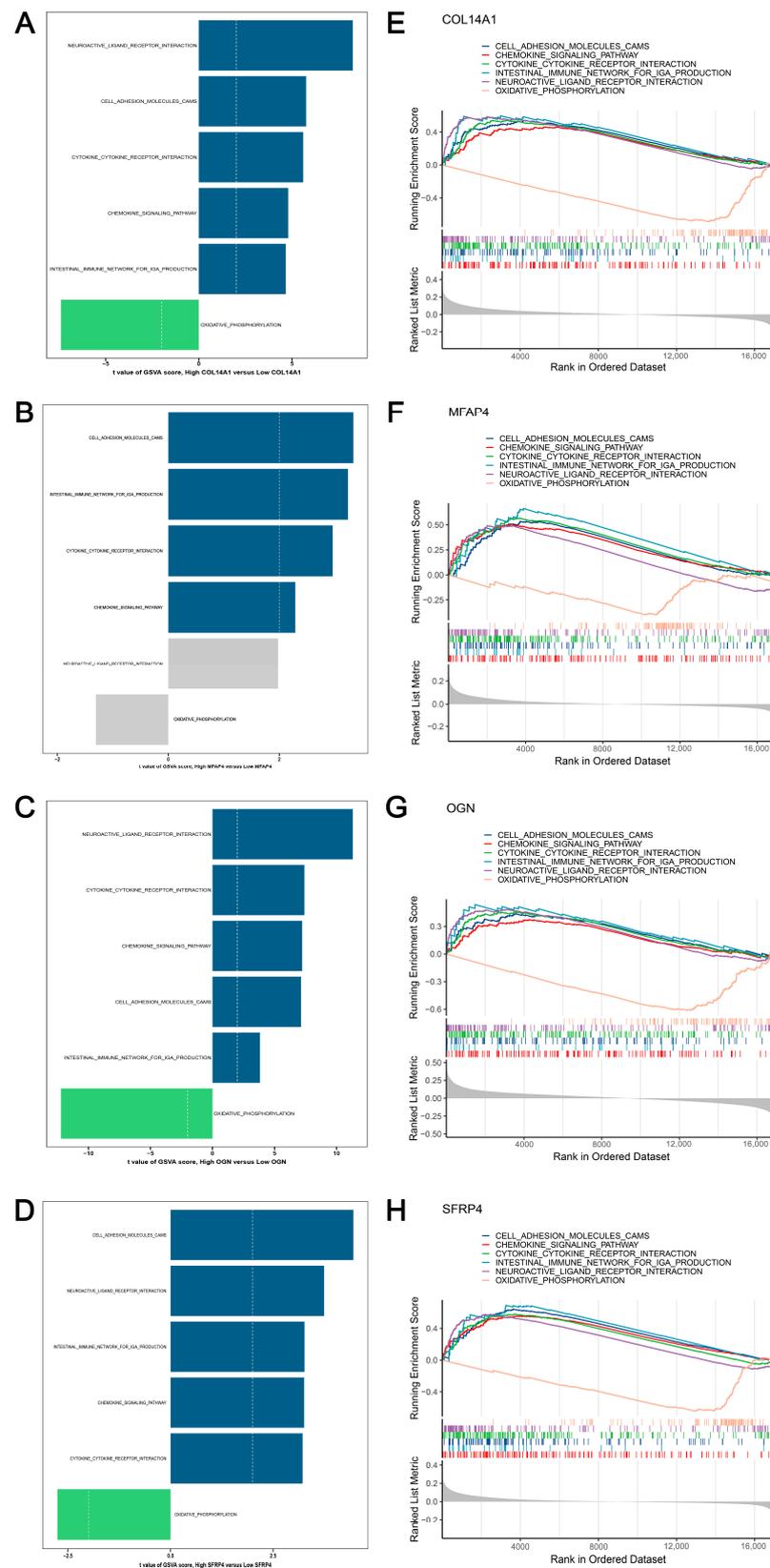
**Figure 6.** ScRNA-seq analysis uncovered the cell-specific expression patterns of hub genes in HF. (A) UMAP clustering of 4221 cells isolated from hearts under HF condition. (B) Annotation on clusters using specific gene markers of CM, EC, FB, SMC, and MP. (C) Seven clusters were annotated as five cell populations (CM, EC, FB, SMC, and MP). Dotted box indicates markers of five cell types. (D) Expression levels of hub genes in five cell populations. (E) Density estimation of hub genes in five cell populations. CM, cardiomyocyte; EC, endothelial cell; FB, fibroblast; SMC, smooth muscle cell; MP, macrophage.

### 3.8. GSVA and GSEA Revealed the Activated and Suppressed Pathways Co-Regulated by Hub Genes in HF

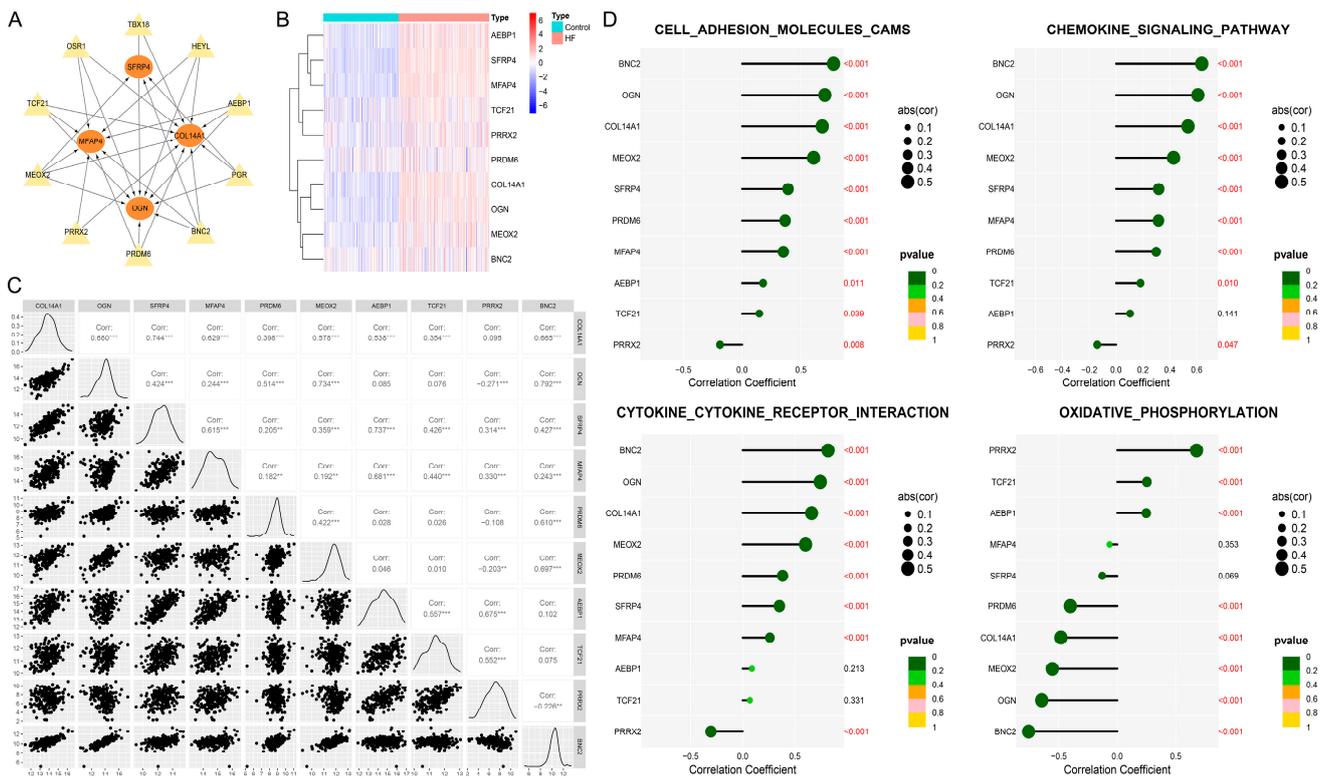
To reveal the underlying pathways co-regulated by four hub genes, we performed GSVA and GSEA using 186 annotated KEGG pathways on the expression profiles of HF. First, we implemented GSVA on the training cohort and compared the GSVA scores of pathways between the low expression and high expression of each hub gene, respectively. The significantly activated pathways ( $t > 2$ ,  $p < 0.05$ ) and suppressed pathways ( $t < 2$ ,  $p < 0.05$ ) via GSVA were shown in Figure 7A–D and Figure S10A–D. Second, GSEA was performed on each gene within the training and test cohorts. As shown in Figure 7E–H, we observed that the chemokine signaling pathway, intestinal immune network for IgA production, cell adhesion molecules, cytokine–cytokine receptor interaction, and neuroactive ligand receptor interaction were significantly activated ( $NES > 1$ ,  $p < 0.05$ ,  $FDR < 0.25$ ) in the training cohort. And oxidative phosphorylation was significantly inactivated ( $NES < 1$ ,  $p < 0.05$ ,  $FDR < 0.25$ ) in the training cohort. These six significant GSEA-enriched pathways of hub genes in the training cohort were also validated in the test cohort (Figure S10). Third, we took the intersection between GSVA and GSEA of each hub gene (Figure S11). After comprehensively comparing the results from GSVA and GSEA, we suggested that cell adhesion molecules, the chemokine signaling pathway, cytokine–cytokine receptor interaction, and oxidative phosphorylation might be downstream mechanisms co-regulated by four hub genes in HF progression. Additionally, we implemented GSVA and GSEA analyses of these shared mechanisms in HF compared to control cases. Importantly, the evident activation of cell adhesion molecules, the chemokine signaling pathway, and cytokine–cytokine receptor interaction was observed in HF, but oxidative phosphorylation was inhibited (Figure S12). We subsequently conducted BN inference on the pathways mentioned above, and the predicted GRNs were visualized in Figure S13. The four hub genes showed higher expressions, more edges, and strengths in GRNs, suggesting their critical roles in regulating these pathways.

### 3.9. *BNC2* and *MEOX2* Were TFs Actively Targeting Hub Genes

On the basis of the TRRUST v2 database, we enrolled the top 10 TFs (*OSR1*, *PRDM6*, *AEBP1*, *TBX18*, *MEOX2*, *BNC2*, *HEYL*, *TCF21*, *PRRX2*, and *PGR*) targeting hub genes (Table S10). A total of six TFs (*AEBP1*, *TCF21*, *PRRX2*, *PRDM6*, *BNC2*, and *MEOX2*) that increased in HF within the training and test cohorts (Figures S14 and S15) were regarded as key TFs that were activated in HF progression. The regulatory network involving four hub genes and six key TFs is shown in Figure 8A. Notably, *BNC2* and *MEOX2* showed regulatory interactions with four hub genes. The clustering heatmap then demonstrated the differential expression patterns of six key TFs and four hub genes in the HF and control groups (Figure 8B), in which the expressions of *AEBP1*, *TCF21*, and *PRRX2* were similar to those of *SFRP4* and *MFAP4*. *PRDM6*, *BNC2*, and *MEOX2* showed expression patterns similar to those of *COL14A1* and *OGN*. Figure 8C further showed the correlation patterns of 10 key TFs and four hub genes in HF, in which *BNC2* and *MEOX2* were significantly correlated to hub genes. We then quantified the enrichment levels of cell adhesion molecules, the chemokine signaling pathway, cytokine–cytokine receptor interaction, and oxidative phosphorylation in the HF group using ssGSEA. To evaluate whether the TFs could actively or negatively affect the hub genes-regulated pathways, we further related the expressions of key TFs to the ssGSEA enrichment scores of pathways. Figure 8D illustrates the significant positive relationships between two TFs (*BNC2*, *MEOX2*) and three pathways (including cell adhesion molecules, the chemokine signaling pathway, and cytokine–cytokine receptor interaction). However, oxidative phosphorylation was negatively correlated to *BNC2* and *MEOX2*, which indicated that these TFs inhibited oxidative phosphorylation. Notably, *PRRX2* exhibited a distinct correlation pattern compared to other key TFs, which was positively correlated with oxidative phosphorylation but negatively related to cell adhesion molecules, the chemokine signaling pathway, and cytokine–cytokine receptor interaction, suggesting that *PRRX2* might be a TF that suppressed hub genes in HF.



