Multivariable prognostic model for heart failure in Chinese Han population-based setting

Man Huang1,2, Lei Xiao1,2, Yang Sun1,2, Dong Hu1,2, Yanghui Chen1,2, Yan Wang1,2 and Dao Wen Wang1,2*

1Division of Cardiology, Department of Internal Medicine, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, 430030, PR China; and 2Hubei Key Laboratory of Genetics and Molecular Mechanism of Cardiologic Disorders, Huazhong University of Science and Technology, Wuhan, PR China

Abstract

Aims The prognosis of heart failure (HF) depends on genetic predisposition, and recent studies have shown that impaired autophagy is involved in HF. This study was aimed to construct a prognostic model combining polygenetic background based on the autophagy pathway and other traditional risk factors (TRF) of HF prognosis.

Methods and results Via re-analysing the transcriptomic data of 50 failing and 14 non-failing donors, differentially expressed autophagy-related genes (ARGs) were chosen for further comparison and analysis with whole exome sequencing and follow-up data of 1000 HF patients. By searching from reported articles, prognosis-related polymorphisms were identified. ARGs and prognosis-related polymorphisms were used to develop genetic risk score (GRS) and genetic risk factor (GRF), respectively. We compared the predictive power of five models [Model 1, GRS; Model 2, composite of TRF and N-terminal B-type natriuretic peptide (NT-proBNP); Model 3, composite of TRF, NT-proBNP, and GRS; Model 4, composite of TRF, NT-proBNP, and GRF; and Model 5, composite of TRF, NT-proBNP, GRF, and GRS] by applying receiver operating characteristic curves. Twenty-four prognosis-related polymorphisms were used to construct GRF and 11 variants among 48 differentially expressed ARGs associated with clinical outcomes of HF patients were applied for GRS. GRS was strongly associated with cardiac mortality of HF patients, independent of TRF and GRF (95% confidence interval 1.273–1.739, P = 5.78 × 10⁻⁷). Comparing with patients with lowest GRS tertile, those with highest tertile had higher risks of developing worse clinical outcomes (hazard ratio = 1.866; 95% confidence interval 1.352–2.575, P = 1.47 × 10⁻⁴). The discrimination power of the model including GRS, TRF, GRF, and NT-proBNP is most considerable (area under curve = 0.777), especially in men, patients over 60, patients with hypertension, patients without diabetes or hyperlipidaemia.

Conclusions The model combining autophagy-related GRS, TRF, GRF, and NT-proBNP performs well in distinguishing between worse-prognosis and better-prognosis HF patients, leading a promising strategy for HF treatment and HF prevention.

Keywords Genetic risk score; Genetic risk factor; Heart failure; Autophagy; Prognosis; Model

Introduction

Heart failure (HF) is the common terminal stage of many cardiovascular disease pathogenically characterized by structural or functional impairment of ventricular filling or ejection of blood. Given its high morbidity and mortality, HF is the leading cause of death in industrialized nations. Over the past several decades, tremendous efforts had been paid on the treatment of HF, yet only a third of patients are reported to benefit from existing treatments. The remaining two-thirds are therefore our target population in need of earlier intervention and intensive treatment. At present, the management of HF patients focus heavily on symptomatic treatment rather than aetiological treatment. The patients, however, are already in a state of serious illness that is difficult to reverse. Therefore, discrimination of the population with higher risk of HF and earlier prevention are extremely essential to achieve a favourable prognosis.
including degradation of long-lived, aggregated, and misfolded proteins and clearance of damaged organelles. However, it is a double-edged sword for HF patients as recent studies have shown that impaired autophagy is involved in the development of HF. Currently, several studies have reported that autophagosomes accumulated in cardiomyocytes in animal models of HF. In a study conducted by Nakai et al., the adult mice developed into left ventricular dilation and HF under pressure-overload conditions with cardiac-specific deficiency of Atg5, a protein involved in autophagy. Ghosh and Pattison have summarized the autophagy-related genes (ARGs) from the key studies of autophagy in cardiac pathology. However, in contrast to the numerous studies in cells and mouse, little is known about their roles in patients with cardiac dysfunction.

A large number of prognostic markers including several genetic polymorphisms of death and/or HF hospitalization have been identified in patients with HF. However, the clinical applicability of conventional predictive model consisting of several indicators is limited and precise risk stratification in HF remains challenging. In addition, any related single polymorphism cannot determine the survival of HF. Genetic risk score (GRS), a model that can summarize the accumulation of trait-associated single nucleotide polymorphisms (SNPs) from high-throughput sequencing data into a single variable for an individual, is an alternative strategy to predict the prognosis of HF, a genetic predisposition disease identified by Framingham Heart Study and Swedish Nationwide Adoption Study. Unlike traditional predictive model for HF, there are few papers describing about genetic predictive model, and this is why research focused on this kind of model to predict the prognosis of HF is urgently needed.

Considering that autophagy is correlated with HF, an HF-related GRS consisted of 11 loci located in ARGs associated with the prognosis of HF patients for assessment of clinical outcome was constructed, as well as a 24 disease-causing-SNPs genetic risk factor (GRF) by searching from articles for reported polymorphisms. GRS, GRF, and other traditional biomarkers were made up as a multivariable prognostic model for HF.

Materials and methods

Study population

We have recruited 1000 chronic heart failure (CHF) patients from Tongji Hospital between March 2008 and November 2017 based on the inclusion and exclusion criteria mentioned before. According to the follow-up protocol, demographic variables, medical history, family history, anthropometric measurements, clinical characteristics, and endpoint events were obtained from structured questionnaires, medical records, face-to-face interviews, and/or phone-call interviews. By the end of the study, only two (0.2%) patients were lost to follow-up, yielding a high follow-up rate of 99.8% (998/1000). We defined the primary endpoints as composite of heart transplantation or cardiovascular death that was confirmed by hospital death certificates or medical records. The traditional risk factors (TRF) were included as follows: sex, age, left ventricular ejection fraction (LVEF), systolic blood pressure, diastolic blood pressure, and history of hypertension, diabetes mellitus, hyperlipaemia, and smoking. The condition of beta-blocker taking as an adjusted factor was collected meanwhile. Written informed consent was obtained from all participants. The study was approved by the ethics committee of Tongji Hospital affiliated with Tongji Medical School and Huazhong University of Science and Technology and was conducted in accordance with the Declaration of Helsinki and the International Conference on Harmonization Guidelines for Good Clinical Practice.

