InnoVationinfo

Journal of
Biomedical Research and Reviews

ISSN: 2581-7388

Volume 4: 1

J Biomed Res Rev 2021

Novel Signature Genes and Pathways Identified for Human Left Ventricle Cardiomyopathies Rise from different Etiologies

Jiao Tian[1,2†]
Zheng Yuan Wu[1†]
Ying Ying He[*3]
Shubai Liu[*1]

[1]*State Key Laboratory of Phytochemistry and Plant Resources in West China, Kunming Institute of Botany, Chinese Academy of Sciences, PR China*
[2]*School of Life Sciences, Yunnan University, China*
[3]*School of Chemical Science & Technology, Yunnan University, China*
[†]*These authors contributed equally to this work*

Abstract

Cardiomyopathy, a heart disease that arises from different etiologies, places a huge burden on global health care society. Clinical cases of cardiomyopathy were examined independently from the genomic database to identify potential biomarkers and pathways for cardiomyopathy. Exploration of these biopsies as a whole transcription disorder pattern by WGCNA (Weighted Gene Co-expression Network Analysis) to discover signature genes for different cardiomyopathy subtypes. Narrow genes and key pathways correlated with cardiomyopathy traits have been identified through co-expression and protein-protein interaction (PPI) network enrichment analysis. Discovered hub genes have been blast through the Cardiovascular Disease Portal to verify human cardiomyopathy-related functions. Three common axes of signature genes have been identified for five subtypes of cardiomyopathy: 1) Four common genes (MDM4, CFLAR, RPS6KB1, PKD1L2) have been identified for ischemic and ischemic cardiomyopathy subgroups; 2) Subtypes of cardiomyopathy (ischemic, post. partum, familiar and idiopathic) have been shared with eight genes (MAPK1, MAPK11, MAPK14, LMNA, RAC1, PECAM1, XIAP, CREB1); 3) TFAM and RHEB have been identified as the common signature genes for subtypes of cardiomyopathy (viral, post. partum, familiar, and idiopathic) (viral, post. partum, familiar, and idiopathic). Major enriched pathways included the MAPK signaling pathway, the protein processing pathway in the endoplasmic reticulum, and so on. Abnormally regulating these signature genes and pathways caused metabolic process disorders and cellular malfunctions that generally contribute to cardiac dysregulation and functional relapse into cardiomyopathies. In summary, these novel signature genes may work as potential biomarkers for the diagnosis of cardiomyopathy and benefit patients with improved outcomes.

Keywords: Heart failure; Cardiomyopathy; Ventricle cardiomyocytes; WGCNA; Signature gene.

List of Abbreviations: PCA--principal component analysis; GO--gene ontology; TOM--Topological Overlap Matrix ; GS--gene significance; MM--module membership ; ME--module eigengene; MS--module significance; WGCNA--weighted gene co-expression network analysis; ID--idiopathic dilated; IS--ischemic; IDCM--idiopathic cardiomyopathy; FCM--familial cardiomyopathy; PCM--post-partum cardiomyopathy; ISCM--ischemic cardiomyopathy; VCM--viral cardiomyopathy.

Introduction

Cardiomyopathy is characterized by the inability of the heart to pump and/or to fill blood as required by the body, and the worst state

Article Information

Article Type: Research Article
Article Number: JBRR-150
Received Date: 12 May, 2021
Accepted Date: 11 June, 2021
Published Date: 19 June, 2021

***Corresponding author:** Shubai Liu, State Key Laboratory of Phytochemistry and Plant Resources in West China, Kunming Institute of Botany, Chinese Academy of Sciences, 132 Blue Black Street, Kunming, 650201 Yunnan, PR China. Tel: +86 871-65223309; E-mail: liushubai@mail.kib.ac.cn

Ying Ying He, School of Chemical Science & Technology, Yunnan University, Kunming, Yunnan 650091, China. E-mail: yingying.he10@gmail.com

Citation: Liu S, He YY, Wu ZY, Tian J (2021) Novel Signature Genes and Pathways Identified for Human Left Ventricle Cardiomyopathies Rise from different Etiologies. J Biomed Res Rev Vol: 4, Issu: 1. (31-50).

lapses into heart failure, a complex pathophysiological condition with left ventricle myocytes dysfunction. Globally, at least 26 million people have suffered from heart failure and have spent more than $30 billion on health in the world [1]. Mortality of heart failure is as high as ~50% over a five-year period [2]. Although cardiomyopathy is the primary condition for heart failure, pathogenic damages (abnormal physical structure and dysfunction) has often occurred in the patient's cardiomyopathy tissue. Identification of biomarkers may be very useful for early diagnosis of cardiomyopathy, interruption of the disease procession to heart failure, and reduction of the mortality.

Cardiomyopathy etiologies are multiple factors, which make cardiomyopathy a heterogeneous complex cardiovascular disease. Major stimuli include changing physiological conditions (such as pregnancy and delivery) and stressful pathological conditions (for example, ischemia, hypertension, diabetes, viral infection, and so on). Cardiomyopathy is divided into several subgroups, which usually have specific morphological features, according to the distinct etiologies [3]. The most common subtypes are hypertrophic cardiomyopathy (HCM), dilated cardiomyopathy (DCM), viral cardiomyopathy (VCM), familial cardiomyopathy (FCM), post-partum cardiomyopathy (PCM) and ischemic cardiomyopathy (ISCM)(3). Early clinical investigations have identified cardiomyopathic cases caused by abnormal gene expression, without somatic genetic alteration, indicating the impact of epigenetics and transcriptional changes on cardiomyopathy [4]. Complex etiologies, however, can lead to a variety of abnormal expression genes, making it difficult to identify common cardiomyopathy biomarkers with limited number cases.

Computerization methodologies have been applied to the discovery of signature genes as potential biomarkers of diseases. Among many computerization methodologies, the Weighted Gene Co-expression Network Analysis (WGCNA) is a useful approach that analyzes gene expression profiling to find out co-expression network based on their functional features, and to discover the genetic disorders in diseases [5]. WGCNA has been widely used in screening novel biomarkers of cancers and has been demonstrated to be a reliable and powerful tool [6-8]. In this study, WGCNA was employed to analyze the gene expression profiles of several cardiomyopathies, containing 90 human biopsies from GEO databases in NCBI. Genes have been discovered that significantly associated with cardiomyopathies. The signature genes and key biological pathways identified by WGCNA were further validated through bioportal database, an independent annotation database of cardiovascular diseases.

Material and Methods

Screening gene expression datasets of cardiomyopathy

The gene expression dataset of cardiomyopathy used for data analysis were screened from the Gene Expression Omnibus (GEO) database (NCBI). Numbers of patients (more than 20 cases), variety of subgroups of cardiomyopathy, and clearly annotations with clinical trials were the major metrics for screening. A dataset, with a GEO tracking number GSE1145 and a platform entry number GPL570 (https:// www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE114 5), was screened out for further analysis. In this datasets, cardiac transcription profiles were established from cardiomyopathy patients of who were undergoing heart failure and planed for a cardiac transplantation. Human left ventricle samples were collected from biopsies of cardiomyopathy patients, or from "normal" organ donors whose hearts cannot be used for transplants. The heart failure of these cardiomyopathy patients arises from different etiologies (Supplementary Table.1). The transcriptional profiles were measured by Affymetrix Human Genome U133 Plus 2.0 Array. Changes in transcriptional profiles were correlated with the physiologic profile of failure hearts acquired at the time of transplantation. Sample collection and microarray dataset were performed by the cardio-genomics lab, Department of Bauer Center for Genomic Research, Harvard University.