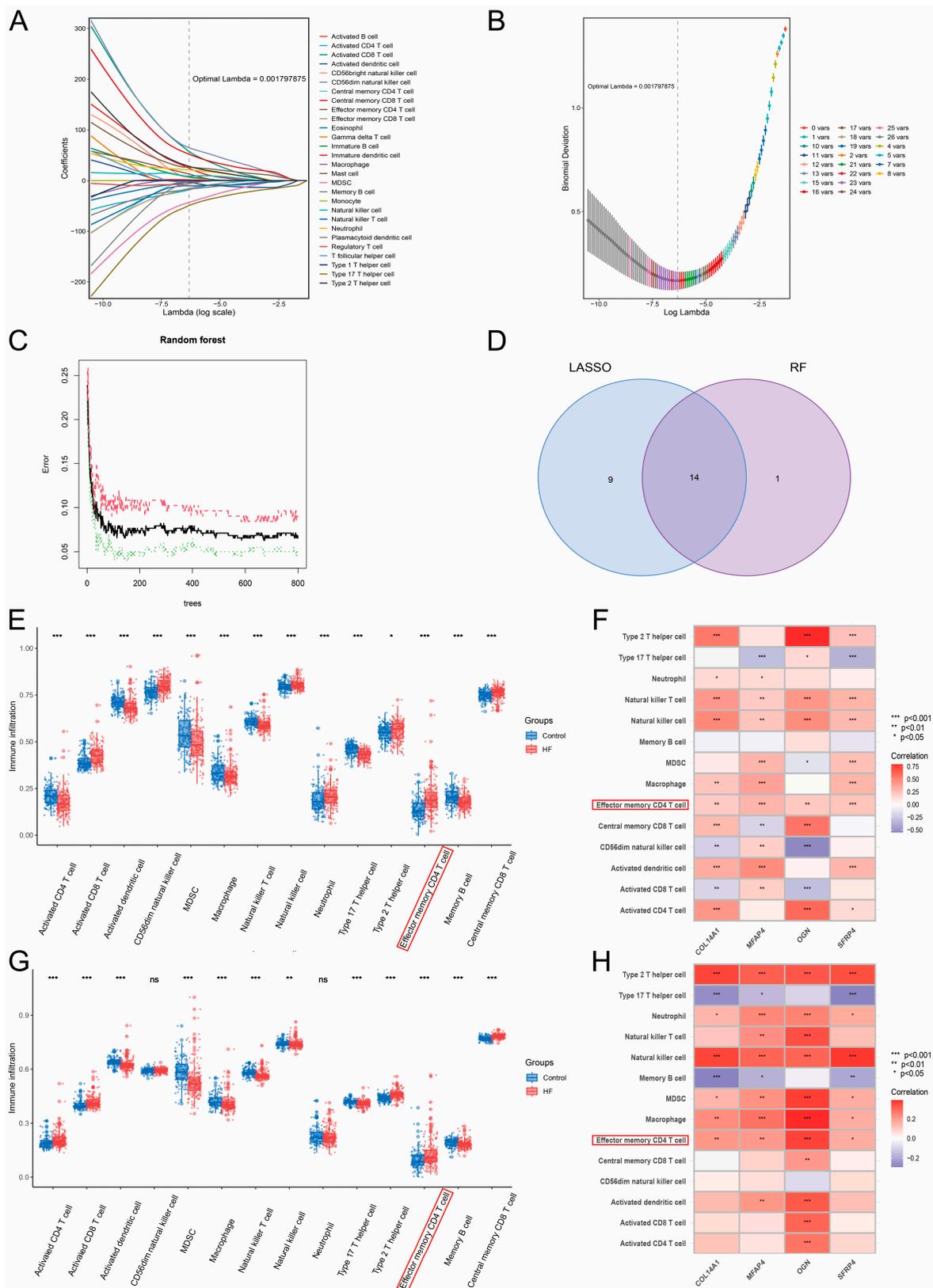
**Figure 7.** GSEA and GSA of low and high expressions of hub genes involved in HF. **(A)** GSA of *COL14A1*. **(B)** GSA of *MFAP4*. **(C)** GSA of *OGN*. **(D)** GSA of *SFRP4*. **(E)** Selected top KEGG terms using GSEA of *COL14A1*. **(F)** Selected top KEGG terms using GSEA of *MFAP4*. **(G)** Selected top KEGG terms using GSEA of *OGN*. **(H)** Selected top KEGG terms using GSEA of *SFRP4*. Dotted line **(A–D)** suggests that the t value is greater than or less than 2.



**Figure 8.** Interactions between TFs and hub genes-related pathways. (A) Hub gene–TF regulatory network: Circle symbolizes hub genes and triangle for TFs. (B) Heatmap showing the expression patterns of TFs and hub genes in HF and control groups. (C) Correlation heatmap of TFs and hub genes expressed in HF. (D) Lollipop plots showing the correlations between TF expression levels and ssGSEA scores of hub genes-related pathways. \*\*,  $p < 0.01$ ; \*\*\*,  $p < 0.001$ .

### 3.10. Elevated Infiltration Levels of Effector Memory CD4+ T Cells Were Highly Related to Hub Genes

In the previous analysis, we observed that the expressions of hub genes may co-regulate two significantly activated immune-related pathways in HF, including the chemokine signaling pathway and cytokine–cytokine receptor interaction. Accordingly, we wondered whether the immune landscape altered and could further exert certain functions in HF. First, we performed ssGSEA to quantify the infiltration levels of 28 immune cell types between the HF and control groups within the training and test cohorts. We then carried out Lasso and RF to perform feature selection on the infiltration profiles of 28 immune cell types to identify the key types, respectively. Under the optimal  $\lambda$  (0.001797875), Lasso obtained 23 key immune cell types (Table S11) with regression coefficients larger than zero (Figure 9A,B). RF constructed with 586 decision trees retained 15 immune cell types (Table S12) with a mean decrease in Gini larger than zero (Figure 9C). A set of 14 immune cell types (activated CD4+ T cell, activated CD8+ T cell, activated dendritic cell, CD56dim natural killer cell, MDSC, macrophage, natural killer T cell, natural killer cell, neutrophil, Type 17 T helper cell, Type 2 T helper cell, effector memory CD4+ T cell, memory B cell, central memory CD8+ T cell) overlapping between Lasso and RF was considered to contain the key immune cell types of HF. Within the training and test cohorts, we demonstrated the elevated infiltrations of activated CD8+ T cells, Type 2 T helper cells, and effector memory CD4+ T cells in HF (Figure 9D,F). We also noticed that the infiltrations of natural killer cells and effector memory CD4+ T cells positively related to expressions of hub genes in the training and test cohorts (Figure 9E,G). Effector memory CD4+ T cells showed significantly up-regulated infiltrations and positive correlations with hub genes in HF, suggesting its crucial role involved in immune-related pathways regulated by hub genes.



**Figure 9.** Exploring the immune landscape of HF and hub gene-related immune infiltration cells. (A) Lasso coefficient profiles of 28 immune infiltration cells. (B) Selection of optimal penalization

coefficient ( $\lambda$ ) under 10-fold cross-validation. (C) The effect of tree number on error variation of RF. Green and red stand for the error rate of the control and HF groups, respectively, and black represents the overall error rate. (D) Integration of Lasso and RF identified 14 key immune infiltration cells of HF. (E) Comparisons of 14 key immune infiltration cells between the HF and control groups in the training cohort. (F) Correlation analysis among 14 key immune infiltration cells and hub genes in the training cohort. (G) Comparisons of 14 key immune infiltration cells between the HF and control groups in the test cohort. (H) Correlation analysis among 14 key immune infiltration cells and hub genes in the test cohort. Effector memory CD4+ T cells are highlighted in the red frame, indicating that the infiltration of effector memory CD4+ T cells elevates in HF and is positively correlated to hub genes. ns, not statistically significant; \*,  $p < 0.05$ ; \*\*,  $p < 0.01$ ; \*\*\*,  $p < 0.001$ .

### 3.11. Small-Molecule Agents Targeting Active TFs and Hub Genes Could Serve as Candidate Drugs for Alleviating HF

A set of 78 small-molecule agents potentially targeting two key TFs or four hub genes was identified via the Enrichr database by accessing DSigDB (Table S13). We then identified captopril (targeting *BNC2*), aldosterone (targeting *MEOX2*), cyclopenthiiazide (targeting *MEOX2*), estradiol (targeting *COL14A1*), tolazoline (targeting *COL14A1*), and genistein (targeting *SFRP4*) previously have been reported to treat or alleviate HF [64–69], as candidate therapeutic agents for HF. Notably, captopril and cyclopenthiiazide belong to Angiotensin-Converting Enzyme Inhibitors (ACEIs) and thiazide diuretics, which have been put into clinical treatment of HF. Moreover, the therapeutic value of aldosterone antagonists in curving HF has been underscored recently. Given the protective effects of estrogens and phytoestrogens for alleviating HF, studies from animal models have demonstrated that genistein and estradiol could attenuate pathological hypertension in the progression of HF, suggesting their HF-related preventative roles. AutoDock Vina software (version 1.1.2) was subsequently performed to dock the six selected small-molecule agents with their targeted TFs (*BNC2*, *MEOX2*) or hub genes (*COL14A1*, *SFRP4*). Furthermore, we visualized the docking models with the lowest binding affinity and Root Mean Squared Error (RMSE) lower than 2 Å through PyMOL (Figure S16). The formation of hydrogen bonds that small-molecule agents bind to the amino acids of TFs or hub genes was also analyzed (Table 2, Figure S16), thus revealing the specific affinity patterns.