Whole exome sequencing

Genomic DNA was extracted from the peripheral blood leukocytes using FastPure DNA Isolation Mini Kit (Vazyme) and Tiangen commercially available kit (Tiangen). Experimental workflow, sample preparation, and sequencing were performed as protocol. All gDNA were of high quality and were determined through spectrophotometric and electrophoretic analyses. First, genomic DNA was disrupted into 300 bp fragments by an ultrasonicator (Covaris). Next, we used SureSelectXT exon V6 kit (Agilent) to prepare library, capture and index the fragments, with adapters. Then these fragments were amplified using Herculase II Fusion DNA Polymerase (Agilent). Following standard Illumina protocol, paired end, 300 bp read-length sequencing was performed to the amplicons by an Illumina HiSeq X Ten sequencer (Illumina).

Data processing and quality control

The whole exome sequencing (WES) data were processed according to GATK Best Practices recommendations. The sequence reads for the exome sequence of each individual were aligned to hg19 human reference genome (GRCh37 Genome Browser) using Burrows-Wheeler Alignment Tool (BWA) 0.7.17. The picard (http://picard.sourceforge.net) was used in sorting output bamfiles and removing duplicated reads. GATK Version 3.4 was applied for INDEL realignment, base quality recalibration, variant discovery, and variant quality score recalibration (VQSR). The WES data were stored with Variant Call Format (VCF). The VCFtools (https://github.com/vcftools/vcftools) was used to perform data analysis, and invalid data were eliminated before establishing
available data pools. The variants exclusion criteria were low coverage (<20×), low quality score (<20), and low average quality (<3). Qualified samples were defined as variants of over 80% of the individuals that reached the read coverage of 20×. Considering the repeatability of data processing, we employed appropriate quality control procedures to suit the WES summary statistics adapted for minor allele frequency (MAF). PLINK17 was used to control imputation quality and Hardy–Weinberg equilibrium. The multidimensional scaling (MDS) and genetic principal components analysis (PCA) was used for cohort structure quality control. We used ANNOVAR18 to annotate the qualified variants.

Integrating autophagy-related genes and transcriptome analysis

The flowchart of this study is provided in Supporting Information, Figure S1. We searched ARGs from Human Autophagy Database (HAdb; http://www.autophagy.lu/index.html) and GeneCards database with relevance score > 7. The two gene sets were combined and integrated into an ARG set. We obtained CHF RNA-sequencing results of 64 samples from human left ventricular tissue in GEO database19 (Accession Number GSE116250) (http://www.ncbi.nlm.nih.gov/geo/), which were composed of 14 non-failing donors, 37 dilated cardiomyopathy (DCM), and 13 ischaemic cardiomyopathy (ICM) samples. After taking the intersection between ARG set and expressed genes in RNA-seq data, we identified differentially expressed genes (DEGs) via comparing the gene expression of CHF and control groups. We used a false discovery rate (FDR) < 0.05 and a \(|\log2\text{fold change}| (|\log2\text{FC}|) > 0.5\) as screening criteria to obtain the differentially expressed ARGs, which were selected as candidate genes for further analysis.

Genetic risk factor

All available articles about HF prognosis and gene polymorphisms published from July 1998 to September 2020 were searched for from electronic databases, PubMed and the Chinese National Knowledge Infrastructure (CNKI), by combining the following search terms: ‘heart failure’, ‘HF’, ‘polymorphism’, ‘variant’, ‘genetic’, ‘prognosis’, ‘outcome’, ‘survival’, and ‘mortality’. Accessible variants were filtered via aligning to the local WES data. After obtaining hazard ratios (HRs) for three genotypes of each variant through Cox proportional hazard regression analysis, all the HRs then be pooled into GRF, by following formula:

\[
\text{GRF}_m = \sum_{t}^{n} \text{HR}_{mn}
\]

where \(\text{GRF}_m\) denotes the summed weight of HR of each polymorphism for subject \(m\) (\(m = 1, \ldots, 1000\)). \(\text{HR}_{mn}\) is the HR of variant \(n\) for subject \(m\) (\(m = 1, \ldots, 1000\)). GRF is a self-defined term, presenting the collective effects of all available reported polymorphisms that have different related effects sizes to the survival of HF that is used to be the addition to the subsequent comparison of predictive power.

Construction of genetic risk score

The following analyses were based on both dominant model and recessive model.20 From our whole exome sequence data, we selected common SNPs with MAF > 0.05 among the ARGs. SNPs that significantly associated with the prognosis of CHF were selected through univariate Cox regression analysis, with a threshold of \(P < 0.05\). After calculating the linkage disequilibrium (LD), only one SNP was selected as a target SNP if the variants were in strong LD (\(r^2 > 0.9\)) with each other. Through Cox proportional hazard, we can get HRs for SNPs. Next, we calculated the weight for each of the 11 SNPs included into the model, according to the following formulas:

\[
\omega_{ij} = \begin{cases} 
1 & \text{if } \text{HR}_j < 1 \text{ and } X_{ij} = 1 \\
\text{HR}_j & \text{if } \text{HR}_j > 1 \text{ and } X_{ij} = 1 \\
1 & \text{if } X_{ij} = 0 
\end{cases}
\]

where \(\omega_{ij}\) denotes the weight of SNP \(j\) (\(i = 1, \ldots, 11\)) for subject \(i\) (\(i = 1, \ldots, 1000\)). The HR for SNP \(j\) is \(\text{HR}_j\), and let \(X_{ij}\) be the status of genotype for SNP \(j\) in subject \(i\) (1 represents those genotypes with higher mortality risk for HF, and 0 represents other genotypes). The multilocus GRS for each subject was constructed by taking sum across the 11 SNPs, through the formula below:

\[
\text{GRS}_i = \sum_{j=1}^{11} \omega_{ij}
\]

where GRS\(_i\) is the GRS for subject \(i\), which can be used for predicting the prognosis of HF.