Construction of weighted gene co-expression network

WGCNA package of R (version 1.63) was downloaded (http://www.Rproject.org) and setup by following the protocol previously [9]. Each gene expression value from the downloaded dataset (GSE1145) was normalized by compared with inter reference genes, and performed a log2 transformation. Microarray quality was tested by sample clustering according to the distance between different samples in Pearson's correlation matrices. Outliers were identified with a height cut of 170000. Outliers and samples with excessive missing values were excluded from next analysis. Data quality was checked by the principal component analysis (PCA). The co-expression network was constructed, setting the soft threshold $\beta = 7$, which means that R^2 is equal to 0.98 and indicates the constructed network is close to a scale-free network. The plot of log10 (p(k)) versus log10(k) (Supplementary Figure S2A&B) indicates that the network is close to a scale-free network by using $\beta = 7$, where k is the whole network connectivity and p(k) is the corresponding frequency distribution (Supplementary Table 2). Once β value was determined, the Topological Overlap Matrix (TOM) and dissTOM = 1−TOM were obtained automatically. Modules were identified based on TOM and dissTOM. The hierarchical clustering analysis was used to identify gene modules and color to indicate modules, which is a cluster of densely interconnected genes in terms of co-expression. Genes that were not co-expressed were assigned into a grey module. The significant *p*-value of candidate genes was calculated via T-test. The association between the modules and diseases were evaluated by Gene significance (GS) and Module significance (MS); GS was defined as mediated *p*-value of each gene (GS = lgP); MS was defined as the average GS of all the genes involved in the same module. The cut-off significant standard was setup as *p*-value lower than 0.05. Importance of a gene within a module was measured by the module membership (MM). MM was calculated as MM(i) = cor (xi, ME), where i represents gene contained in module; ME (module eigengene) is defined as the first principal component of the module and represents

the overall expression level of the module [10]. Modules that significantly associated with the traits of different etiologies were identified by calculation the correlation of MEs with clinical pathological features.

Function & Pathway enrichment analysis for gene significance in module

Significant genes that were related to different pathological phenotype were blast through the web-based GenCLiP 2.0. Correlation analysis that biological functions and molecular networks involved with the genes were performed [11] The connection strength of a gene to other genes in a global functional pathway and network was measured by gene connectivity. Genes associated with cardiomyopathy were filtered from the identified significant genes by blast in the Cardiovascular Disease Portal, which contains annotations of 854 genes with verified functions related to human cardiomyopathies [12,13]. Genes that have been reported to link with pathological features were labelled with purple border.

Identification of hub genes

Hub gene in a module, a significant gene that widely connects with other significant genes, is the key interconnected nodes within a functionally network and plays the most important roles in the biological processes [14]. Hub genes were identified by two methods: co-expression network and PPI network analysis. Potential hub genes among each significant module were selected by co-expression network, using GS (> 0.2), MM (> 0.8, with a threshold of *p*-value <0.05), and correlated to certain clinical traits as the screening criteria. In parallel, protein-protein interaction (PPI) network of the module genes were built in a selected significant module through the STRING database. Interaction between genes was defined as positive, as cutoff > 0.4; potential hub gene was filtered in PPI network analysis, if its connectivity degree of ≥ 8 through STRING database. Overlapped potential hub genes in both co-expression network and PPI network analysis were the "real" hub genes.

Validations of expression of hub genes and signature genes

Significant genes were validated by comparison of their expression levels among the subgroups of cardiomyopathy and health group that was used as the benchmark. Individual gene expression in each group was presented as means ± standard error of the mean (SEM). Significance of differences was determined by Student t-test (Prism, GraphPad, San Diego, CA). Differences were significant if $p < 0.05$ (*). When $p < 0.05$, The standard of significance was setup as up-expression (Fold change > 1.0, $p < 0.05$) or down-expression (Fold change < 1.0, $p < 0.05$).

Results

Clinical information of dataset

From GEO database (NCBI), dataset GSE1145 was screened out and chosen for WGCNA analysis. In this dataset, it contained 90 left ventricle biopsies of patients. These biopsy samples represented 8 subgroups: 1, normal hearts, used as control (Health, n=11); 2, idiopathic dilated (ID, n=15); 3, ischemic (IS, n=11); 4, idiopathic cardiomyopathy (IDCM, n=12); 5, familial cardiomyopathy (FCM, n=5); 6, post-partum cardiomyopathy (PCM, n=4); 7, ischemic cardiomyopathy (ISCM, n=20); 8, viral cardiomyopathy (VCM, n=7). Using Affymetrix Human Genome U133 Plus 2.0 Array, the dataset contains expression data of 20,283 target genes for each biopsy sample. Each probe-set was linked with gene symbol through the Affymetrix annotation file GLP570. Besides gene symbols, their related function and clinical traits were also annotated (Figure 1, Supplementary Table 1).

Identified gene modules correlation with pathological traits

The cardiomyopathic and health samples were separated in the PCA plot (Supplementary Figure S1). Four normal samples (GSM18444/18445/18447/18448) were detected as outliers and ignored in the subsequent analysis. The total 86 samples were used for next step analysis (Figure 1). Co-expression networks with different clinical traits were built using the Pearson correlation analysis (Figure 2A). Calculated with a dynamic tree cutting algorithm, the distinct co-expression modules were identified that significantly related to different pathological features (Figure 2B, Figure 3A). Twenty-three modules were detected through the dataset (Figure 2B, Table 1). The numbers of significant genes containing in modules were varied from 121 to 14938. The correlation significance of module and pathological features was determined by module significance (MS) correlation and statistics p-value, and significant module varied from different subtype cases Through calculation of the linear mixed-effects model, significant modules were identified for specific pathological feature (Figure 3B). The higher value of module eigengene (ME) correlation, the module is closer correlated to cardiomyopathies (Supplementary Figure S3-9). In the idiopathic dilated group, six modules were significantly associated with its status, including turquoise module (t-value = 0.75, p-value = 5e−17), light-cyan module (t-value = 0.54, p-value = 7e−08), tan module (t-value = 0.56, p-value = 2e−08), grey module (t-value = 0.7, p-value = 8e−14), green module (t-value = 0.45, p-value = 1e−05) and the blue module (t-value = 0.72, p-value = 8e−15) (Figure 2B, Figure 3B, Table 1). In ischemic group, six significant modules were identified, including light-yellow module (t-value = 0.43, p-value = 3e−05), green-yellow module (t-value = 0.41, p-value = 8e−05), light-green module (t-value = 0.4, p-value = 1e−04), red module (t-value = 0.35, p-value = 8e−04), green module (t-value = 0.24, p-value = 0.03) and the blue module (t-value = 0.44, p-value = 2e−05) (Figure 2B, Figure 3B, Table 1). In idiopathic cardiomyopathy (IdCM) group, three modules were significantly linked to pathological trait, listed as module of Magenta (t-value = 0.43, p-value = 4e−05), Purple (t-value = 0.31, p-value = 0.003) and Brown (t-value = 0.25, p-value = 0.02). In familial cardiomyopathy group, only Magenta module (t-value = 0.22, p-value = 0.04) was significantly correlated to this trait (Figure 2B, Figure 3B). For the Hypertrophic cardiomyopathy (HCM) group, none of module was identified and ignored in next step analysis (Figure 2B). In the Post. Partum cardiomyopathy (PCM) group, Magenta module (t-value = 0.24, p-value=0.03)