**Table 2.** Molecular docking results of drugs with targets.

Drugs	Targets	Affinity (kcal/mol)	Bond	Protein's Residues
Captopril	<i>BNC2</i>	−3.9	H-bond	GLN-199
Aldosterone	<i>MEOX2</i>	−7.1	H-bond	GLU-203
Cyclopenthiiazide	<i>MEOX2</i>	−5.7	H-bond	LYS-241
				THR-192
Estradiol	<i>COL14A1</i>	−8.2	H-bond	ASN-237
				ARG-171
Tolazoline	<i>COL14A1</i>	−5.9	H-bond	VAL-930
Genistein	<i>SFRP4</i>	−3.8	H-bond	ASP-1148
				GLU-127
				ARG-262

### 3.12. Plasma OGN Elevated in HF with Robust Diagnostic Value and Positive Causal Correlation to HF Risk

To study whether the four hub genes could serve as plasma biomarkers for diagnosing HF, we conducted qRT-PCR analysis on plasma collected from three HF patients and three non-HF controls. As shown in Figure 10A, OGN was highly expressed in HF, elevating about 1.5 times compared to the control. Additionally, OGN demonstrated an extremely robust diagnostic value (AUC = 1) in distinguishing HF from control (Figure 10B). A positive correlation between OGN and NT-proBNP ( $R = 0.95$ ,  $p = 0.0036$ ) is shown in Figure 10C. Leveraging GSE135055 as an external validation cohort, we noted that in comparison with

the control, the expressions of plasma *OGN* in patients with HF also significantly increased (Figure 10D). Moreover, the robust diagnostic power of *OGN* was again demonstrated (AUC = 0.852, Figure 10E). Meanwhile, no significant relationship between *OGN* and EF was observed ( $p = 0.83$ , Figure 10F). Figure 10G further showed the expression patterns of *OGN* among diverse groups of clinical characteristics, including HF type, gender, smoking history, hypertension, and age. Interestingly, we found that the expression of *OGN* may not fluctuate through these clinical factors (all  $p > 0.05$ ), which were subsequently validated in the training and test cohorts (Figure S17). To investigate the potential of plasma *OGN* as a diagnostic biomarker benefiting from genetic variation, we performed a two-sample MR analysis based on large-scale GWAS data. We initially integrated the exposure data (plasma *OGN*) and outcome data (HF risk) to filter the overlapping significant SNPs. Following the three basic assumptions of the two-sample MR approach, we selected 25 significant SNPs as strong IVs to implement the analysis. As depicted in Figure 10H, we related the effect sizes of SNPs on plasma *OGN* and SNPs on HF, and there was an overall positive fitted relationship, suggesting the direct casual association between plasma *OGN* and HF. The forest plot and funnel plot of each SNP on HF can be found in Figure S18A,B, respectively. Using the IVW approach, which showed the highest statistical effectiveness among the five MR methods, we demonstrated that plasma *OGN* was positively correlated to the risk of HF with an estimated OR of 1.12805 (95% CI: 1.05431–1.20694,  $p < 0.05$ , Figure 10I). The positive causal relationship between plasma *OGN* and HF was also estimated via the other four MR approaches (Figure 10H), showing a consistent directional change (all OR > 1). Thus, we herein speculated that a higher concentration of plasma *OGN* corresponds to an increased risk of HF. To guarantee the reliability of this result, we conducted a series of sensitivity tests (heterogeneity, horizontal pleiotropy, and leave-one-out analyses). No heterogeneity and horizontal pleiotropy were observed (all  $p > 0.05$ , Table S14). In the leave-one-out analysis, the step-by-step removal of each SNP was not recognized to greatly influence the result (all SNP > 0, Figure S18C), indicating all incorporated SNPs achieved significant causality rather than the occurrence of a dominate SNP. In a word, these data suggested the great potential of *OGN* as a plasma biomarker for HF diagnosis and the evaluation of HF progression.

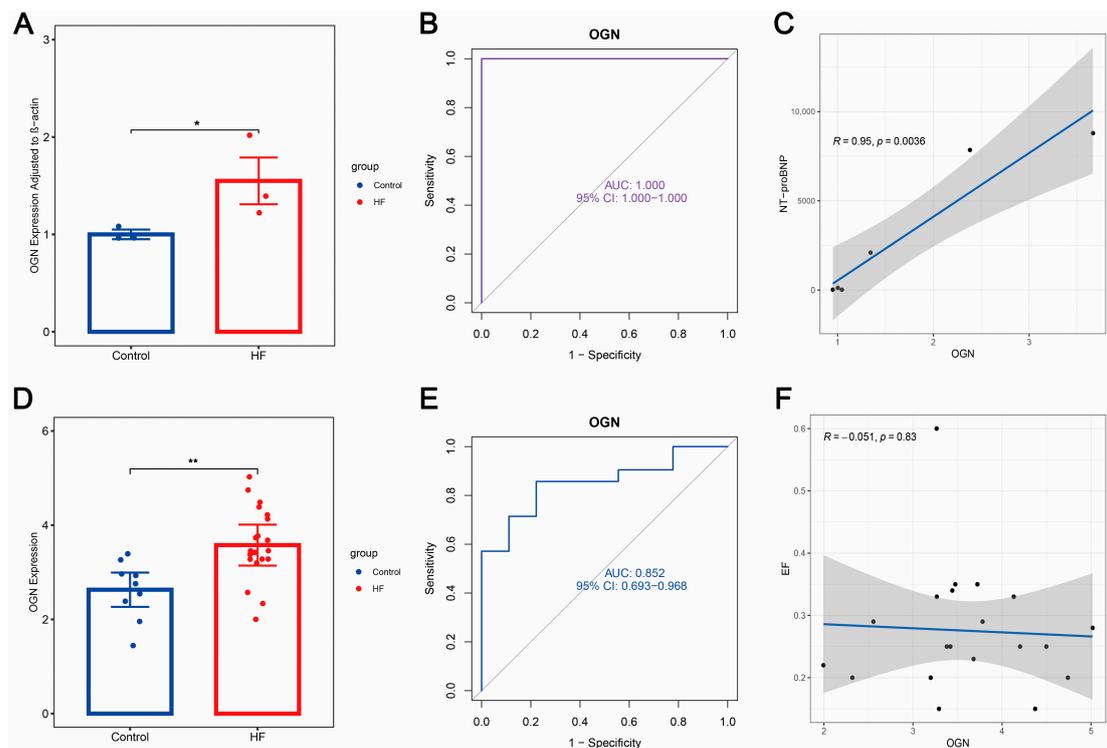
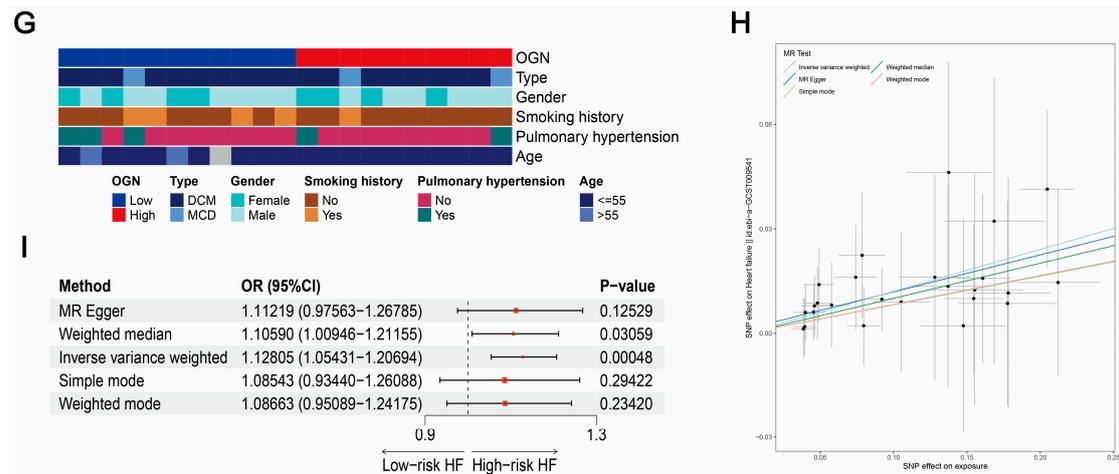


Figure 10. Cont.

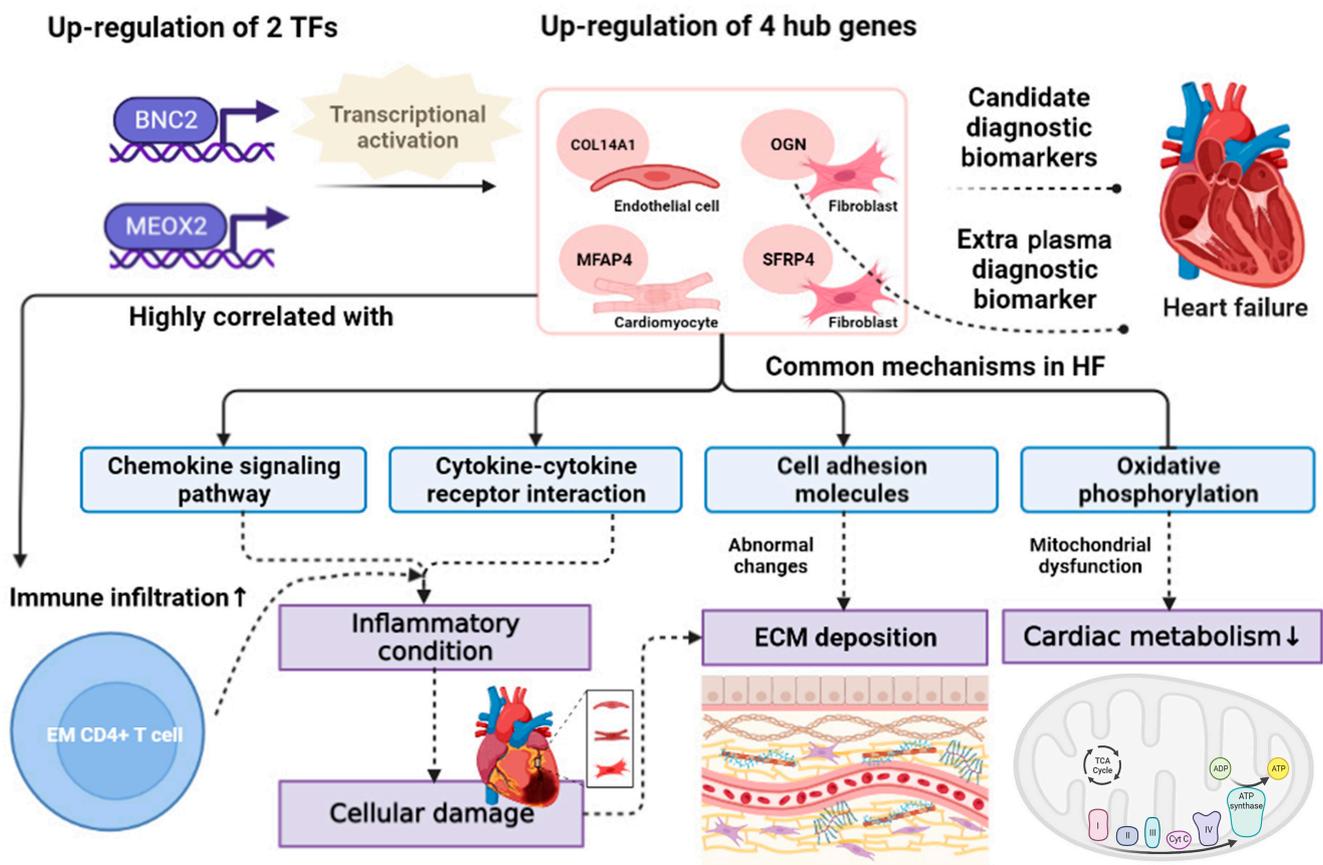


**Figure 10.** Evaluation of plasma *OGN* as a promising diagnostic biomarker of HF. (A) The qPCR-detected mRNA expression level of *OGN* (normalized by  $\beta$ -actin) in plasma of HF patients compared to control cases. (B) ROC curve of *OGN*. (C) Correlation between expressions of *OGN* and NT-proBNP. (D) The expressions of *OGN* in HF patients compared to control cases in GSE135055. (E) ROC curve of *OGN* in GSE135055. (F) Correlation between expressions of *OGN* and EF in GSE135055. (G) Heatmap displaying the expression patterns of *OGN* in groups of different clinical characteristics in GSE135055. (H) A scatterplot shows the positive association between the SNP effect on plasma *OGN* (x-axis) and HF (y-axis). (I) Forest plot indicating the estimated OR effects and *p* values generated from using five approaches on assessing the causal effect of plasma *OGN* on the risk of AMI. The size of red box indicates standard error. \*, *p* < 0.05; \*\*, *p* < 0.01.