Statistical analysis

Data were expressed as mean ± standard deviation (SD) for continuous variable, while percentage or median value for binary or categorical variables. DEGs and volcano plots were performed via the R packages ‘limma’ and ‘ggplot2’, respectively. LD analysis was performed via Haplview 4.1. Student’s t-test was used to compare the correlation between the GRS, TRF, and GRF. The cohort was divided into higher-risk, middle-risk, and lower-risk groups according to the tertiles of GRS and was analysed by the Kaplan–Meier method. Receiver operating character-
Results

Characteristics of study population

In total, 1000 patients with CHF were enrolled in the present study, with 787 caused by primary DCM and 213 caused by ICM (Table 1), noting that ICM was developed from myocardial infarction (MI) or arrhythmia mostly. There were 743 male patients (74.3%), and all patients of the cohort were Chinese Han. The mean age at diagnosis was 57.00 ± 14.19 years old. Upon enrolment, the mean LVEF was 57.00 ± 14.19%. Fifty-six (5.60%) patients have been suffering from stroke and 271 (27.10%) had a history of arrhythmia. Three hundred ninety-two (39.20%) patients are complicated with hypertension, 175 (17.50%) with diabetes, and 50 (5.00%) with hyperlipidaemia. More detailed clinical information is shown in Table 1. During the follow-up with a median time of 28.5 months, 258 (25.8%) primary endpoint events occurred.

Differentially expressed autophagy-related genes between patients with heart failure and healthy controls

Firstly, to get more comprehensive ARGs, HAdb database and GeneCards database were adopted simultaneously, and then 334 ARGs were filtered, as shown in Supporting Information, Figure S1. Secondly, transcriptomic dataset (GSE116250) of left ventricular tissue with HF in the GEO database was analysed for DEGs using GEO2R. Sixty-eight common genes between ARGs and HF-related genes were visualized in a venn diagram (Supporting Information, Figure S2A). With the following cut-off criteria of FDR < 0.05 and \(| \log_{2} FC | > 0.5\), 48 DEGs were ultimately identified between CHF tissues and non-failing tissues (Supporting Information, Table S2), including 19 DEGs down-regulated and 29 DEGs up-regulated (Supporting Information, Figure S2B). The following ARGs were down-regulated: ATG2B, ATG3, BECNJ1, C9orf72, CALCOCO2, CYCS, DNM1L, EIF4EBP1, GSK3B, ITGB1, KIF5B, MAP 3K5, MAPK9, NAMPT, OPTN, PIK3CB, RHEB, VDAC1, and WDFY3. The following ARGs were up-regulated: CXCR4, CX3CL1, ARSA, BAG3, DAPK3, EDDEM1, GAA, GNAI3, HGS, HSP90AB1, HSPB8, ITGB4, ITPR1, JUN, LAMP1, MAP2K7, MCL1, PELP1, PPP1R15A, PRKAB1, RAB33B, RAC1, RPTOR, SIRT1, SPNS1, TM9SF1, TMEM208, TP53INP2, and ULKI.

Screening single nucleotide polymorphisms in whole exome sequencing data associated with prognosis of heart failure

We obtained 149 common variants located within the above 48 ARGs (Supporting Information, Table S2). Subsequent univariate Cox regression analysis identified 10 and 5 common SNPs that were significantly correlated with the prognosis of HF patients in dominant and recessive model (Figure 1), respectively. Although two loci, rs11258194 and rs10412007, reached significance in both dominant and recessive models, we adopted recessive model for the loci because of the more significant association with HF in recessive model. To eliminate LD among the 13 SNPs, we conducted a population-based linkage analysis using Plink and found a high LD score \((r^2 > 0.99)\) between rs2289622 and rs9323945 and between rs3759601 and rs3759602 (Supporting Information, Figure S3). Therefore, rs2289622 and rs3759601 were taken into the further analysis. To sum up, seven SNPs in dominant model and four SNPs in recessive model (Supporting Information, Table S3) were included into the following construction of GRS. In summary, four variants (rs2289622, rs1822372, rs12434329, and rs3759601) were located in ATG2B, and two variants (rs11258210 and rs11258194) were located in OPTN among the 11 SNPs. There were seven SNPs in exon regions, with four (rs10122902, rs1822372, rs12434329, and rs1567962) synonymous variants and three (rs2289622, rs3759601, and rs11258194) missense variants. The MAF in the East Asian population ranged from 0.09 to 0.76 according to the 1000 Genomes Project Database.