	Case type	Idiopathic Dilated		Ischemic		Idiopathic-CM		Familial-CM		Hypertrophic-CM		Post.Partum-CM		Ischemic-CM		Viral-CM	
Module	**Genes Size**	**Corr**	**p value**	**Corr**	**p value**	**Corr**	**p value**	**Corr**	**p value**	**Corr**	**p value**	**Corr**	**p value**	**Corr**	**p value**	**Corr**	**p value**
Royalbule	209	-0.28	4.00E-05	-0.12	8.40E-02	-0.22	1.40E-03	-0.27	7.70E-05	-0.33	1.10E-06	-0.28	4.00E-05	-0.17	0.014	-0.25	2.60E-04
Darkgreen	121	-0.17	6.20E-02	0.11	2.30E-01	0.22	1.50E-02	-0.27	2.70E-03	-0.23	1.10E-02	0.061	5.10E-01	-0.16	0.08	0.25	5.70E-03
Black	1099	0.43	4.00E-05	0.23	8.00E-03	-0.01	8.40E-01	-0.14	3.20E-06	-0.05	7.90E-02	-0.2	2.20E-11	**0.71**	**2.80E-169**	-0.15	5.90E-07
Grey60	355	-0.25	1.80E-06	-0.29	2.60E-08	0.48	7.40E-22	-0.10	6.00E-02	-0.35	1.10E-11	-0.19	3.20E-04	-0.12	2.40E-02	-0.26	6.80E-07
Midnightblue	447	0.11	1.00E-03	0.50	1.20E-29	0.02	7.40E-01	0.07	1.70E-01	-0.04	4.40E-01	-0.058	2.20E-01	**0.39**	**1.10E-17**	-0.12	1.10E-02
Greenyellow	759	0.039	2.80E-01	**0.34**	**8.00E-05**	0.35	2.70E-23	0.14	1.10E-04	0.09	1.20E-02	0.14	1.10E-04	0.11	0.0024	0.23	1.40E-10
Lightyellow	256	0.088	1.60E-01	**0.13**	**3.00E-05**	0.09	1.40E-01	-0.10	1.30E-01	0.08	2.10E-01	-0.15	1.60E-02	0.15	0.016	-0.052	4.10E-01
Pink	1007	0.19	1.20E-09	-0.088	5.20E-03	-0.18	8.80E-09	-0.18	8.80E-09	-0.02	4.50E-01	-0.1	1.50E-03	**0.30**	**7.00E-03**	-0.097	2.10E-03
Red	1301	0.44	2.00E-05	**0.30**	**8.00E-04**	-0.20	3.30E-13	-0.11	7.00E-05	0.00	9.70E-01	-0.16	6.50E-09	**0.18**	**6.20E-11**	-0.13	2.50E-06
Yellow	1881	0.64	2.00E-09	-0.02	4.10E-01	-0.09	6.40E-05	-0.08	3.10E-04	-0.09	1.10E-04	-0.18	3.70E-15	**0.63**	**1.00E-08**	-0.093	5.40E-05
Cyan	616	0.73	5.00E-07	0.29	2.10E-13	0.32	3.90E-16	0.17	2.20E-05	-0.09	2.70E-02	0.037	3.60E-01	**0.52**	**2.00E-03**	**0.39**	**8.20E-24**
Darkred	143	0.59	2.00E-07	0.17	4.20E-02	0.03	7.70E-01	-0.04	6.00E-01	-0.07	3.90E-01	-0.3	2.70E-04	**0.67**	**3.00E-05**	0.061	4.70E-01
Brown	2539	0.047	1.80E-02	-0.07	4.20E-04	0.26	1.70E-40	-0.01	7.20E-01	-0.09	1.80E-05	0.089	7.10E-06	0.71	<1e-200	-0.085	1.80E-05
Magenta	945	0.23	5.00E-07	0.23	8.20E-13	**0.65**	**1.50E-114**	**0.31**	**1.70E-22**	-0.14	1.60E-05	**0.30**	**4.20E-21**	0.098	2.60E-03	**0.16**	**7.70E-07**
Purple	781	0.29	2.00E-04	0.06	1.20E-01	**0.56**	**1.10E-65**	0.31	7.40E-19	0.04	2.90E-01	0.17	1.80E-06	0.097	6.70E-03	0.21	3.10E-09
Salmon	643	-0.19	1.20E-06	-0.12	2.30E-03	-0.04	2.90E-01	-0.03	4.50E-01	-0.05	1.80E-01	-0.089	2.40E-02	-0.084	3.30E-02	-0.014	7.20E-01
Lightgreen	317	-0.054	3.40E-01	**0.41**	**0.00**	-0.07	2.20E-01	-0.05	3.90E-01	-0.17	2.40E-03	-0.088	1.20E-01	-0.025	6.60E-01	-0.12	3.30E-02
Tan	672	**0.44**	**2.00E-08**	-0.22	8.30E-09	-0.01	7.40E-01	-0.23	1.60E-09	-0.21	3.90E-08	-0.24	2.90E-10	0.048	2.10E-01	-0.15	9.50E-05
Lightcyan	358	**0.63**	**7.00E-08**	-0.12	2.30E-02	0.08	1.50E-01	-0.06	2.80E-01	-0.14	8.00E-03	-0.18	6.20E-04	0.27	2.10E-07	-0.072	1.70E-01
Turquoise	11003	**0.77**	**<1e-200**	-0.10	1.70E-23	0.12	1.40E-36	0.00	8.00E-01	-0.14	2.80E-49	-0.14	2.80E-49	0.3	3.00E-04	-0.13	1.10E-42
Blue	3898	**0.71**	**8.00E-15**	**0.33**	**0.00**	0.31	1.40E-87	0.16	9.10E-24	0.05	1.40E-03	0.035	2.90E-02	0.55	2.00E-06	0.17	1.20E-26
Green	1484	**0.33**	**1.00E-05**	**0.12**	**0.00**	0.10	1.30E-04	-0.01	7.90E-01	-0.17	4.40E-11	-0.11	2.20E-05	0.23	2.90E-19	-0.11	2.20E-05
Grey	14938	**0.67**	**8.00E-14**	-0.12	4.90E-49	0.05	9.70E-10	-0.01	2.30E-01	-0.063	1.30E-14	-0.055	1.70E-11	0.21	1.00E-03	-0.046	1.90E-08
Significant Module		Turquoise		Light yellow		Magenta		Magenta		N/A		Magenta		Yellow		Cyan	

Table 1: Correlations between significant modules with different etiologies for cardiomyopathies.