#### 4. Discussion

HF remains high in incidence and mortality worldwide, and a novel approach is urgently needed to improve this situation. In the past few decades, HF has been proven to be a heterogeneous and progressive clinical syndrome that is reflected in its complex pathophysiology [6].

Herein, we performed a comprehensive bioinformatics analysis on bulk RNA-seq and sc-RNA seq datasets from HF patients and non-HF controls. We briefly outlined our main findings shown in a schematic diagram (Figure 11). First, we identified 926 distinct DEGs (HF versus control), including 648 up-regulated genes and 278 down-regulated genes in HF. The KEGG pathway enrichment analysis then showed that DEGs were mainly involved in ECM and immune-related pathways (cytokine–cytokine receptor interaction, cell adhesion molecules, the T cell receptor signaling pathway, ECM–receptor interaction, complement and coagulation cascades). Impaired signaling through corresponding cytokine receptors accelerates myocardial apoptosis and tissue damage after acute cardiac stress, suggesting an adverse role of cytokine–cytokine receptor interactions in HF [70]. Furthermore, cell adhesion molecules play a crucial role in modulating cardiac inflammation and pathological cardiac remodeling by facilitating the recruitment of T-cells to the left ventricle, ultimately contributing to cardiac dysfunction of HF [71]. Excessive extracellular matrix remodeling has extensively been known to precipitate HF. Coagulation proteases orchestrate clotting-independent activation of protease-activated receptors (PARs) to modulate TGF- $\beta$ 1 signaling, influencing cardiac fibrosis in HF [72]. Notably, a large body of evidence has demonstrated the abnormal ECM and immune-related functions or mechanisms altered in HF development and progression.



**Figure 11.** Schematic diagram illustrating the speculative pathogenic mechanisms of four hub genes that participated in HF.

Gene co-expression analysis is commonly used to identify gene modules linked to disease traits [73]. Next, we performed WGCNA and MEGENA gene co-expression analyses on DEGs. Leveraging the gene-pairwise correlation and agglomerative clustering, WGCNA allows the assignment of each gene to certain modules, which helps recognize the modules highly correlated to disease traits. We initially identified two crucial modules from WGCNA, including ME turquoise and ME blue modules. Of note, these two modules all displayed significant strong correlations to HF ( $R > 0.7$ ,  $p < 0.05$ ), and genes recognized in these modules showed positive associations between GS and MM. These results suggest that these two modules exhibited a strong linkage with HF, which could be used as functional gene sets of HF. Contrary to the traditional clustering by WGCNA, MEGENA is implemented based on divisive hierarchical clustering [74]. Therefore, more compact and functionally coherent modules could be screened out through MEGENA, and we recognized two large gene modules from MEGENA, including c1\_2 and c1\_6 modules. WGCNA and MEGENA could complement each other [74] and comprehensively capture gene modules strongly related to HF rather than rely on a single method. Four gene modules were obtained, including two WGCNA-identified gene modules (ME turquoise and ME blue modules) and two MEGENA-identified gene modules (c1\_2 and c1\_6 modules). More interestingly, we obtained two large intersections between the four identified gene modules, which were 248 genes shared by the ME turquoise and c1\_6 modules and 234 genes common to the ME blue and c1\_2 modules. Although WGCNA and MEGENA successfully seized these two preserved co-expression gene patterns associated with HF, their biological significance of HF remained uncertain. Therefore, we implemented enrichment analysis to annotate human disease, gene functions, and pathways to determine whether the HF-related terms are over-enriched or under-enriched in the two overlapping gene sets. Here, we selected gene sets with more HF-representative terms to guarantee

the algorithmic and biological capture of HF characteristics. The DO, GO, and KEGG enrichment analyses demonstrated that 248 genes were mainly involved in diverse cardiovascular disorders and HF-related biological functions or pathways. In comparison, fewer HF-related enriched terms were shown in 234 genes. Consequently, we retained 248 genes and performed feature selection via a machine learning-based integration pipeline under the LOOCV framework. A total of 134 feature genes were selected using the Enet ( $\alpha = 0.1$ ) model with the highest AUC values in the training and test cohorts. Together with the PPI network establishment and five ranking algorithms of Degree, MNC, MCC, EPC, and DMNC, we further obtained 10 pivotal genes, including *KIT*, *SLC6A1*, *SLC6A4*, *COL14A1*, *MME*, *OGN*, *SFRP4*, *CD1C*, *MFAP4*, *HTR2B*. Most of these pivotal genes' activations have been reported to contribute to HF, as summarized in Table S15. ROC curve analysis then identified four key genes (*COL14A1*, *OGN*, *SFRP4*, *MFAP4*) performed with robust AUC scores ( $AUC > 0.85$ ) in both the training and test cohorts. The up-regulation patterns of four key genes in HF and their positive correlation patterns were observed within the training and test cohorts. Leveraging sc-RNA seq analysis, we uncovered the cellular composition of failed hearts, mainly including cardiomyocytes, endothelial cells, fibroblasts, smooth muscle cells, and macrophages. Moreover, distinct cell-specific expression patterns were revealed in which *OGN* and *SFRP4* were mainly expressed in cardiac fibroblasts, *COL14A1* was mainly expressed in cardiac endothelial cells, and *MFAP4* was expressed in cardiomyocytes.

Of these four key genes, Collagen Type XIV Alpha 1 Chain (*COL14A1*) has previously been reported to affect arterial remodeling [75] and showed up-regulation mRNA and protein levels in the ischemic heart of HF patients [76]. *Col14a1* deficiency would contribute to defective ventricular morphogenesis and chaotic collagen fibril organization in mice, thereby indicating its maintenance of structural integrity during early heart development [77]. Given the specific expression pattern of *COL14A1* in endothelial cells uncovered by sc-RNA analysis, we speculate that the close association of *COL14A1*-activated endothelial cells and dysfunction in collagen constituents contributes to myocardial hypertrophy of HF progression. Osteoglycin (*OGN*), which encoded a class III small leucine-rich proteoglycan (SLRP) member, was first discovered to act as a key regulator in rat left ventricular mass (LVM) through modulating the TGF- $\beta$  pathway [78]. It was also proposed that *OGN* might be involved in the regulation of cardiac dysfunction and adverse remodeling after myocardial infarction in HF [79]. Moreover, *OGN* was reported to modulate fibrosis and inflammation, resulting in diastolic dysfunction shown in hypertensive heart disease [80]. More recently, Fang et al. have revealed the specific mechanism of *Ogn* in myocarditis mice by which its silencing suppressed the Wnt signaling pathway, thus inhibiting myocardial fibrosis proliferation [81]. The close linkage between activated cardiac fibroblast and adverse outcomes following HF (fibrosis, remodeling, and dysfunction) has been extensively expounded [82]. Therefore, it is reasonable to deduce the expression of *OGN* active in cardiac fibroblasts in the context of HF, leading to myocardial fibrosis and remodeling post-HF. Microfibril-associated glycoprotein 4 (*MFAP4*) has been proven to regulate calcium-dependent cell adhesion or intercellular interactions, leading to inflammation and fibrosis, which was previously verified to engage in remodeling-related diseases such as liver fibrosis, atherosclerosis, and arterial injury stimulated remodeling [83,84]. Apart from the close association between *MFAP4* and remodeling-related disorders, it was also shown that *MFAP4* could also serve as a candidate biomarker for cardiovascular diseases such as HF [84]. For instance, clinical cohort-based research demonstrated that the protein levels of *MFAP4* were significantly increased in serum derived from HF patients compared to control cases [85]. Mechanistically, *Mfap4* deletion could attenuate cardiac fibrosis and ventricular arrhythmias, suggesting its potential as a therapeutic target in HF prevention [86]. Additionally, Dorn et al. showed that *Mfap4* knockout affected pressure overload-induced cardiac remodeling, leading to elevated cardiac hypertrophy and exacerbating cardiac function [87]. According to our previous analysis, *MFAP4* was identified to be specifically expressed in cardiomyocytes, which largely implies its significant activation

in directly exacerbating hypertrophy of cardiomyocytes in the condition of HF. Secreted frizzled-related protein 4 (*SFRP4*) has been reported to be expressed in human ventricular myocardium and correlates with apoptosis-related gene expression [88]. Zeng et al. showed that the increased *Sfrp4* in mice myocardial infarction model and knockout of *Sfrp4* could protect against myocardial ischemia and reperfusion injury through attenuating apoptosis of cardiomyocytes [89]. Additionally, evidence from the rat heart ischemic model showed that *Sfrp4* recombinant protein injection could exert a cardio-protective effect and improve the cardiac function of the ischemic heart, which suggested the therapeutic potential of *SFRP4* in humans [90]. Regarding the evident expression of *SFRP4* in cardiac fibroblasts of failed hearts, we mechanically speculated that elevated *SFRP4* affects myocardial fibrosis and remodeling post-HF. Ji et al. found that plasma *SFRP4* concentrations were increased in patients with coronary artery disease (CAD) and acted as an independent factor [91]. Altogether, findings from prior experimental and clinical evidence have indicated the great potential of four hub genes in developing ECM fibrotic, immune, or cardiac remodeling-related biomarkers and therapeutic molecular targets of HF.