Genetic risk factor

After searching from articles for reported polymorphisms, 24 available variants (Supporting Information, Table S4) were then selected when compared with the WES data. It is known from Supporting Information, Table S3 that activation of the renin-angiotensin-aldosterone system (RAAS) and the adrenergic system is closely related to HF. It is now well acknowledged that common variants in genes that encode neurohormonal, adrenergic, and intracellular proteins can modulate clinical consequences of HF patients. According to the formula mentioned previously, GRFs for every patient were summed up ranging from 20.639 to 29.154.
| Characteristics | Total cohort (N = 1000) | Low-risk GRS tertile (N = 336) | Intermediate-risk GRS tertile (N = 337) | High-risk GRS tertile (N = 327) | P  |
|-----------------|------------------------|-------------------------------|----------------------------------------|-------------------------------|----|
| Men, yes        | 743 (74.30%)           | 247 (73.51%)                  | 255 (75.67%)                           | 241 (73.70%)                  | 0.78 |
| Age, years      | 57.00 ± 14.19          | 56.19 ± 14.49                 | 56.71 ± 13.51                          | 58.12 ± 14.51                 | 0.19 |
| SBP, mmHg       | 127.42 ± 24.47         | 128.14 ± 24.60                | 126.92 ± 24.15                         | 127.20 ± 24.72                | 0.83 |
| DBP, mmHg       | 80.65 ± 17.12          | 80.89 ± 16.79                 | 80.91 ± 16.90                          | 80.12 ± 17.84                 | 0.41 |
| NT-proBNP, pg/mL| 3750.00 (1555.00, 8645.00) | 3913.00 (1640.00, 8170.00)   | 3468.50 (1546.75, 8925.00)             | 3782 (1480, 8584)             | 0.28 |
| TC, mmol/L      | 3.91 ± 1.31            | 3.88 ± 0.99                   | 3.86 ± 1.02                            | 4.01 ± 1.81                   | 0.9  |
| TG, mmol/L      | 1.40 ± 1.13            | 1.36 ± 0.89                   | 1.42 ± 0.88                            | 1.41 ± 1.54                   | 0.51 |
| HDL-C, mmol/L   | 0.96 ± 0.31            | 0.97 ± 0.31                   | 0.98 ± 0.32                            | 0.93 ± 0.29                   | 0.46 |
| LDL-C, mmol/L   | 2.42 ± 0.87            | 2.43 ± 0.79                   | 2.39 ± 0.86                            | 2.42 ± 0.95                   | 0.54 |
| NYHA functional class ≤2 | 373 (37.30%) | 135 (40.18%) | 123 (36.50%) | 115 (35.17%) | 0.38 |
| Smoking, yes    | 390 (39.00%)           | 143 (42.56%)                  | 131 (38.87%)                           | 116 (35.47%)                  | 0.52 |
| LVEF (%)        | 34.55 ± 12.40          | 34.00 ± 11.94                 | 34.66 ± 12.82                          | 35.75 ± 12.38                 | 0.25 |
| Diagnosed as DCM, yes | 787 (78.70%) | 270 (80.36%) | 268 (79.53%) | 249 (76.15%) | 0.42 |
| Diagnosed as ICM, yes | 213 (21.30%) | 66 (19.64%) | 69 (20.47%) | 77 (23.55%) | 0.42 |
| Hypertension, yes | 392 (39.20%) | 127 (37.80%) | 132 (39.17%) | 133 (40.67%) | 0.79 |
| Diabetes, yes   | 175 (17.50%)           | 56 (16.67%)                   | 64 (18.99%)                            | 55 (16.82%)                   | 0.72 |
| Hyperlipidaemia, yes | 50 (5.00%)     | 14 (4.17%)                    | 21 (6.23%)                             | 15 (4.59%)                    | 0.42 |
| Prior arrhythmia, yes | 271 (27.10%) | 97 (28.87%) | 92 (27.30%) | 82 (25.08%) | 0.56 |
| Prior stroke, yes | 56 (5.60%)     | 15 (4.46%)                    | 22 (6.53%)                             | 19 (5.81%)                    | 0.50 |
| Beta-blocker use, yes | 435 (43.50%) | 176 (52.38%) | 136 (40.36%) | 123 (37.61%) | 0.0003 |
| ACEI use, yes   | 468 (46.80%)           | 161 (47.92%)                  | 162 (48.07%)                           | 145 (44.34%)                  | 0.54 |

ACEI, angiotensin-converting enzyme inhibitor; DBP, diastolic blood pressure; DCM, dilated cardiomyopathy; GRS, genetic risk score; HDL-C, high-density lipoprotein cholesterol; ICM, ischaemic cardiomyopathy; LDL-C, low-density lipoprotein cholesterol; LVHF, left ventricular ejection fraction; NT-proBNP, N-terminal B-type natriuretic peptide; NYHA, New York Heart Association; SBP, systolic blood pressure; TC, total cholesterol; TG, triglyceride.
Genetic risk score for heart failure

The distribution of GRS calculated using the above 11 SNPs was shown in Supporting Information, Figure S4, ranging from 11.000 to 18.005, with a mean value of 13.913. When divided into tertiles based on GRS, baseline characteristics by genetic risk category based on the GRS are presented in Table 1. Patients with high-risk category were less likely to take beta-blocker. After adjusting for TRF and beta-blocker taking, GRF and GRS were significantly and independently associated with the prognosis of HF [GRS: 95% confidence interval (CI) 1.273–1.739, \( P = 5.78 \times 10^{-7} \); GRF: 95% CI 1.249–1.713, \( P = 2 \times 10^{-6} \)], respectively (Table 2). Kaplan–Meier curves showed a clear pattern towards greater risk of developing worse clinical outcomes for intermediate-risk and high-risk GRS patients (Figure 2). As shown in Table 3, individuals with a high genetic risk had a 2.225-fold increased hazard for the primary endpoints.

### Table 2 Results of univariable and multivariable Cox proportional hazard analysis for cardiac events

| Variables                  | Univariable analysis\(^a\) | Multivariable analysis\(^b\) |
|----------------------------|-----------------------------|-----------------------------|
|                            | \( P \) value  | HR  | 95% CI        | \( P \) value  | HR  | 95% CI        |
| Gender                     | 0.033           | 1.338 | 1.024–1.747  | 0.366           | 1.137 | 0.861–1.503  |
| Age                        | <0.001          | 1.029 | 1.019–1.039  | <0.001          | 1.018 | 1.008–1.028  |
| Hypertension               | 0.468           | 0.911 | 0.708–1.172  | 0.382           | 0.877 | 0.655–1.176  |
| Diabetes mellitus          | 0.036           | 1.371 | 1.021–1.840  | 0.073           | 0.752 | 0.551–1.027  |
| Hyperlipidaemia            | 0.511           | 0.816 | 0.446–1.494  | 0.630           | 0.934 | 0.709–1.232  |
| Smoking                    | 0.809           | 1.031 | 0.804–1.323  | 0.129           | 1.073 | 0.980–1.175  |
| SBP                        | <0.001          | 0.987 | 0.981–0.993  | 0.028           | 0.991 | 0.983–0.999  |
| DBP                        | <0.001          | 0.980 | 0.972–0.988  | 0.158           | 0.992 | 0.981–1.003  |
| LVEF                       | <0.001          | 0.974 | 0.963–0.986  | <0.001          | 0.978 | 0.967–0.990  |
| Beta-blocker use           | <0.001          | 5.657 | 3.960–8.080  | <0.001          | 5.119 | 3.531–7.421  |
| Genetic risk factor        | 1.46 \times 10^{-3} | 1.804 | 1.543–2.109  | 2 \times 10^{-6} | 1.463 | 1.249–1.713  |
| Genetic risk score         | 6.31 \times 10^{-6} | 1.696 | 1.459–1.972  | 5.78 \times 10^{-7} | 1.488 | 1.273–1.739  |

DBP, diastolic blood pressure; LVEF, left ventricular ejection fraction; SBP, systolic blood pressure.