ISSN: 2581-7388

Gene Expression Profiling of Cardiomyopathies Induced by Different Etiologies

Case Groups	Label	Organism	Resource	Disease state	Case No.
Health	Ctrl	Human	Left Ventricle Biopsy	Normal	7
Idiopathic Dilated	ID	Human	Left Ventricle Biopsy	Cardiomyopathy	15
Ischemic	IS	Human	Left Ventricle Biopsy	Cardiomyopathy	11
Idiopathic Cardiomyopathy	IDCM	Human	Left Ventricle Biopsy	Cardiomyopathy	12
Familial Cardiomyopathy	FCM	Human	Left Ventricle Biopsy	Cardiomyopathy	5
Hypertrophic Cardiomyopathy	HCM	Human	Left Ventricle Biopsy	Cardiomyopathy	5
Post-partum Cardiomyopathy	PCM	Human	Left Ventricle Biopsy	Cardiomyopathy	4
Ischemic Cardiomyopathy	ISCM	Human	Left Ventricle Biopsy	Cardiomyopathy	20
Viral Cardiomyopathy	VCM	Human	Left Ventricle Biopsy	Cardiomyopathy	7

(reference: GSE1145)

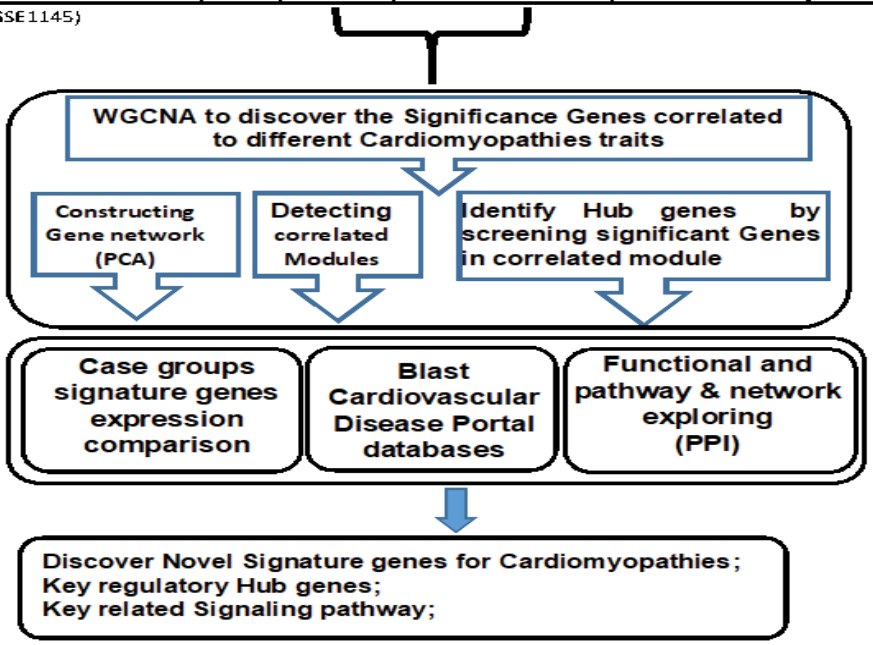

Figure 1: Work flow of the different etiologies caused cardiomyopathies biopsy samples for WGCNA.

was correlated to its trait (Figure 2B, Figure 3B, Table 1). In ischemic cardiomyopathy group, seven modules were significantly correlated to pathological trait, including Black (t-value = 0.56, p-value = 2e−08), Midnight-blue (t-value = 0.49, p-value = 2e−06), Dark-red (t-value = 0.43, p-value = 3e−05), Cyan (t-value = 0.34, p-value = 0.02), Yellow (t-value = 0.57, p-value = 1e−08), Pink (t-value = 0.29, p-value = 0.007) and Red (t-value = 0.26, p-value = 0.02) (Figure 3B). For Viral cardiomyopathy (VCM), two significant modules, Cyan (t-value = 0.35, p-value=0.001) and Magenta (t-value = 0.23, p-value = 0.03), were identified (Figure 2B, Figure 3B).

Enriched genes significance related to different cardiomyopathies

Compared the module memberships (MM) correlation and Genes Significance (GS) among the all significant modules, the module with most significant value was defined as the best candidate for pathological traits correlation analysis (Table 1, Figure 3B), respectively. These candidates were listed as turquoise module (cor=0.77, p<1.0e−200, GS=0.4244) (Idiopathic dilated group, Supplementary Figure S3), light-yellow module (cor=0.13, p=3.0e−05, GS=0.2733) (Ischemic group, Supplementary Figure S4, Supplementary Table 3), magenta module (cor=0.65, p=1.5e−114, GS=0.2783) (Idiopathic cardiomyopathy group, Supplementary Figure S5, Supplementary Table 3), magenta module (cor=0.31, p=1.7e−22, GS=0.1502) (Familial cardiomyopathy group, Supplementary Figure S6, Supplementary Table 3), magenta module (cor=0.30, p=4.2e−21, GS=0.1583) (Post. partum cardiomyopathy group, Supplementary Figure S7, Supplementary Table 3), yellow module (cor=0.63, p<1.0e−200, GS=0.3259) (Ischemic cardiomyopathy group, Supplementary Figure S8, Supplementary Table 3) and cyan module (cor=0.39, p=8.2e−24, GS=0.2245) (Viral cardiomyopathy group, Supplementary Figure S9, Table 3. Supplementary Table 3), respectively. In addition, the scatter plot of multiple module memberships (MM) was plotted against the Genes Significance (GS) in each significant module, and the point was represented each gene contained in a module.

Hierarchical clustering of eigengene profiles with cardiomyopathies traits

Based on the ME's values, the hierarchical clustering was performed between all modules and different cardiomyopathies traits to identify their relationships. Furthermore, the eigengene dendrogram analysis was performed to build the correlation of candidate module with