Subsequently, we integrated GSVA and GSEA algorithms to explore the shared pathways regulated by the four hub genes in HF. Three significantly activated pathways (including cell adhesion molecules, the chemokine signaling pathway, and cytokine–cytokine receptor interaction) were co-identified through the combination of GSVA and GSEA. In contrast, oxidative phosphorylation was identified as a common suppressed pathway. Notably, the induction of ECM remodeling, inflammatory cytokines, and chemokines have been proven to contribute to the pathogenesis of HF, such as adverse remodeling, myocardial injury, myocardial fibrosis, and dysfunction [70,71,92,93]. Moreover, it was shown that impaired mitochondrial respiratory and decreased oxidative phosphorylation were detected in the hearts of HF patients [94]. More strikingly, we observed the same phenomenon through GSVA and GSEA algorithms: cell adhesion molecules, chemokine signaling pathway, and cytokine–cytokine receptor interaction were active, whereas oxidative phosphorylation was inhibited in HF compared to control. We noticed that four hub genes participated in diverse pathogenic pathways involving ECM, immune factors, and oxidative phosphorylation, suggesting the pathophysiologic heterogeneity of HF. Given the close interplays between these pathogenic pathways in HF progression, we speculated that the continuous induction of chemokines or cytokines stimulated the formation of inflammatory conditions in HF that promoted ECM deposition and impaired oxidative phosphorylation capability, thereby exerting an adverse effect on cardiac function (Figure 11). Regarding the upstream regulatory mechanism that could activate the transcription of hub genes, the transcription factors were predicted using the TRRUST v2 database. After systemically exploring the correlation patterns between the up-regulated TFs in HF and the hub gene-regulated pathogenic pathways, we identified *BNC2* and *MEOX2* as upstream regulators to activate the transcription of hub genes and drive the pathogenic pathways. Basonuclin 2 (*BNC2*), a zinc finger TF, was recently identified as a core TF essential for myofibroblastic activation in fibrosis, causing ECM deposition during fibrogenesis [95]. Mesenchyme Homeobox 2 (*MEOX2*) has previously been reported to regulate the proliferation, differentiation, and migration of vascular endothelial cells and cardiomyocytes [96,97]. *BNC2* and *MEOX2* are closely linked with cardiac functions and may contribute to HF-related pathogenic pathways, which could be confident upstream regulators during HF progression. Taken together, we uncovered a comprehensive mechanism of hub genes involved in HF, which was a compact regulatory interaction of TFs–hub genes–pathways: The up-regulations of TFs (*BNC2* and *MEOX2*) activate the transcriptions of four hub genes (*COL14A1*, *OGN*, *MFAP4*, *SFRP4*) and subsequently drive the activation or suppression of downstream signaling pathways (Figure 11). The interplays between activated pathways related to immune factors and ECM and the suppressed oxidative phosphorylation pathway may further cause abnormal cardiac functions, thereby promoting the development and progression of HF.

Given that the immunity-related pathways involving chemokine and cytokine were observed in hub genes, we wondered which main type of immune cell altered in HF could regulate or participate in these pathways. Leveraging the ssGSEA algorithm, we quantified the infiltration levels of 28 types of immune cells based on the gene expression profiles. Combining with Lasso and RF models, we then obtained a total of 14 key immune cell types with HF (activated CD4+ T cell, activated CD8+ T cell, activated dendritic cell, CD56dim natural killer cell, MDSC, macrophage, natural killer T cell, natural killer cell, neutrophil, Type 17 T helper cell, Type 2 T helper cell, effector memory CD4+ T cell, memory B cell, and central memory CD8+ T cell). Together with expression and correlation analyses on the training and test cohorts, we finally suggested effector memory CD4+ T cell, which demonstrated both increased infiltrations in HF and positive relationships with hub genes, as an essential immune cell type that participated in the chemokine signaling pathway and cytokine–cytokine receptor interaction with hub genes. Preliminary evidence has shown that the infiltrations of effector memory CD4+ T cells caused cardiac fibroblast activation and subsequent fibrosis, influencing pressure-overload-induced HF [98]. This result motivated us to consider that the elevated infiltrations of effector memory CD4+ T cells could activate the chemokine signaling pathway and cytokine–cytokine receptor interaction—potentially by the recruitment of chemokines and cytokines.

Considering that four hub genes (*COL14A1*, *OGN*, *MFAP4*, and *SFRP4*) and two key TFs (*BNC2* and *MEOX2*) may serve as potential therapeutic targets of HF, several small-molecule agents targeting hub genes were predicted. Then, molecular docking showed that the close binding patterns between ligands (small-molecule agents) and receptors (hub genes or TFs), including captopril-targeting *BNC2*, aldosterone-targeting *MEOX2*, cyclopentiazide-targeting *MEOX2*, estradiol-targeting *COL14A1*, tolazoline-targeting *COL14A1*, and genistein-targeting *SFRP4*. We noticed that the estimated RMSE from the docked model was lower than 2 Å, indicating the high credibility of molecular docking. Furthermore, we collected the high-quality structures of receptors and ligands, which largely guarantee the accuracy of molecular docking. Using the molecular docking technique, we illustrated the treatment potential of these TFs and genes. Although the effectiveness of these molecules as drug targets in curing HF was demonstrated from an in-silico perspective, further animal studies and clinical trials are needed. Notably, these small-molecule agents were previously reported to cure or alleviate HF symptoms [64–69]. Therefore, the identified small-molecule agents may reverse the over-expression levels of two hub genes (*COL14A1*, *SFRP4*) and two TFs (*BNC2*, *MEOX2*) in HF, thus inhibiting the regulatory network of TFs–hub genes–pathways and then alleviating HF progression.

In the previous analysis, we illustrated the great potential of four hub genes as molecular diagnostic biomarkers and therapeutic targets for HF. To further investigate whether the four hub genes could serve as extra plasma biomarkers of HF for clinic applications, we collected six whole blood samples derived from three HF patients and three control cases. Then, the qRT-PCR results showed that *OGN* was highly expressed in the plasma of HF patients compared to the control cases. Additionally, *OGN* demonstrated robust diagnostic accuracy in HF recognition and a significantly positive correlation to NT-proBNP concentration. The elevated pattern and robust diagnostic power of plasma *OGN* were subsequently verified in an external cohort, including 21 HF and nine control cases. Some clinical factors, particularly gender and age, have been reported to cause differences in the detection of HF biomarkers [99]. For instance, gender-related differences in Cardiac Troponin (cTn) values have been evident in patients with HF [100] with higher values commonly shown in males. However, most clinicians barely consider these factors when using these biomarkers [100]. Thus, more effort should be put into investigating the interference on diagnostic values of biomarkers, considering the underlying differences brought by the confounding factors. Herein, we deduced that the successful clinically used biomarker should be stably detectable for reflecting disease rather than being disturbed via different characteristics of patients. Accordingly, we wondered whether the expression levels of *OGN* varied in different clinical characteristics, mainly including HF type, gender, and age.

Notably, no significant expressions of *OGN* differed in groups of these clinical factors. This result suggests that *OGN* may serve as a stable biomarker for HF detection with limited affection brought by the confounding factors such as gender and age. Given the promising application of plasma *OGN* in discerning HF, we also studied the causal association of plasma *OGN* and the risk of HF from the genetic insight using GWAS datasets. We used the SNPs as IVs for a two-sample MR study based on the exposure (plasma *OGN*) and outcome (HF) GWAS data. As a result, we found that plasma *OGN* was significantly causally related to the increased risk of HF with an overall OR estimated larger than one measured by five MR approaches. Strikingly, the genetic variant in *OGN* was proven to promote HF occurrence on the large-scale GWAS data, which largely sustained the great potential of *OGN* as an HF biomarker. Thus, these results demonstrated that *OGN* could serve as a plasma detective biomarker that is more indicative of HF patients and HF progression.

Nevertheless, some limitations remain in our present study. First, these findings were mainly generated from an *in silico* analysis of transcriptional datasets and a need for *in-vitro* confirmative studies. We are pushing ahead with a larger-scale investigation by integrating more *in silico* and *in vitro* datasets related to HF. However, it remained an effective means to rapidly explore the candidate diagnostic biomarkers and underlying pathogenic mechanisms of HF, thus providing meaningful and constructive references for understanding HF. Second, we detected the up-regulated *OGN* expressions in the plasma of three HF patients compared to three control cases. The sample sizes were relatively limited. Further validation should be conducted within a prospective and larger multi-center collaboration, which guarantees the credibility of *OGN* in diagnostic applications. Given our understanding of the biomarkers-regulated pathogenic mechanisms in HF, more in-depth molecular studies, such as genetic knockout and pharmacological exploration on cell or animal models, are needed to substantiate our work.

## 5. Conclusions

Four hub genes (*COL14A1*, *OGN*, *MFAP4*, and *SFRP4*) were identified as candidate molecular biomarkers for the diagnosis of HF. Subsequently, the pathogenic mechanisms of hub genes were expounded. The up-regulation of two key TFs (*BNC2* and *MEOX2*) prompted the expression of hub genes and further activated several ECM and immune-related signaling pathways, including cell adhesion molecules, the chemokine signaling pathway, and cytokine–cytokine receptor interaction while inhibiting the oxidative phosphorylation pathway. Moreover, elevated infiltrations of effector memory CD4+ T cells in HF were closely related to the expressions of hub genes. They may contribute to activating the chemokine signaling pathway and cytokine–cytokine receptor interaction, leading to the inflammatory condition of HF. Six small-molecule agents, including captopril, aldosterone antagonist, cyclopentiazide, estradiol, tolazoline, and genistein, were predicted *in silico* to exhibit interactions with two key TFs (*BNC2* and *MEOX2*) and two hub genes (*COL14A1* and *SFRP4*). These agents could be developed into therapeutic drugs to treat HF by curbing the up-regulation of these TFs or hub genes. Finally, our *in vitro* study on clinical specimens suggested that *OGN* could serve as a plasma diagnostic biomarker of HF with robust diagnostic accuracy and positive correlation with NT-proBNP concentration. Two-sample MR analysis further demonstrated the positive causal association between plasma *OGN* and the increased risk of HF. Altogether, our study proposes *OGN* as a candidate diagnostic biomarker of HF and provides novel insight into the pathogenesis of HF.

**Supplementary Materials:** The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/biom14020179/s1>, Table S1: The clinical information of HF patients and control cases in GSE141910; Table S2: The clinical information of HF patients and control cases in GSE57338; Table S3: The clinical information of HF patients and control cases in GSE42955; Table S4: The clinical information of HF patients and control cases in GSE135055; Table S5: The clinical information of HF patients and control cases of our in-house cohort; Table S6: The primer sequences of hub genes and  $\beta$ -actin; Table S7: Detailed information of 926 DEGs; Table S8: Genes

in WGCNA-identified modules; Table S9: Genes in MEGENA-identified modules; Table S10: Top 10 predicted TFs with hub genes; Table S11: Lasso-selected immune cell types; Table S12: RF-selected immune cell types; Table S13: Predicted drugs targeting TFs and hub genes; Table S14: Sensitivity analysis of two-sample MR; Table S15: Summary of biological functions of ten pivotal genes; Figure S1: Workflow of the study; Figure S2: PCA plot of merging test cohort. Figure S3: Construction of weighted gene co-expression network on DEGs. Figure S4: Systematic screening of pivotal genes using five algorithms of cyto-Hubba. Figure S5: ROC curve analysis to evaluate the diagnostic accuracies of pivotal genes. Figure S6: The diagnostic accuracies of four hub genes in the test cohort. Figure S7: Comparison of AUC scores on our identified biomarkers with current biomarkers; Figure S8: Quality control and dimension reduction in sc-RNA seq data. Figure S9: AUCCell scores of different gene markers in each cell population. Figure S10: GSVA and GSEA of hub genes. Figure S11: The activated and suppressed pathways are shared by hub genes; Figure S12: GSVA and GSEA of shared pathways by hub genes within the control and HF cases. Figure S13: The Bayesian networks show the GRN of common pathways regulated by hub genes. Figure S14: Expression levels of 10 predicted TFs in HF compared to the control group in the training cohort. Figure S15: Expression levels of 10 predicted TFs in HF compared to the control group in the test cohort. Figure S16: Molecular docking of candidate drugs with TFs and hub genes. Figure S17: Expression patterns of OGN across diverse clinical characteristics of HF patients. Figure S18: Two-sample MR analysis.