\(^a\)HR and 95% CI, hazard ratios and 95% confidence interval that were calculated with univariate Cox proportional hazard model.

\(^b\)HR and 95% CI, hazard ratios and 95% confidence interval that were calculated with multivariable Cox proportional hazard model.

---

**Genetic risk score for heart failure**

The distribution of GRS calculated using the above 11 SNPs was shown in Supporting Information, Figure S4, ranging from 11.000 to 18.005, with a mean value of 13.913. When divided into tertiles based on GRS, baseline characteristics by genetic risk category based on the GRS are presented in Table 1. Patients with high-risk category were less likely to take beta-blocker. After adjusting for TRF and beta-blocker taking, GRF and GRS were significantly and independently associated with the prognosis of HF [GRS: 95% confidence interval (CI) 1.273–1.739, \( P = 5.78 \times 10^{-7} \); GRF: 95% CI 1.249–1.713, \( P = 2 \times 10^{-6} \)], respectively (Table 2). Kaplan–Meier curves showed a clear pattern towards greater risk of developing worse clinical outcomes for intermediate-risk and high-risk GRS patients (Figure 2). As shown in Table 3, individuals with a high genetic risk had a 2.225-fold increased hazard for the primary endpoints.
(HR = 2.225, 95% CI 1.617–3.062, \( P = 9.03 \times 10^{-7} \)), whereas individuals with intermediate genetic risk had a 1.693-fold increased hazard compared with those with low genetic risk (HR = 1.693, 95% CI 1.214–3.263, \( P = 1.93 \times 10^{-3} \)). After adjusting for TRF and beta-blocker taking, the results were not significantly changed (HR = 2.549, 95% CI 1.790–3.630, \( P = 3.29 \times 10^{-7} \); HR = 1.957, 95% CI 1.356–2.823, \( P = 3.29 \times 10^{-4} \)).

The power of five models in predicting the prognosis of heart failure

We next examined the predictive ability and discriminative power of the following models. Considering that NT-proBNP is a classical prognosis predictor of HF, we constructed five risk prediction models: Model 1, GRS; Model 2, composite of TRF and NT-proBNP; Model 3, composite of TRF, NT-proBNP, and GRS; Model 4, composite of TRF, NT-proBNP, and GRF; and Model 5, composite of TRF, NT-proBNP, GRF, and GRS. The average area under curves of these five models were 0.618 (95% CI 0.587–0.648), 0.723 (95% CI 0.694–0.751), 0.751 (95% CI 0.723–0.778), 0.755 (95% CI 0.727–0.782), and 0.777 (95% CI 0.750–0.803), respectively (Figure 3 and Supporting Information, Table S5). The Model 3 and Model 4 could effectively predict the clinical outcome of HF with similar power, better than the GRS-based model and Model 5 by integrating the TRF information, GRF, and NT-proBNP with the GRS generated a significant increase in the power of predicting outcomes than Model 3 and Model 4 (\( P = 0.0032 \) for Model 3 vs. Model 5, \( P = 0.0029 \) for Model 4 vs. Model 5) (Supporting Information, Tables S5 and S7), indicating the improvement of the accuracy and precision in identifying CHF patients with worse prognosis. Further subgroup analysis stratified by age, gender, or comorbidities was shown in Supporting Information, Tables S6 and S7. Among men, patients over 60, patients with hypertension, patients without diabetes or hyperlipidaemia, the discriminatory power of Model 5 is superior to other models (for men: \( P = 0.0196 \) for Model 3 vs. Model 5, \( P = 0.0478 \) for Model 4 vs. Model 5; for patients over 60: \( P = 0.0126 \) for Model 3 vs. Model 5, \( P = 0.0026 \) for Model 4 vs. Model 5; for patients with hypertension: \( P = 0.0171 \) for Model 3 vs. Model 5, \( P = 0.0318 \) for Model 4 vs. Model 5; for patients without diabetes: \( P = 0.0153 \) for Model 3 vs. Model 5, \( P = 0.0070 \) for Model 4 vs. Model 5; and for patients without hyperlipidaemia: \( P = 0.0059 \) for Model 3 vs. Model 5, \( P = 0.0023 \) for Model 4 vs. Model 5).

Discussion

In this study, three principal observations were highlighted when implementing GRS and GRF into clinical practice. Firstly, we identified 11 common SNPs that were significantly associated with clinical outcome of patients with CHF by multi-omics analysis combining local data of WES and dataset from GEO database, and 24 known disease-causing SNPs from reported articles. Secondly, GRS composed of the 11 SNPs and GRF composed of the 24 SNPs are independent of TRF to assess long-term cardiovascular risk, respectively. Thirdly, integrating GRS and TRF, GRF, and NT-proBNP increased the effectiveness of predicting in clinical outcomes of patients with HF. In the present study, through integrating an ARG set from 2 databases and analysis of the case–control transcriptome sequencing dataset of 64 human left ventricular tissues, 48 ARGs were successfully identified that were...
differentially expressed between failing and normal heart tissue, including 19 down-regulated genes and 29 up-regulated genes. With screening out of loci that both existed in dominant and recessive models as well as with high LD, 11 SNPs were finally confirmed to be significantly associated with the mortality risk of HF. The 11 SNPs were located in the following 7 genes: inositol 1,4,5-trisphosphate receptor type 1 (ITPR1), C9orf72-SMCR8 complex subunit (C9orf72), optineurin (OPTN), autophagy-related 2B (ATG2B), regulatory-associated protein of MTOR complex 1 (RPTOR), heat shock protein 90 alpha family class B member 1 (HSP90AB1), and mitogen-activated protein kinase 7 (MAP2K7). ITPR1 encodes IP3 receptor type 1 (IP3R1) functioning on the endoplasmic reticulum (ER) membrane in a tetrameric form and modulating intracellular calcium homeostasis and signalling. C9orf72 regulates endosomal trafficking, which is associated with 9p-linked ALS (amyotrophic lateral sclerosis) and FTD (frontotemporal dementia). Optineurin is a binding partner for adenoviral protein E3-14.7K utilizing tumour necrosis factor-alpha or Fas-ligand pathways to mediate apoptosis, inflammation, or vasoconstriction. ATG2B is the main regulator in the process of autophagy and is required for elongation and development of the isolation membrane. RPTOR is regulatory-associated protein of mTOR, which can develop cardiac dysfunction and HF. HSP90AB1 encodes a member of the heat shock protein 90 family and is part of the complex in long-term cardiac remodelling fibroblasts. MAP2K7 is found to be an essential role in cardiomyocytes for protecting the heart from hypertrophic insults, thereby preventing the transition to HF. Meanwhile, after searching for reported pathogenic polymorphisms and compared with the WES data, 24 SNPs were finally involved into the construction of GRF. Most of them were related to RAAS or adrenergic system.