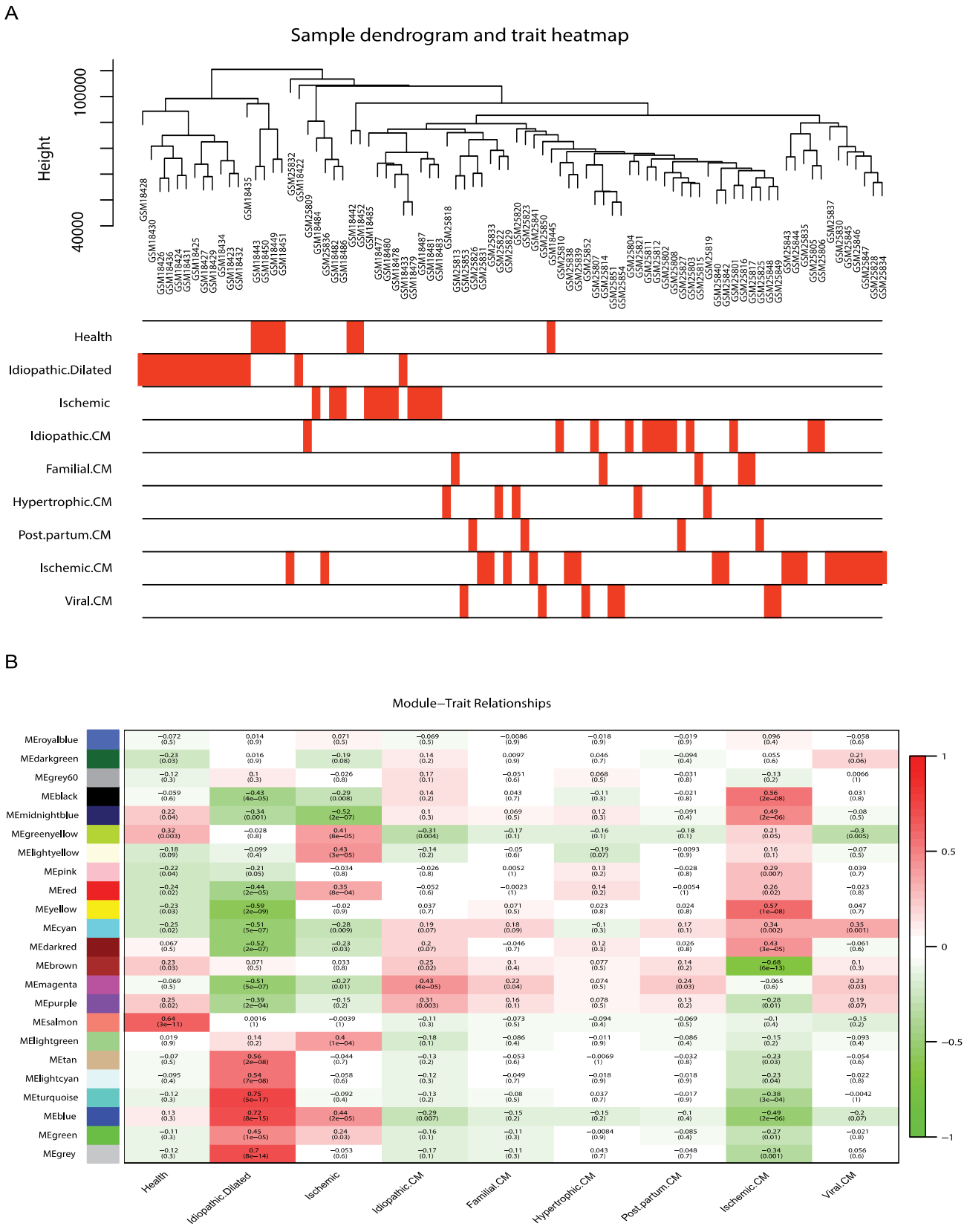

Figure 2: Sample dendrogram and Clinic Feature traits heatmap**. A.** Clustering dendrogram of samples based on their Euclidean distance. The different etiologies caused cardiomyopathies (CM) were idiopathic dilated (ID, n=15), ischemic (IS, n=11), idiopathic cardiomyopathy (IdCM, n=12), familial cardiomyopathy (FCM, n=5), post-partum cardiomyopathy (PCM, n=4), ischemic cardiomyopathy (IsCM, n=20) and viral cardiomyopathy (VCM, n=7). The white color means a low value, red means a high value. B. The correlation heatmap of Module-Cardiomyopathy etiologies traits. Each row corresponds to a module; each column corresponds to a different etiologies feature. Identification of modules significant associated with clinical features and Clustering dendrogram of genes, with dissimilarity based on topological overlap, is merged with assigned module colors. Grey bars represent Genes that do not belong to any other modules and are not co-expressed. Each cell contains the test statistic value and its corresponding p value from the linear mixed-effects model. CM is the abbreviation of Cardiomyopathy.

different subtypes of cardiomyopathy feature, respectively (Supplementary Figures S10-16). In the idiopathic dilated group, turquoise module was tightly clustered with idiopathic dilated (Supplementary Figure S10). In the ischemic group, green yellow and light yellow modules were the closest branch clustered with ischemic (Supplementary Figure S11). In previous step analysis, green-yellow module (t-value = 0.41, p-value = 8e−05, GS = 0.2418) was the secondary higher correlation with ischemic status (Figure 2B, Figure 3B). It suggests that genes containing in green-yellow module would be involve the progress of ischemic. In idiopathic cardiomyopathy group, modules of brown, magenta and purple were clustered with idiopathic cardiomyopathy in a separate branch, and magenta and purple module allocated in same cluster (Supplementary Figure 12). It suggests that module of purple and magenta would be the top two significant associated with cardiomyopathy status (Figure 2B, Table 1). In the familial cardiomyopathy and post. partum cardiomyopathy groups, modules of brown, magenta and purple were tightly clustered with familiar cardiomyopathy, while the magenta module was the most significant associated with disease status (Supplementary Figures S13-14). In the hypertrophic cardiomyopathy group, no module was associated with its pathological feature. In the ischemic cardiomyopathy group, although module of black and midnightblue were clustered in closer branch, the yellow module was allocated in the adjacent branch (Supplementary Figure S15). Combined with the module-trait relationship correlation and gene significance results, it suggested that module of black, midnightblue and yellow were significant associated with ischemic cardiomyopathy (Figure 2B, Figure 3B, Table 1). The module of brown and blue were ignored for next analysis as with higher negative GS values (Table 1, Figure 2B). The yellow module was the most significant correlation to ischemic cardiomyopathy. In the viral cardiomyopathy group, the modules of magenta, purple and brown were clustered with viral cardiomyopathy in a separated branch, and cyan module had the highest GS value associated with pathological feature (t-value = 0.35, p-value = 1e−03, GS = 0.2245) (Figure 2B, Figure 3B, Supplementary Figure S16). It suggested that magenta module containing genes involve the progress of viral cardiomyopathy.

Functions and pathways enrichment analysis

Blast through GenClip2, the enrichments analysis of significance genes contained in the correlated module were summarized as bar chart, including Biological Process, Molecular Function and Cellular Component (Ischemic, Supplementary Figure S17A; Idiopathic Cardiomyopathy, Supplementary Figure S18A; Familial Cardiomyopathy, Supplementary Figure S19A; Post-Partum Cardiomyopathy, Supplementary Figure S20A; Ischemic Cardiomyopathy, Supplementary Figure S21A; Viral Cardiomyopathy, Supplementary Figure S22A). The Biological Processes were mainly concentrated in subgroups of cellular macromolecular metabolic process, protein metabolic process, organic substance metabolic process and macromolecule modification. The genes significant enriched in molecular functions were summarized and listed (Supplementary Table 4). The molecular functions were linked to endoplasmic reticulum functions, cell functions (migration, death, growth, division), DNA binding, protein-protein interaction, kinases activity and signal transduction, iron binding and nucleotide binding, etc. The cellular components were mainly enriched in membrane-bounded organelle, intracellular organelle part, etc. The pathways concentrated on mitogen-activated protein kinase (MAPK) signaling pathway, protein processing in endoplasmic reticulum, regulation of actin cytoskeleton, etc. These results suggested that dysregulation of cardiac functions would be associated with metabolism abnormal and accelerated progress of cardiomyopathies. Furthermore, through gene network literature mining and clustering analysis, these significance genes were clustered and labelled according to cellular functions and pathological feature keywords literature corresponding (Ischemic, Supplementary Figure S17; Idiopathic Cardiomyopathy, Supplementary Figure S18; Familial Cardiomyopathy, Supplementary Figure S19; Post-Partum Cardiomyopathy, Supplementary Figure S20; Ischemic Cardiomyopathy, Supplementary Figure S21; Viral Cardiomyopathy, Supplementary Figure S22).