**Author Contributions:** Conceptualization, Y.Z. (Yihao Zhu) and Y.Z. (Yao Zu); methodology, Y.Z. (Yihao Zhu); software, Y.Z. (Yihao Zhu); validation, Y.Z. (Yihao Zhu); formal analysis, Y.Z. (Yihao Zhu); investigation, Y.Z. (Yihao Zhu); resources, Y.Z. (Yihao Zhu); data curation, Y.Z. (Yihao Zhu); writing—original draft preparation, Y.Z. (Yihao Zhu); writing—review and editing, Y.Z. (Yihao Zhu), Y.Z. (Yao Zu), and B.C.; visualization, Y.Z. (Yihao Zhu); supervision, Y.Z. (Yao Zu) and B.C.; project administration, Y.Z. (Yihao Zhu) and Y.Z. (Yao Zu); funding acquisition, Y.Z. (Yao Zu). All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the National Natural Science Foundation of China (grant numbers 32170423,31501166); Chenguang Program from the Shanghai Education Committee (grant number 14CG49); Shanghai Sailing Program of Science and Technology Commission of Shanghai Municipality (grant number 15YF1405000); and SciTech Funding ZY-2023-0001 by CSPFTZ Lingang Special Area Marine Biomedical Innovation Platform.

**Institutional Review Board Statement:** The studies involving human participants were reviewed and approved by the medical ethics committee of the Shanghai Sixth People's Hospital Affiliated to Shanghai Jiao Tong University School of Medicine (Lin-gang) (Ethic code: 2021-KY-108(K); Approved date: 26 November 2021).

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** Data used in this study, including GSE141910, GSE57338, GSE42955, GSE121893, and GSE135055 can be downloaded from the GEO database (<https://www.ncbi.nlm.nih.gov/geo/>, accessed on 30 November 2022).

**Acknowledgments:** We thank Xiaojing Yang for her assistance in the work of qRT-PCR analysis in Figure 10A. We thank Yuxi Chen for her assistance in the work of signature collection in Figure S7.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. McDonagh, T.A.; Metra, M.; Adamo, M.; Gardner, R.S.; Baumbach, A.; Böhm, M.; Burri, H.; Butler, J.; Čelutkienė, J.; Chioncel, O.; et al. 2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: Developed by the Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC). With the special contribution of the Heart Failure Association (HFA) of the ESC. *Eur. J. Heart Fail.* **2022**, *24*, 4–131. [[PubMed](#)]
2. Castiglione, V.; Aimo, A.; Vergaro, G.; Saccaro, L.; Passino, C.; Emdin, M. Biomarkers for the diagnosis and management of heart failure. *Heart Fail. Rev.* **2022**, *27*, 625–643. [[CrossRef](#)] [[PubMed](#)]
3. Savarese, G.; Lund, L.H. Global Public Health Burden of Heart Failure. *Card. Fail. Rev.* **2017**, *3*, 7–11. [[CrossRef](#)] [[PubMed](#)]
4. Bozkurt, B.; Coats, A.J.; Tsutsui, H.; Abdelhamid, M.; Adamopoulos, S.; Albert, N.; Anker, S.D.; Atherton, J.; Böhm, M.; Butler, J.; et al. Universal Definition and Classification of Heart Failure: A Report of the Heart Failure Society of America, Heart Failure

- Association of the European Society of Cardiology, Japanese Heart Failure Society and Writing Committee of the Universal Definition of Heart Failure. *J. Card. Fail.* **2021**, *27*, 387–413.
5. Sarhene, M.; Wang, Y.; Wei, J.; Huang, Y.; Li, M.; Li, L.; Acheampong, E.; Zhengcan, Z.; Xiaoyan, Q.; Yunsheng, X.; et al. Biomarkers in heart failure: The past, current and future. *Heart Fail. Rev.* **2019**, *24*, 867–903. [[CrossRef](#)]
  6. Chow, S.L.; Maisel, A.S.; Anand, I.; Bozkurt, B.; de Boer, R.A.; Felker, G.M.; Fonarow, G.C.; Greenberg, B.; Januzzi, J.J.; Kiernan, M.S.; et al. Role of Biomarkers for the Prevention, Assessment, and Management of Heart Failure: A Scientific Statement From the American Heart Association. *Circulation* **2017**, *135*, e1054–e1091. [[CrossRef](#)]
  7. Sun, C.; Ni, M.; Song, B.; Cao, L. Circulating Circular RNAs: Novel Biomarkers for Heart Failure. *Front. Pharmacol.* **2020**, *11*, 560537. [[CrossRef](#)]
  8. Takase, H.; Dohi, Y. Kidney function crucially affects B-type natriuretic peptide (BNP), N-terminal proBNP and their relationship. *Eur. J. Clin. Investig.* **2014**, *44*, 303–308. [[CrossRef](#)]
  9. Kucher, N.; Printzen, G.; Goldhaber, S.Z. Prognostic role of brain natriuretic peptide in acute pulmonary embolism. *Circulation* **2003**, *107*, 2545–2547. [[CrossRef](#)]
  10. Madamanchi, C.; Alhosaini, H.; Sumida, A.; Runge, M.S. Obesity and natriuretic peptides, BNP and NT-proBNP: Mechanisms and diagnostic implications for heart failure. *Int. J. Cardiol.* **2014**, *176*, 611–617. [[CrossRef](#)]
  11. Braunwald, E. Biomarkers in heart failure. *N. Engl. J. Med.* **2008**, *358*, 2148–2159. [[CrossRef](#)]
  12. Ahmad, T.; Fiuzat, M.; Pencina, M.J.; Geller, N.L.; Zannad, F.; Cleland, J.G.; Snider, J.V.; Blankenberg, S.; Adams, K.F.; Redberg, R.F.; et al. Charting a roadmap for heart failure biomarker studies. *JACC Heart Fail.* **2014**, *2*, 477–488. [[CrossRef](#)]
  13. Piek, A.; Du, W.; de Boer, R.A.; Silljé, H. Novel heart failure biomarkers: Why do we fail to exploit their potential? *Crit. Rev. Clin. Lab. Sci.* **2018**, *55*, 246–263. [[CrossRef](#)]
  14. Savic-Radojevic, A.; Pljesa-Ercegovac, M.; Matic, M.; Simic, D.; Radovanovic, S.; Simic, T. Novel Biomarkers of Heart Failure. *Adv. Clin. Chem.* **2017**, *79*, 93–152.
  15. Weinberg, E.O.; Shimp, M.; Hurwitz, S.; Tominaga, S.; Rouleau, J.L.; Lee, R.T. Identification of serum soluble ST2 receptor as a novel heart failure biomarker. *Circulation* **2003**, *107*, 721–726. [[CrossRef](#)] [[PubMed](#)]
  16. Yu, Y.D.; Xue, Y.T.; Li, Y. Identification and verification of feature biomarkers associated in heart failure by bioinformatics analysis. *Sci. Rep.* **2023**, *13*, 3488. [[CrossRef](#)]
  17. Kolar, V.; Vastrad, B.; Vastrad, C.; Kotturshetti, S.; Tengli, A. Identification of candidate biomarkers and therapeutic agents for heart failure by bioinformatics analysis. *BMC Cardiovasc. Disord.* **2021**, *21*, 329. [[CrossRef](#)] [[PubMed](#)]
  18. Fan, S.; Hu, Y. Integrative analyses of biomarkers and pathways for heart failure. *BMC Med. Genom.* **2022**, *15*, 72. [[CrossRef](#)] [[PubMed](#)]
  19. Zhou, J.; Zhang, W.; Wei, C.; Zhang, Z.; Yi, D.; Peng, X.; Peng, J.; Yin, R.; Zheng, Z.; Qi, H.; et al. Weighted correlation network bioinformatics uncovers a key molecular biosignature driving the left-sided heart failure. *BMC Med. Genom.* **2020**, *13*, 93. [[CrossRef](#)] [[PubMed](#)]
  20. Flam, E.; Jang, C.; Murashige, D.; Yang, Y.; Morley, M.P.; Jung, S.; Kantner, D.S.; Pepper, H.; Bedi, K.J.; Brandimarto, J.; et al. Integrated landscape of cardiac metabolism in end-stage human nonischemic dilated cardiomyopathy. *Nat. Cardiovasc. Res.* **2022**, *1*, 817–829. [[CrossRef](#)] [[PubMed](#)]
  21. Molina-Navarro, M.M.; Roselló-Lletí, E.; Ortega, A.; Tarazón, E.; Otero, M.; Martínez-Dolz, L.; Lago, F.; González-Juanatey, J.R.; España, F.; García-Pavía, P.; et al. Differential gene expression of cardiac ion channels in human dilated cardiomyopathy. *PLoS ONE* **2013**, *8*, e79792. [[CrossRef](#)] [[PubMed](#)]
  22. Liu, Y.; Morley, M.; Brandimarto, J.; Hannenhalli, S.; Hu, Y.; Ashley, E.A.; Tang, W.H.; Moravec, C.S.; Margulies, K.B.; Cappola, T.P.; et al. RNA-Seq identifies novel myocardial gene expression signatures of heart failure. *Genomics* **2015**, *105*, 83–89. [[CrossRef](#)] [[PubMed](#)]
  23. Hua, X.; Wang, Y.Y.; Jia, P.; Xiong, Q.; Hu, Y.; Chang, Y.; Lai, S.; Xu, Y.; Zhao, Z.; Song, J. Multi-level transcriptome sequencing identifies COL1A1 as a candidate marker in human heart failure progression. *BMC Med.* **2020**, *18*, 2. [[CrossRef](#)] [[PubMed](#)]
  24. Kerr, M.K. Linear models for microarray data analysis: Hidden similarities and differences. *J. Comput. Biol.* **2003**, *10*, 891–901. [[CrossRef](#)] [[PubMed](#)]
  25. Rohart, F.; Gautier, B.; Singh, A.; Lê, C.K. mixOmics: An R package for ‘omics feature selection and multiple data integration. *PLoS Comput. Biol.* **2017**, *13*, e1005752. [[CrossRef](#)]
  26. Iqbal, N.; Kumar, P. Integrated COVID-19 Predictor: Differential expression analysis to reveal potential biomarkers and prediction of coronavirus using RNA-Seq profile data. *Comput. Biol. Med.* **2022**, *147*, 105684. [[CrossRef](#)]
  27. Yu, G.; Wang, L.G.; Han, Y.; He, Q.Y. clusterProfiler: An R package for comparing biological themes among gene clusters. *OMICS* **2012**, *16*, 284–287. [[CrossRef](#)]
  28. Langfelder, P.; Horvath, S. WGCNA: An R package for weighted correlation network analysis. *BMC Bioinform.* **2008**, *9*, 559. [[CrossRef](#)] [[PubMed](#)]
  29. Song, W.M.; Zhang, B. Multiscale Embedded Gene Co-expression Network Analysis. *PLoS Comput. Biol.* **2015**, *11*, e1004574. [[CrossRef](#)] [[PubMed](#)]
  30. Cao, J.; Liu, Z.; Liu, J.; Li, C.; Zhang, G.; Shi, R. Bioinformatics Analysis and Identification of Genes and Pathways in Ischemic Cardiomyopathy. *Int. J. Gen. Med.* **2021**, *14*, 5927–5937. [[CrossRef](#)]