Heart failure is an evolving disease that shows different clinical symptoms and prognosis at different stages. Obtaining genetic predictive models to predict the prognosis of HF remains challenging so far. Researches have recently found that autophagy is a key process in the pathogenesis of HF, while little is known about its role in the prognosis of HF. GRS construction using the total 11 SNPs is a new strategy to predict prognosis of HF through autophagy pathway. Our study demonstrated that GRS with HF in individuals is free of TRF and GRF, suggesting a great diagnostic value of clinical application using such a genetic score to identify people who are at risk, which was validated in our current cohort of Chinese Han patients with HF. GRF was built as a prognostic factor, just the same as TRF, which was a part of the multivariable prognostic model.

Furthermore, we evaluated the discrimination and predictive power of five models in our own HF population. Several classical models have been widely used in clinical practice. For example, Seattle HF model that easily captures clinical characteristics can accurately predict the survival of HF.

### Table 3

| Group   | GRS     | Number | Incident cases | P value | Crude HR 95% CI | Adjusted HR 95% CI |
|---------|---------|--------|----------------|---------|----------------|-------------------|
| Lower   | 11.000–13.631 | 336 | 57            | Reference | Reference | Reference |
| Middle  | 13.634–13.988 | 337 | 89           | P = 1.93 × 10–3 | 1.693 | 1.214–2.363 |
| Higher  | 13.988–18.005 | 327 | 112          | 9.03 × 10–7 | 2.225 | 1.617–3.062 |

CI, confidence interval; GRS, genetic risk score.
patients without NT-proBNP and GRFs. MAGGIC risk score was calculated by 13 highly significant independent predictors of mortality including clinical characteristics, laboratory examinations, and medicine taking. The predictive power of the model consists of autophagy-related GRS, TRS, GRF, and NT-proBNP, which had a highest discriminating ability among the five models and can distinguish patients with better or worse prognosis more accurately was comparable with the previous models, suggesting a promising way for HF treatment and prevention. Especially via subgroup analysis, men, patients over 60, and patients with hypertension, patients without diabetes or hyperlipidaemia were further proven can benefit from applying this model. Noting the power of Model 3 and Model 4, we presumed that GRS and GRF might play the roles to a similar extent in the prognosis prediction, which means GRS likely to be clinically applicable due to these reported variants that have been demonstrated to be correlated to the clinical outcomes of HF. Besides, it could be inferred that more risk variants are to be discovered through GWAS or multi-omics analysis; meanwhile, larger sample size or functional experiments are also required for further validation.

However, several limitations of this study need to be considered. Firstly, the inclusion of reported autophagy-related suspicious HF SNPs might be incomplete; more potential loci need to be identified via some ways like expanding the sample of RNA-seq to get more DEGs. Secondly, the time of searching for reported prognosis-related polymorphisms was also limited; wider range of time is needed to obtain more variants. Thirdly, it was a single-centre study; larger studies need to be performed to validate the results. Fourthly, there were racial heterogeneity between RNA-seq samples and WES subjects, thus in need of additional population and can generalize the results to the other population. In addition, a prospective cohort study is warranted in the future to validate our findings of the present study.

In conclusion, we constructed an autophagy-related GRS based on 11 SNPs associated with prognosis of HF patients and a GRF based on 24 SNPs that enable to discriminate the mortality risk of HF. The model composed of GRS, GRF, TRF, and NT-proBNP allowed clinicians to stratify HF individuals into different risks of HF and to provide intensive treatment and early prevention accordingly, leading to a new insight for them into the treatment strategies for different conditions of HF patients in the future.

**Conflict of interest**

None declared.

**Funding**

This study is supported by the key project from the National Natural Science Foundation of China (Nos. 81600236, 81630010, 81770351, 81790624, and 91439203) and the National Key R&D Program of China (No. 2017YFC0909400).
Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1. 48 differentially expressed genes.

Table S2. 149 common variants with Minor Allele Frequency (MAF) > 0.05 in the 48 ARGs.

Table S3. Statistically significant common SNPs in the dominant and recessive models by univariate Cox regression analysis.

Table S4. Genetic variants reported in related to the prognosis in HF patients.

Table S5. AUC for CHF with 5 different models.

Table S6. AUC for CHF subgroups of 5 models.

Table S7. Pairwise comparison of ROC curves in 3 models.

Figure S1. The flow chart of the study.

Figure S2. Screening candidate ARGs through combining transcriptome data and whole exome sequencing data.

Figure S3. LD Structure and Haplotype Blocks of the 13 SNPs.

Figure S4. Histogram of distribution of genetic risk score.

References

1. Yancy CW, Jessup M, Bozkurt B, Butler J, Casey DE Jr, Drazner MH, Fonarow GC, Geraci SA, Horwich T, Januzzi JL, Johnson MR, Kasper EK, Levy WC, Masoudi FA, McBride PE, McMurry JJ, Mitchell JE, Peterson PN, Riegel B, Sam F, Stevenson LW, Tang WH, Tsai EJ, Wilkoff BL. ACCF/AHA guideline for the management of heart failure: a report of the American College of Cardiology Foundation/American Heart Association Task Force on practice guidelines. Circulation 2013, 2014; 128: e240–e327.