Moreover, the network and connectivity of significance genes were identified as node-connection map, which was correlated to different traits (Ischemic, Supplementary Figure S17-B; Idiopathic Cardiomyopathy, Supplementary Figure S18B; Familial Cardiomyopathy, Supplementary Figure S19B; Post-Partum Cardiomyopathy, Supplementary Figure S20B; Ischemic Cardiomyopathy, Supplementary Figure S21B; Viral Cardiomyopathy, Supplementary Figure S22B). The key regulatory genes were labelled with purple border, which was reported in pathological cases, and majority of genes involved the pathway were labelled as un-reported. More interesting, no related key regulatory gene was identified in Post-Partum cardiomyopathy group.

Identification of hub genes for cardiomyopathies

Through co-expression network (MM-GS) filtered, the candidates of hub gene were identified in different groups (5 genes in Ischemic group, figure 4A; 113 genes in Idiopathic Cardiomyopathy group, figure 4B; 41 genes in Familiar Cardiomyopathy group, figure 4C; 65 genes in Post-Partum Cardiomyopathy group, figure 4D; 83 genes in Ischemic Cardiomyopathy group, figure 4E; 60 genes in Viral Cardiomyopathy group, figure 4F).

Calculated by the PPI network method, the candidates of hub gene were summarized (12 genes in Ischemic group, Figure 4A; 120 genes in idiopathic cardiomyopathy group, Figure 4B; 277 genes in Familiar Cardiomyopathy group, Figure 4C; 119 genes in Post-Partum Cardiomyopathy group, Figure 4D; 348 genes in Ischemic Cardiomyopathy group, Figure 4E; 49 genes in Viral Cardiomyopathy group, Figure 4F). These real hub genes were determined as described in method section (Table 2), and the numbers of real hub genes were listed (Figure 4A-F). The Idiopathic Dilated group was dismissed for further analysis as no identified real hub gene. Through Venn diagrams analysis, three common axes of hub genes were discovered among these cardiomyopathy groups (Figure 4G).

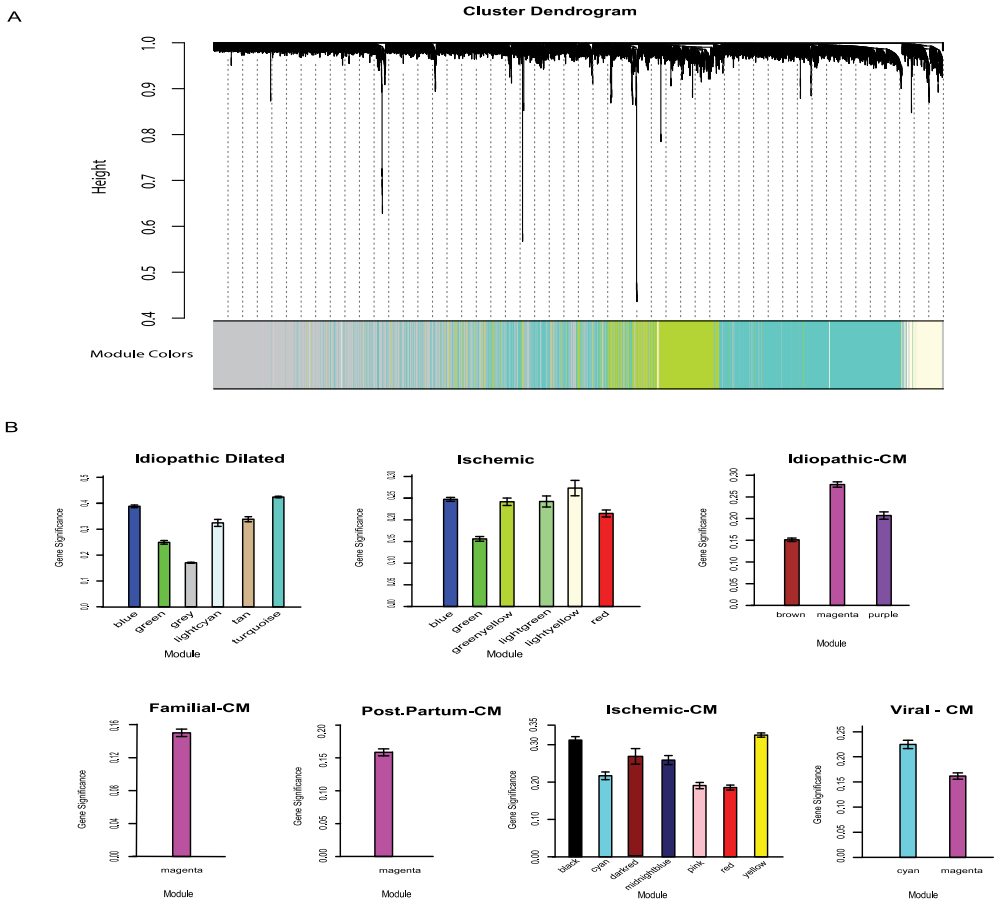

Figure 3: Gene dendrogram and module membership for Clinic Feature traits. (A). Dendrogram of all differentially expressed genes clustered based on a dissimilarity measure. (B). The total Module memberships vs. gene significance cluster for etiologies trait. Distribution of average gene significance and errors in the significant modules associated with the different etiologies.

The first axis was PICALM, which shared by Ischemic Cardiomyopathy, Idiopathic Cardiomyopathy and Post. Partum Cardiomyopathy groups (Figure 4G), and significantly up-regulated in Idiopathic Cardiomyopathy and Ischemic Cardiomyopathy groups (Table 3). PICALM is key regulator in iron homeostasis, clathrin-mediated endocytosis [15,16]. Overexpression of PICALM impaired endocytosis of Transferrin (Tf) Receptor (TfR) and Epidermal Growth Factor Receptor (EGFR) and disturbed the iron homeostasis [16,17]. Up to now, it is still illusive that the exactly role and deregulatory mechanism of PICALM in cardiomyopathies. It is strongly suggesting that PICALM work as potential novel biomarker and therapy target for these subcases of cardiomyopathies.