31. Yu, G.; Wang, L.G.; Yan, G.R.; He, Q.Y. DOSE: An R/Bioconductor package for disease ontology semantic and enrichment analysis. *Bioinformatics* **2015**, *31*, 608–609. [[CrossRef](#)] [[PubMed](#)]
32. Liu, Z.; Liu, L.; Weng, S.; Guo, C.; Dang, Q.; Xu, H.; Wang, L.; Lu, T.; Zhang, Y.; Sun, Z.; et al. Machine learning-based integration develops an immune-derived lncRNA signature for improving outcomes in colorectal cancer. *Nat. Commun.* **2022**, *13*, 816. [[CrossRef](#)] [[PubMed](#)]
33. Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N.S.; Wang, J.T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **2003**, *13*, 2498–2504. [[CrossRef](#)]
34. Szklarczyk, D.; Franceschini, A.; Wyder, S.; Forslund, K.; Heller, D.; Huerta-Cepas, J.; Simonovic, M.; Roth, A.; Santos, A.; Tsafou, K.P.; et al. STRING v10: Protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* **2015**, *43*, D447–D452. [[CrossRef](#)]
35. Chin, C.H.; Chen, S.H.; Wu, H.H.; Ho, C.W.; Ko, M.T.; Lin, C.Y. cytoHubba: Identifying hub objects and sub-networks from complex interactome. *BMC Syst. Biol.* **2014**, *8* (Suppl. S4), S11. [[CrossRef](#)]
36. Robin, X.; Turck, N.; Hainard, A.; Tiberti, N.; Lisacek, F.; Sanchez, J.C.; Müller, M. pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinform.* **2011**, *12*, 77. [[CrossRef](#)] [[PubMed](#)]
37. Wang, L.; Yu, P.; Zhou, B.; Song, J.; Li, Z.; Zhang, M.; Guo, G.; Wang, Y.; Chen, X.; Han, L.; et al. Single-cell reconstruction of the adult human heart during heart failure and recovery reveals the cellular landscape underlying cardiac function. *Nat. Cell Biol.* **2020**, *22*, 108–119. [[CrossRef](#)]
38. Hao, Y.; Hao, S.; Andersen-Nissen, E.; Mauck, W.R.; Zheng, S.; Butler, A.; Lee, M.J.; Wilk, A.J.; Darby, C.; Zager, M.; et al. Integrated analysis of multimodal single-cell data. *Cell* **2021**, *184*, 3573–3587. [[CrossRef](#)]
39. Aibar, S.; González-Blas, C.B.; Moerman, T.; Huynh-Thu, V.A.; Imrichova, H.; Hulselmans, G.; Rambow, F.; Marine, J.C.; Geurts, P.; Aerts, J.; et al. SCENIC: Single-cell regulatory network inference and clustering. *Nat. Methods* **2017**, *14*, 1083–1086. [[CrossRef](#)]
40. Alquicira-Hernandez, J.; Powell, J.E. Nebulosa recovers single-cell gene expression signals by kernel density estimation. *Bioinformatics* **2021**, *37*, 2485–2487. [[CrossRef](#)]
41. Liberzon, A.; Birger, C.; Thorvaldsdóttir, H.; Ghandi, M.; Mesirov, J.P.; Tamayo, P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* **2015**, *1*, 417–425. [[CrossRef](#)]
42. Hänzelmann, S.; Castelo, R.; Guinney, J. GSEA: Gene set variation analysis for microarray and RNA-seq data. *BMC Bioinform.* **2013**, *14*, 7. [[CrossRef](#)]
43. Cao, R.; Yuan, L.; Ma, B.; Wang, G.; Qiu, W.; Tian, Y. An EMT-related gene signature for the prognosis of human bladder cancer. *J. Cell. Mol. Med.* **2020**, *24*, 605–617. [[CrossRef](#)]
44. Subramanian, A.; Tamayo, P.; Mootha, V.K.; Mukherjee, S.; Ebert, B.L.; Gillette, M.A.; Paulovich, A.; Pomeroy, S.L.; Golub, T.R.; Lander, E.S.; et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 15545–15550. [[CrossRef](#)]
45. Sato, N.; Tamada, Y.; Yu, G.; Okuno, Y. CBNplot: Bayesian network plots for enrichment analysis. *Bioinformatics* **2022**, *38*, 2959–2960. [[CrossRef](#)]
46. Han, H.; Cho, J.W.; Lee, S.; Yun, A.; Kim, H.; Bae, D.; Yang, S.; Kim, C.Y.; Lee, M.; Kim, E.; et al. TRRUST v2: An expanded reference database of human and mouse transcriptional regulatory interactions. *Nucleic Acids Res.* **2018**, *46*, D380–D386. [[CrossRef](#)] [[PubMed](#)]
47. Li, L.; Wang, X. Identification of gastric cancer subtypes based on pathway clustering. *NPJ Precis. Oncol.* **2021**, *5*, 46. [[CrossRef](#)]
48. Yu, H.; Yu, M.; Li, Z.; Zhang, E.; Ma, H. Identification and analysis of mitochondria-related key genes of heart failure. *J. Transl. Med.* **2022**, *20*, 410. [[CrossRef](#)] [[PubMed](#)]
49. Wang, T.; Tian, J.; Jin, Y. VCAM1 expression in the myocardium is associated with the risk of heart failure and immune cell infiltration in myocardium. *Sci. Rep.* **2021**, *11*, 19488. [[CrossRef](#)] [[PubMed](#)]
50. Tu, P.; Li, X.; Cao, L.; Zhong, M.; Xie, Z.; Wu, Z. Machine learning and BP neural network revealed abnormal B cell infiltration predicts the survival of lung cancer patients. *Front. Oncol.* **2022**, *12*, 882018. [[CrossRef](#)]
51. Engebretsen, S.; Bohlin, J. Statistical predictions with glmnet. *Clin. Epigenet.* **2019**, *11*, 123. [[CrossRef](#)]
52. Statnikov, A.; Wang, L.; Aliferis, C.F. A comprehensive comparison of random forests and support vector machines for microarray-based cancer classification. *BMC Bioinform.* **2008**, *9*, 319. [[CrossRef](#)]
53. Yoo, M.; Shin, J.; Kim, J.; Ryall, K.A.; Lee, K.; Lee, S.; Jeon, M.; Kang, J.; Tan, A.C. DSigDB: Drug signatures database for gene set analysis. *Bioinformatics* **2015**, *31*, 3069–3071. [[CrossRef](#)]
54. Pagadala, N.S.; Syed, K.; Tuszynski, J. Software for molecular docking: A review. *Biophys. Rev.* **2017**, *9*, 91–102. [[CrossRef](#)]
55. Trott, O.; Olson, A.J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **2010**, *31*, 455–461. [[CrossRef](#)]
56. Emdin, C.A.; Khera, A.V.; Kathiresan, S. Mendelian Randomization. *JAMA* **2017**, *318*, 1925–1926. [[CrossRef](#)] [[PubMed](#)]
57. Xie, J.; Huang, H.; Liu, Z.; Li, Y.; Yu, C.; Xu, L.; Xu, C. The associations between modifiable risk factors and nonalcoholic fatty liver disease: A comprehensive Mendelian randomization study. *Hepatology* **2023**, *77*, 949–964. [[CrossRef](#)] [[PubMed](#)]
58. Fang, Y.; Huang, S.; Han, L.; Wang, S.; Xiong, B. Comprehensive Analysis of Peritoneal Metastasis Sequencing Data to Identify LINC00924 as a Prognostic Biomarker in Gastric Cancer. *Cancer Manag. Res.* **2021**, *13*, 5599–5611. [[CrossRef](#)]
59. Francis, S.H.; Busch, J.L.; Corbin, J.D.; Sibley, D. cGMP-dependent protein kinases and cGMP phosphodiesterases in nitric oxide and cGMP action. *Pharmacol. Rev.* **2010**, *62*, 525–563. [[CrossRef](#)] [[PubMed](#)]