2. Komajda M, Charron P, Tesson F. Genetic aspects of heart failure. Eur J Heart Fail 1999; 1: 121–126.

3. Lavine KJ, Greenberg MJ. Beyond genetics-technological advances improving the molecular characterization and precision treatment of heart failure. Heart Fail Rev 2021; 26: 405–415.

4. Ravanap S, Srikumar IF, Talwar P. Autophagy: the spotlight for cellular stress responses. Life Sci 2017; 188: 53–67.

5. Du J, Liu Y, Fu J. Autophagy and heart failure. Ady Exp Med Biol 2020; 1207: 223–227.

6. Nakai A, Yamaguchi O, Takeda T, Higuchi Y, Hikos S, Taniki M, Omiya S, Mizote I, Matsumura Y, Asahi M, Nishida K, Hori M, Mizushima N, Otsu K. The role of autophagy in cardiomyocytes in the basal state and in response to hemodynamic stress. Nat Med 2007; 13: 619–624.

7. Ghosh R, Pattison JS. Macroautophagy and chaperone-mediated autophagy in heart failure: the known and the unknown. Oxid Med Cell Longev. 2018; 2018: 1, 860241–22.

8. Ponikowski P, Voors AA, Anker SD, Bueno H, Cellden JGF, Coats AJS, Falk V, González-Juanatey JR, Harjola VP, Jankowska EA, Jessup M, Linde C, Nikoyannopoulos P, Parisi JT, Piecke B, Riley JP, Rosano GMC, Ruilope LM, Ruschitzka F, Rutten FH, van der Meer P. ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: the Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC). Developed with the special contribution of the Heart Failure Association (HFA) of the ESC. Eur Heart J 2016; 2016; 37: 2129–2200.

9. Dron JS, Hegele RA. The evolution of genetic-based risk scores for lipids and cardiovascular disease. Curr Opin Lipidol. 2019; 30: 71–81.

10. Lee DS, Penna MJ, Benjamin EJ, Wang T, Levy D, O’Donnell CJ, Nam BH, Larson MG, D’Agostino RB, Vasan RS. Association of parental heart failure with risk of heart failure in offspring. N Eng J Med. 2006; 355: 138–147.

11. Lindgren MP, Pirouzifard M, Smith JG, Sundquist J, Sundquist K, Zöller B. A Swedish nationwide adoption study of the heritability of heart failure. JAMA Cardiol. 2018; 3: 703–710.

12. Li S, Sun Y, Hu S, Hu D, Li C, Xiao L, Chen Y, Li H, Cui G, Wang DW. Genetic risk scores to predict the prognosis of chronic heart failure patients in Chinese. J Cell Mol Med 2020; 24: 285–293.

13. Liggert SB, Cresci S, Kelly RJ, Syed FM, Markovitch SJ, Hahn HS, Diwan A, Martini JS, Sparks L, Parekh RR, Sperut JS, Koch WJ, Kardia SL, Dorn GW 2nd. A GRK5 polymorphism that inhibits beta-adrenergic receptor signaling is protective in heart failure. Nat Med. 2008; 14: 510–517.

14. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philpikas AA, del Angel G, Rivas RA, Henn A, McKenna A, Fennell TJ, Kernstyk AM, Sivachenko AY, Cibulskis K, Gabriel SB, Altshuler D, Daly MJ. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet. 2011; 43: 491–498.

15. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009; 25: 1754–1760.

16. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Mort GT, Sherry ST, McVean G, Durbin R. The variant call format and VCFtools. Bioinformatics. 2011; 27: 2156–2158.

17. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007; 81: 559–575.

18. Wang K, Li H, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 2010; 38: e164.

19. Sweet ME, Coccoiolo A, Slavos D, Jones K, Sreet JR, Graw SL, Reece TB, Ambardekar AV, Ristow MR, Mestroni L, Taylor MRG. Transcriptome analysis of human heart failure reveals dysregulated cell adhesion in dilated cardiomyopathy and activated immune pathways in ischemic heart failure. BMC Genomics. 2020; 39: 812.

20. Davies RW, Wells GA, Stewart AF, Erdmann J, Shah SH, Ferguson JF, Hall AS, Anand SS, Burnett MS, Epstein SE, Dandona S, Chen L, Nahrstaedt J, Loley C, König IR, Kraus WE, Granger CB, Engert JC, Hengstenberg C, Wichmann HE, Schreiber S, Tang WH, Ellis SG, Rader DJ, Hazen SL, Reilly MP, Samani NJ, Schukert H, Roberts R, McPherson R. A genome-wide association study for coronary artery disease identifies a novel susceptibility locus in the major histocompatibility complex. Circ Cardiovasc Genet. 2012; 5: 217–225.

21. Börjesson M, Magnusson Y, Hjalmarson A, Andersson B. A novel polymorphism in the gene coding for the beta1-adrenergic receptor associated with survival in patients with heart failure. Eur J Heart Fail. 2000; 21: 1853–1858.

22. Bernardz-Pereira S, Santos PC, Krieger JE, Mansur AJ, Pereira AC. ActN3 R577X polymorphism and long-term survival in patients with chronic heart failure: the known and the unknown. Nat Rev Cardiol. 2013; 10: 559–575.

23. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007; 81: 559–575.
failure. *BMC Cardiovasc. Disord.* 2014; 14: 90.

23. Nonen S, Okamoto H, Fujio Y, Takemoto Y, Yoshiyama M, Hamaguchi T, Matsui Y, Yoshikawa J, Kitabatake A, Azuma J. Polymorphisms of norepinephrine transporter and adrenergic receptor α1β are associated with the response to β-blockers in dilated cardiomyopathy. *Pharmacogenomics J.* 2008; 8: 78–84.