The secondary axis, contained genes of PRKACB, MOB1A, CDC40, were shared in Post. Partum Cardiomyopathy and Idiopathic Cardiomyopathy groups. In addition, these genes (PRKACB, MOB1A, CDC40) were significantly overexpressed in Idiopathic cardiomyopathy group, and MOB1A was up-regulated in Post. Partum cardiomyopathy group (Table 3).These genes were linked to the cAMP (cyclic AMP)-dependent protein kinase A (PKA) mediated the exciting-contraction coupling in cardiomyocytes, and regulated microtubule stability, cell cycle and cell proliferation & migration, and restrained cardiomyocyte proliferation and size via Hippo pathway [18,19]. *PRKACB* (protein kinase cAMP-activated catalytic subunit β gene) was linked to congenital heart defect with abnormal over-expression [20]. MOB1A (MOB kinase activator 1A) was required for cytokinesis through regulating microtubule stability. It worked as binding partners as well as co-activators of Ndr family protein kinases and mediated phosphor-recognition in core Hippo pathway that restrains cardiomyocyte proliferation during development to control cardiomyocyte size [18,19]. Overexpression of MOB1A induces centrosomes fail to split and cell size dysregulation [21]. CDC40 (Cell Division Cycle 40), a splicing factor of cell division cycle 40 homolog, regulates cell cycle and cell proliferation and migration [22]. Overexpression of CDC40 causes abnormally cell proliferation and migration, and linked with carcinogenesis.

The third axis consisted of five genes (CREB1, DBT, NCOA2, NUDT21, PIK3C2A) and were overlapped among three groups of Familial / Idiopathic / Post. Partum Cardiomyopathy (Figure 4G). The CREB1 (cAMP-responsive element-binding protein) had been identified as the transcription factor and mediated cAMP stimulation by multiple extracellular signals, such as growth factors and hormones. The CREB1 was the key regulator in heart and linked with heart disease via cAMP-PKA pathway dysregulation [23,24]. The DBT (dihydrolipoamide branched chain transacylase E2) is an

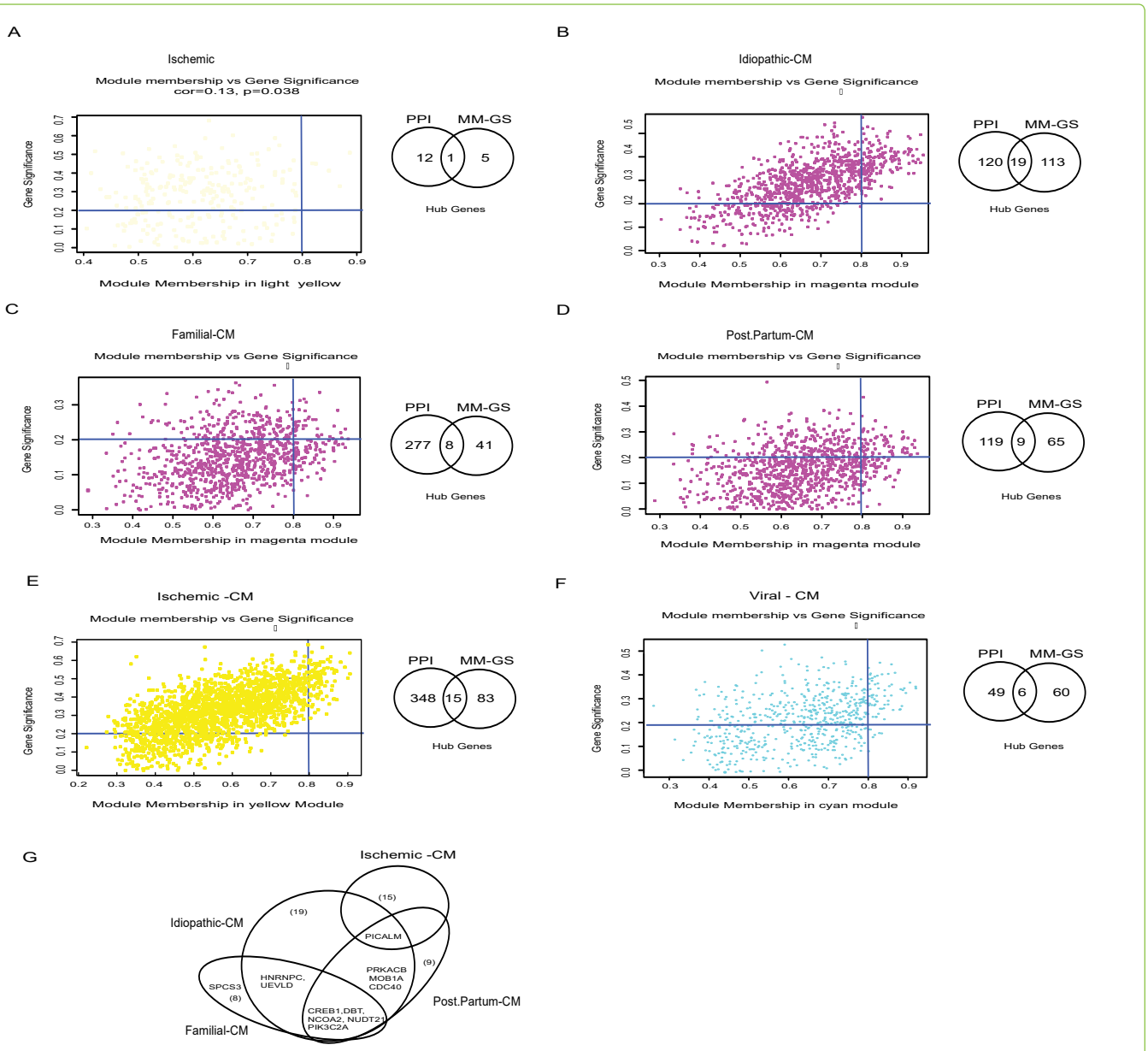

Figure 4: The scatterplots of Gene Significance (GS) for etiologies vs. Module Membership (MM) in the all cases. There is a highly significant correlation between GS and MM in this module that the most important (central) elements of module tend to be highly correlated with cardiomyopathies. Module membership vs gene significance is correlating to different etiologies for cardiomyopathies. Hub genes detected by GS-MM and protein-protein network (PPI). (A) – (F) Scatter plot of module eigengenes in module for all cases, and the real hub genes indicated by Venn diagram for co-expression and PPI network. (G). The Venn diagram showed the overlapped real hub for different cardiomyopathies.

inner-mitochondrial enzyme complex regulated to degrade the branched-chain amino acids isoleucine, leucine, and valine [25]. The DBT was reported as clinical diagnostics biomarker for patients with dilated cardiomyopathy via caused mitochondria dysfunction [26]. NCOA2 (nuclear receptor coactivator 2) is a transcriptional coactivator that functional aid for nuclear hormone receptors, including steroid, thyroid, retinoid, and vitamin D receptors. *NCOA2* promotes muscle cells maintenance and growth, eventually regulates in cardiac cTnT levels [27,28]. Overexpression of NCOA2 regulated cell proliferation in cardiomyopathy. NUDT21 (nudix hydrolase 21) is a novel of cell fate regulator by alternative polyadenylation chromatin signaling, and suppression of NUDT21 will enhance the cell pluripotent, facilitated trans-differentiation into stem cell. NUDT21 regulates cell proliferation through ERK pathway [29]. Up to now, little knows about the function of NUDT21 in cardiomyocytes. PIK3C2A (phosphatidylinositol-4-phosphate 3-kinase catalytic subunit type 2 alpha) is an enzyme belong to phosphorylate the 3'-OH of inositol ring of phosphatidylinositol (PI) superfamily and regulates multiple signaling pathways. PIK3C2A is mainly expressed in endothelial cells, vascular endothelium, and smooth muscle [30]. Lower expression of PIK3C2A in peripheral blood was used as significant biomarker for acute myocardial infarction patients [31]. More interesting, these hub genes indicated different expression pattern. The expression level of DBT, NCOA2, NUDT21 and PIK3C2A were significantly