60. Müller, F.U.; Neumann, J.; Schmitz, W. Transcriptional regulation by cAMP in the heart. *Mol. Cell. Biochem.* **2000**, *212*, 11–17. [[CrossRef](#)] [[PubMed](#)]
61. Leach, J.P.; Heallen, T.; Zhang, M.; Rahmani, M.; Morikawa, Y.; Hill, M.C.; Segura, A.; Willerson, J.T.; Martin, J.F. Hippo pathway deficiency reverses systolic heart failure after infarction. *Nature* **2017**, *550*, 260–264. [[CrossRef](#)]
62. Lohse, M.J.; Engelhardt, S.; Eschenhagen, T. What is the role of beta-adrenergic signaling in heart failure? *Circ. Res.* **2003**, *93*, 896–906. [[CrossRef](#)]
63. Dobaczewski, M.; Chen, W.; Frangogiannis, N.G. Transforming growth factor (TGF)- $\beta$  signaling in cardiac remodeling. *J. Mol. Cell. Cardiol.* **2011**, *51*, 600–606. [[CrossRef](#)] [[PubMed](#)]
64. Rajagopalan, S.; Pitt, B. Aldosterone as a target in congestive heart failure. *Med. Clin. N. Am.* **2003**, *87*, 441–457. [[CrossRef](#)]
65. Jankowska, E.A.; Rozentryt, P.; Ponikowska, B.; Hartmann, O.; Kustrzycka-Kratochwil, D.; Reczuch, K.; Nowak, J.; Borodulin-Nadziejka, L.; Polonski, L.; Banasiak, W.; et al. Circulating estradiol and mortality in men with systolic chronic heart failure. *JAMA* **2009**, *301*, 1892–1901. [[CrossRef](#)] [[PubMed](#)]
66. Crawford, R.J.; Allman, S.; Gibson, W.; Kitchen, S.; Richards, H.H. A comparative study of frusemide-amiloride and cyclopentiazide-potassium chloride in the treatment of congestive cardiac failure in general practice. *J. Int. Med. Res.* **1988**, *16*, 143–149. [[CrossRef](#)]
67. Matori, H.; Umar, S.; Nadadur, R.D.; Sharma, S.; Partow-Navid, R.; Afkhami, M.; Amjedi, M.; Eghbali, M. Genistein, a soy phytoestrogen, reverses severe pulmonary hypertension and prevents right heart failure in rats. *Hypertension* **2012**, *60*, 425–430. [[CrossRef](#)]
68. Sommer, C.M.; Radeleff, B.A. A novel approach for percutaneous treatment of massive nonocclusive mesenteric ischemia: Tolazoline and glycerol trinitrate as effective local vasodilators. *Catheter. Cardiovasc. Interv.* **2009**, *73*, 152–155. [[CrossRef](#)]
69. Pfeffer, M.A.; McMurray, J.J.; Velazquez, E.J.; Rouleau, J.L.; Køber, L.; Maggioni, A.P.; Solomon, S.D.; Swedberg, K.; Van de Werf, F.; White, H.; et al. Valsartan, captopril, or both in myocardial infarction complicated by heart failure, left ventricular dysfunction, or both. *N. Engl. J. Med.* **2003**, *349*, 1893–1906. [[CrossRef](#)] [[PubMed](#)]
70. Baumgarten, G.; Knuefermann, P.; Kalra, D.; Gao, F.; Taffet, G.E.; Michael, L.; Blackshear, P.J.; Carballo, E.; Sivasubramanian, N.; Mann, D.L. Load-dependent and -independent regulation of proinflammatory cytokine and cytokine receptor gene expression in the adult mammalian heart. *Circulation* **2002**, *105*, 2192–2197. [[CrossRef](#)]
71. Salvador, A.M.; Nevers, T.; Velázquez, F.; Aronovitz, M.; Wang, B.; Abadía, M.A.; Jaffe, I.Z.; Karas, R.H.; Blanton, R.M.; Alcaide, P. Intercellular Adhesion Molecule 1 Regulates Left Ventricular Leukocyte Infiltration, Cardiac Remodeling, and Function in Pressure Overload-Induced Heart Failure. *J. Am. Heart Assoc.* **2016**, *5*, e3126. [[CrossRef](#)]
72. Garlapati, V.; Molitor, M.; Michna, T.; Harms, G.S.; Finger, S.; Jung, R.; Lagrange, J.; Efentakis, P.; Wild, J.; Knorr, M.; et al. Targeting myeloid cell coagulation signaling blocks MAP kinase/TGF- $\beta$ 1-driven fibrotic remodeling in ischemic heart failure. *J. Clin. Investig.* **2023**, *133*, e156436. [[CrossRef](#)] [[PubMed](#)]
73. Wang, Q.; Zhang, Y.; Wang, M.; Song, W.M.; Shen, Q.; McKenzie, A.; Choi, I.; Zhou, X.; Pan, P.Y.; Yue, Z.; et al. The landscape of multiscale transcriptomic networks and key regulators in Parkinson’s disease. *Nat. Commun.* **2019**, *10*, 5234. [[CrossRef](#)] [[PubMed](#)]
74. Chella, K.K.; Kurt, Z.; Barrere-Cain, R.; Sabir, S.; Das, A.; Floyd, R.; Vergnes, L.; Zhao, Y.; Che, N.; Charugundla, S.; et al. Integration of Multi-omics Data from Mouse Diversity Panel Highlights Mitochondrial Dysfunction in Non-alcoholic Fatty Liver Disease. *Cell Syst.* **2018**, *6*, 103–115. [[CrossRef](#)] [[PubMed](#)]
75. Ma, Z.; Wang, X.; Lv, Q.; Gong, Y.; Xia, M.; Zhuang, L.; Lu, X.; Yang, Y.; Zhang, W.; Fu, G.; et al. Identification of Underlying Hub Genes Associated with Hypertrophic Cardiomyopathy by Integrated Bioinformatics Analysis. *Pharmgenomics Pers. Med.* **2021**, *14*, 823–837. [[CrossRef](#)] [[PubMed](#)]
76. Gil-Cayuela, C.; Rivera, M.; Ortega, A.; Tarazón, E.; Triviño, J.C.; Lago, F.; González-Juanatey, J.R.; Almenar, L.; Martínez-Dolz, L.; Portolés, M. RNA sequencing analysis identifies new human collagen genes involved in cardiac remodeling. *J. Am. Coll. Cardiol.* **2015**, *65*, 1265–1267. [[CrossRef](#)] [[PubMed](#)]
77. Tao, G.; Levay, A.K.; Peacock, J.D.; Huk, D.J.; Both, S.N.; Purcell, N.H.; Pinto, J.R.; Galantowicz, M.L.; Koch, M.; Lucchesi, P.A.; et al. Collagen XIV is important for growth and structural integrity of the myocardium. *J. Mol. Cell Cardiol.* **2012**, *53*, 626–638. [[CrossRef](#)] [[PubMed](#)]
78. Petretto, E.; Sarwar, R.; Grieve, I.; Lu, H.; Kumaran, M.K.; Muckett, P.J.; Mangion, J.; Schroen, B.; Benson, M.; Punjabi, P.P.; et al. Integrated genomic approaches implicate osteoglycin (Ogn) in the regulation of left ventricular mass. *Nat. Genet.* **2008**, *40*, 546–552. [[CrossRef](#)] [[PubMed](#)]
79. Van Aelst, L.N.; Voss, S.; Carai, P.; Van Leeuwen, R.; Vanhoutte, D.; Sanders-van, W.S.; Eurlings, L.; Swinnen, M.; Verheyen, F.K.; Verbeken, E.; et al. Osteoglycin prevents cardiac dilatation and dysfunction after myocardial infarction through infarct collagen strengthening. *Circ. Res.* **2015**, *116*, 425–436. [[CrossRef](#)]
80. Deckx, S.; Heggermont, W.; Carai, P.; Rienks, M.; Dresselaers, T.; Himmelreich, U.; van Leeuwen, R.; Lommen, W.; van der Velden, J.; Gonzalez, A.; et al. Osteoglycin prevents the development of age-related diastolic dysfunction during pressure overload by reducing cardiac fibrosis and inflammation. *Matrix Biol.* **2018**, *66*, 110–124. [[CrossRef](#)]
81. Fang, Y.; Chang, Z.; Xu, Z.; Hu, J.; Zhou, H.; Yu, S.; Wan, X. Osteoglycin silencing exerts inhibitory effects on myocardial fibrosis and epithelial/endothelial-mesenchymal transformation in a mouse model of myocarditis. *Biofactors* **2020**, *46*, 1018–1030. [[CrossRef](#)]

82. Humeres, C.; Frangogiannis, N.G. Fibroblasts in the Infarcted, Remodeling, and Failing Heart. *JACC Basic Transl. Sci.* **2019**, *4*, 449–467. [[CrossRef](#)]
83. Lindholt, J.S.; Madsen, M.; Kirketerp-Møller, K.L.; Schlosser, A.; Kristensen, K.L.; Andersen, C.B.; Sorensen, G.L. High plasma microfibrillar-associated protein 4 is associated with reduced surgical repair in abdominal aortic aneurysms. *J. Vasc. Surg.* **2020**, *71*, 1921–1929. [[CrossRef](#)]
84. Kanaan, R.; Medlej-Hashim, M.; Jounblat, R.; Pilecki, B.; Sorensen, G.L. Microfibrillar-associated protein 4 in health and disease. *Matrix Biol.* **2022**, *111*, 1–25. [[CrossRef](#)] [[PubMed](#)]
85. Sækmoose, S.G.; Mössner, B.; Christensen, P.B.; Lindvig, K.; Schlosser, A.; Holst, R.; Barington, T.; Holmskov, U.; Sorensen, G.L. Microfibrillar-Associated Protein 4: A Potential Biomarker for Screening for Liver Fibrosis in a Mixed Patient Cohort. *PLoS ONE* **2015**, *10*, e140418. [[CrossRef](#)]
86. Wang, H.B.; Yang, J.; Shuai, W.; Yang, J.; Liu, L.B.; Xu, M.; Tang, Q.Z. Deletion of Microfibrillar-Associated Protein 4 Attenuates Left Ventricular Remodeling and Dysfunction in Heart Failure. *J. Am. Heart Assoc.* **2020**, *9*, e15307. [[CrossRef](#)]
87. Dorn, L.E.; Lawrence, W.; Petrosino, J.M.; Xu, X.; Hund, T.J.; Whitson, B.A.; Stratton, M.S.; Janssen, P.; Mohler, P.J.; Schlosser, A.; et al. Microfibrillar-Associated Protein 4 Regulates Stress-Induced Cardiac Remodeling. *Circ. Res.* **2021**, *128*, 723–737. [[CrossRef](#)]
88. Pawar, N.M.; Rao, P. Secreted frizzled related protein 4 (sFRP4) update: A brief review. *Cell. Signal.* **2018**, *45*, 63–70. [[CrossRef](#)] [[PubMed](#)]
89. Zeng, W.; Cao, Y.; Jiang, W.; Kang, G.; Huang, J.; Xie, S. Knockdown of Sfrp4 attenuates apoptosis to protect against myocardial ischemia/reperfusion injury. *J. Pharmacol. Sci.* **2019**, *140*, 14–19. [[CrossRef](#)] [[PubMed](#)]
90. Matsushima, K.; Suyama, T.; Takenaka, C.; Nishishita, N.; Ikeda, K.; Ikada, Y.; Sawa, Y.; Jakt, L.M.; Mori, H.; Kawamata, S. Secreted frizzled related protein 4 reduces fibrosis scar size and ameliorates cardiac function after ischemic injury. *Tissue Eng. Part. A* **2010**, *16*, 3329–3341. [[CrossRef](#)] [[PubMed](#)]
91. Ji, Q.; Zhang, J.; Du, Y.; Zhu, E.; Wang, Z.; Que, B.; Miao, H.; Shi, S.; Qin, X.; Zhao, Y.; et al. Human epicardial adipose tissue-derived and circulating secreted frizzled-related protein 4 (SFRP4) levels are increased in patients with coronary artery disease. *Cardiovasc. Diabetol.* **2017**, *16*, 133. [[CrossRef](#)] [[PubMed](#)]
92. Zannad, F.; Rossignol, P.; Iraqi, W. Extracellular matrix fibrotic markers in heart failure. *Heart Fail. Rev.* **2010**, *15*, 319–329. [[CrossRef](#)] [[PubMed](#)]
93. Hanna, A.; Frangogiannis, N.G. Inflammatory Cytokines and Chemokines as Therapeutic Targets in Heart Failure. *Cardiovasc. Drugs Ther.* **2020**, *34*, 849–863. [[CrossRef](#)] [[PubMed](#)]
94. Lemieux, H.; Semsroth, S.; Antretter, H.; Höfer, D.; Gnaiger, E. Mitochondrial respiratory control and early defects of oxidative phosphorylation in the failing human heart. *Int. J. Biochem. Cell Biol.* **2011**, *43*, 1729–1738. [[CrossRef](#)]
95. Bobowski-Gerard, M.; Boulet, C.; Zummo, F.P.; Dubois-Chevalier, J.; Gheeraert, C.; Bou, S.M.; Strub, J.M.; Farce, A.; Ploton, M.; Guille, L.; et al. Functional genomics uncovers the transcription factor BNC2 as required for myofibroblastic activation in fibrosis. *Nat. Commun.* **2022**, *13*, 5324. [[CrossRef](#)] [[PubMed](#)]
96. Lin, J.; Friesen, M.T.; Bocangel, P.; Cheung, D.; Rawszer, K.; Wigle, J.T. Characterization of Mesenchyme Homeobox 2 (MEOX2) transcription factor binding to RING finger protein 10. *Mol. Cell. Biochem.* **2005**, *275*, 75–84. [[CrossRef](#)]
97. Douville, J.M.; Cheung, D.Y.; Herbert, K.L.; Moffatt, T.; Wigle, J.T. Mechanisms of MEOX1 and MEOX2 regulation of the cyclin dependent kinase inhibitors p21 and p16 in vascular endothelial cells. *PLoS ONE* **2011**, *6*, e29099. [[CrossRef](#)]
98. Adamo, L.; Rocha-Resende, C.; Prabhu, S.D.; Mann, D.L. Reappraising the role of inflammation in heart failure. *Nat. Rev. Cardiol.* **2020**, *17*, 269–285. [[CrossRef](#)]
99. Sciomer, S.; Moscucci, F.; Salvioni, E.; Marchese, G.; Bussotti, M.; Corrà, U.; Piepoli, M.F. Role of gender, age and BMI in prognosis of heart failure. *Eur. J. Prev. Cardiol.* **2020**, *27* (Suppl. S2), 46–51. [[CrossRef](#)]
100. Cediel, G.; Codina, P.; Spitaleri, G.; Domingo, M.; Santiago-Vacas, E.; Lupón, J.; Bayes-Genis, A. Gender-Related Differences in Heart Failure Biomarkers. *Front. Cardiovasc. Med.* **2020**, *7*, 617705. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.