24. Bristow MR, Murphy GA, Krause-Steinrauf H, Anderson JL, Carlquist JF, Thaneemit-Chen S, Krishnan V, Abrahamsen T, Lowes BD, Port JD, Davis RJ, Liu W, Zi M, Chi H, Jin J, Prehar S, Neyses L, Haddad E, Costello C, Zorzato F. A polymorphism within a regulator of renopeptide polymorphism alters the renopeptide-lowering effects and therapeutic response of the beta-blocker bucindolol in chronic heart failure. *Circ Heart Fail.* 2010; 3: 21–28.

25. Liggett SB, Milet-Perez J, Thaneemit-Chen S, Weber SA, Greis P, Hodge D, Nelson B, Morrison J, Domanski MJ, Wagner LE, Abraham WT, Anderson JL, Carlquist JF, Krause-Steinrauf H, Lazzeroni LC, Port JD, Lavori PW, Liggett SB. An α2c-adrenergic receptor polymorphism predicts outcome in African Americans with heart failure: results from the A-HetT trial. *J Am Coll Cardiol.* 2006; 48: 1277–1282.

26. Agra RM, Gago-Dominguez M, Paradela-Dobarro B, Torres-Espanol M, Alvarez L, Fernandez-Trasancos A, Varela-Roman C, Alazza M, Eiras S, Alvarez E, Carracedo A, Gonzalez-Juaneaty JR. Obesity-related genetic determinants of heart failure prognosis. *Cardiovasc Drugs Ther.* 2019; 33: 415–424.

27. Hansen PR, Nelvég-Kristensen KE, Rasmussen HB, Torp-Pedersen C, Kober L, Nielsen CH, Enevold C. Prognostic role of genetic polymorphisms of the interleukin-6 signaling pathway in patients with severe heart failure. *Pharmacogenomics J.* 2019; 19: 428–437.

28. Mizon-Gérard F, de Groote P, Lamblin N, Bauters C, Helbecque N. Prognostic impact of matrix metalloproteinase gene polymorphisms in patients with heart failure according to the aetiology of left ventricular systolic dysfunction. *Eur Heart J.* 2004; 25: 688–693.

29. McNamara DM, Holubkov R, Postava L, Ramani R, Janosko K, Mathier M, MacGowan GA, Murali S, Feldman AM, London B. Effect of the Asp298 variant of endothelial nitric oxide synthase on survival for patients with congestive heart failure. *Cardiovasc Drugs Ther.* 2020; 34: 345–356.

30. Gombos T, Führheiz Z, Pozsonyi Z, Jánoskuti I, Prohászka Z, Karádi I. Long-term survival and apolipoprotein A1 level in chronic heart failure: interaction with tumor necrosis factor α – 308 G/A polymorphism. *J Card Fail.* 2017; 23: 113–120.

31. van der Meer P, De Boer RA, White HL, van der Meer P, De Boer RA, White HL, van Dijk M, Vercruysse M, Varela-Roman C, Alazza M, Eiras S, Alvarez E, Carracedo A, Gonzalez-Juaneaty JR. Obesity-related genetic determinants of heart failure prognosis. *Cardiovasc Drugs Ther.* 2019; 33: 415–424.

32. Fossett JK, White C, Cheung KH, Mak DO. Inositol trisphosphate receptor Ca2+ release channels. *Physiol Rev.* 2007; 87: 593–658.

33. DeJesus-Hernandez M, Mackenzie IR, Boeve BF, Baker X, Baker M, Rutherford NJ, Nicholson AM, Finch NA, Flynnt H, Adamson J, Kouri N, Wojtas A, Sengdy P, Hsiung GY, Karydas A, Seeley WW, Josephs KA, Coppola G, Geschwind DH, Wszolek ZK, Feldman H, Knopman DS, Petersen RC, Miller BL, Dickson DW, Boylan KB, Graff-Radford NR, Rademakers R. Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. *Neuron.* 2011; 72: 245–256.

34. Toth RP, Atkin JD. Dysfunction of optineurin in amyotrophic lateral sclerosis and glaucoma. *Front Immunol.* 2018; 9: 1017.

35. Kishi-Ikaura C, Koyama-Honda I, Itakura E, Mizushima N. Ultrastructural analysis of autophagosome organization using mammalian autophagy-deficient cells. *J Cell Sci.* 2014; 127(Pt 9): 4089–4102.

36. You KS, Yi YW, Kwak SJ, Seong YS. Inhibition of RPTOR overcomes resistance to EGFR inhibition in triple-negative breast cancer cells. *Int J Oncol.* 2018; 52: 828–840.

37. Shende P, Plaisance I, Morandi C, Pelleix C, Berthonneche C, Zorzato F, Krishnan J, Lerch R, Hall MN, Ruegg MA, Pedrazzini T, Brink M. Cardiac raptor ablation impairs adaptive hypertrophy, alters metabolic gene expression, and causes heart failure in mice. *Circulation.* 2011; 123: 1073–1082.

38. García-R, Merino D, Gómez JM, Nistal JF, Hurlé MA, Cortajarena AL, Villar AV. Extracellular heat shock protein 90 binding to TGFβ receptor I participates in TGFβ-mediated collagen production in myocardial fibroblasts. *Cell Signal.* 2016; 28: 1563–1579.

39. Liu W, Zr I, Chi H, Jin J, Prehar S, Neyes L, Cartwright EJ, Flavell RA, Davis RJ, Wang X. Deprivation of MKK7 in cardiomyocytes provokes heart failure in mice when exposed to pressure overload. *J Mol Cell Cardiol.* 2011; 50: 702–711.

40. Levy WC, Mozaffarian D, Linker DT, Sutradhar SC, Anker SD, Cropp AB, Anand I, Maggioni A, Burton P, Sullivan MD, Pitt B, Poole-Wilson PA, Mann DL, Packer M. The Seattle heart failure model: prediction of survival in heart failure. *Circulation.* 2006; 113: 1424–1433.

41. Pocock SJ, Ariti CA, McMurray JJ, Maggioni A, Kober L, Squire I, Swedberg K, Dobson J, Poppe K, Whalley GA, Dougherty RN. Predicting survival in heart failure: a risk score based on 39 372 patients from 30 studies. *Eur Heart J.* 2013; 34: 1404–1413.