Traits	Module	GeneSymbol	GS.Ischemic	p.GS.Ischemic	MM.lightyellow	p.MM.lightyellow
Ischemic	lightyellow	HNRNPA2B1	0.4496	1.41E-05	0.8263	1.18E-22
Idiopathic-CM	magenta	ACTR2	0.2938	6.05E-03	0.8039	1.19E-20
Idiopathic-CM	magenta	CDC40	0.3870	2.32E-04	0.8170	8.67E-22
Idiopathic-CM	magenta	CREB1	0.2563	1.72E-02	0.8108	3.09E-21
Idiopathic-CM	magenta	DBT	0.4203	5.60E-05	0.8300	5.22E-23
Idiopathic-CM	magenta	HNRNPC	0.4916	1.54E-06	0.8520	2.52E-25
Idiopathic-CM	magenta	MCL1	0.3203	2.64E-03	0.8102	3.46E-21
Idiopathic-CM	magenta	MOB1A	0.3836	2.66E-04	0.8790	9.54E-29
Idiopathic-CM	magenta	NCOA	0.3836	1.93E-03	0.8740	4.67E-28
Idiopathic-CM	magenta	NRAS	0.3464	6.10E-04	0.8112	2.86E-21
Idiopathic-CM	magenta	NUDT21	0.4863	2.06E-06	0.9332	4.37E-39
Idiopathic-CM	magenta	PICALM	0.3291	1.98E-03	0.8621	1.61E-26
Idiopathic-CM	magenta	PIK3C2A	0.3638	5.75E-04	0.9048	6.85E-33
Idiopathic-CM	magenta	PRKACB	0.3670	5.10E-04	0.9329	5.05E-39
Idiopathic-CM	magenta	RHOA	0.2580	1.65E-02	0.8039	1.19E-20
Idiopathic-CM	magenta	SRSF10	0.2282	3.46E-02	0.5498	4.18E-08
Idiopathic-CM	magenta	SRSF11	0.3086	3.84E-03	0.6696	1.80E-12
Idiopathic-CM	magenta	UBE2D3	0.2317	3.18E-02	0.5657	1.38E-08
Idiopathic-CM	magenta	UEVLD	0.3228	2.44E-03	0.8330	2.67E-23
Idiopathic-CM	magenta	USO1	0.3931	1.81E-04	0.8615	1.91E-26
Familial-CM	magenta	CRE 1	0.2286	3.42E-02	0.8108	3.09E-21
Familial-CM	magenta	DBT	0.3175	2.90E-03	0.8300	5.22E-23
Familial-CM	magenta	HNRNP	0.2586	1.62E-02	0.8520	2.52E-25
Familial-CM	magenta	NCOA2	0.2154	4.64E-02	0.8740	4.67E-28
Familial-CM	magenta	NUDT21	0.2316	3.19E-02	0.9332	4.37E-39
Familial-CM	magenta	PIK3C2	0.2401	2.60E-02	0.9048	6.85E-33
Familial-CM	magenta	SPCS3	0.2129	4.91E-02	0.8073	6.17E-21
Familial-CM	magenta	UEVLD	0.2837	8.13E-03	0.8330	2.67E-23
Post.partum CM	magenta	CDC40	0.2346	2.97E-02	0.8170	8.67E-22
Post.partum CM	magenta	CREB1	0.2712	1.16E-02	0.8108	3.09E-21
Post.partum CM	magenta	DBT	0.2183	4.35E-02	0.8300	5.22E-23
Post.partum CM	magenta	MOB1A	0.2116	5.05E-02	0.8790	9.54E-29
Post.partum CM	magenta	NCOA2	0.2853	7.75E-03	0.8740	4.67E-28
Post.partum CM	magenta	NUDT21	0.2405	2.57E-02	0.9332	4.37E-39
Post.partum CM	magenta	PICALM	0.2572	1.68E-02	0.8621	1.61E-26
Post.partum CM	magenta	PIK3C2A	0.2777	9.64E-03	0.9048	6.85E-33
Post.partum CM	magenta	PRKACB	0.2681	1.26E-02	0.9329	5.05E-39
Ischemic CM	yellow	BUB3	0.6406	3.07E-11	0.9073	2.37E-33
Ischemic CM	yellow	CCNC	0.5346	1.15E-07	0.8010	2.06E-20
Ischemic CM	yellow	CHD2	0.4445	1.81E-05	0.8135	1.78E-21
Ischemic CM	yellow	DNAJB14	0.5254	2.06E-07	0.8225	2.71E-22
Ischemic CM	yellow	DNAJC10	0.4687	5.32E-06	0.8324	3.03E-23
Ischemic CM	yellow	DNAJC10	0.4581	9.21E-06	0.7051	3.52E-14
Ischemic CM	yellow	DNAJC9	0.6701	1.71E-12	0.8357	1.40E-23
Ischemic CM	yellow	FMR1	0.3898	2.08E-04	0.8669	3.98E-27
Ischemic CM	yellow	LAMP2	0.5934	1.73E-09	0.8560	8.70E-26
Ischemic CM	yellow	LAMP2	0.3889	2.15E-04	0.6455	1.96E-11
Ischemic CM	yellow	LAMP2	0.3384	1.44E-03	0.6166	2.61E-10
Ischemic CM	yellow	PICALM	0.5353	1.09E-07	0.8013	1.96E-20
Ischemic CM	yellow	RAB11A	0.4722	4.43E-06	0.8183	6.65E-22
Ischemic CM	yellow	RARS2	0.5297	1.57E-07	0.8373	9.69E-22

Ischemic CM	yellow	TAF2	0.5409	7.58E-08	0.8373	6.30E-28
Ischemic CM	yellow	TMX3	0.5732	8.03E-09	0.8283	7.60E-23
Ischemic CM	yellow	U2SURP	0.5857	3.16E-09	0.8014	1.93E-20
Ischemic CM	yellow	U2SURP	0.2703	1.18E-02	0.5509	3.90E-08
Ischemic CM	yellow	ZDHHC17	0.5542	3.11E-08	0.8610	2.17E-26
Viral CM	cyan	EIF3M	0.2842	7.99E-03	0.8171	8.53E-22
Viral CM	cyan	GTF2H	0.2701	1.19E-02	0.8356	1.44E-23
Viral CM	cyan	LSM5	0.2492	2.07E-02	0.8274	9.43E-23
Viral CM	cyan	TTC37	0.3557	7.76E-04	0.9213	3.21E-36
Viral CM	cyan	UBA3	0.2859	7.62E-03	0.8105	3.30E-21
Viral CM	cyan	UFL1	0.2817	8.59E-03	0.8862	8.42E-30

Table 2: The Hub genes of different Module−Traits.

Traits	GeneSymbol	Disease			Health			FC	p Value
		Mean	SEM	No.	Mean	SEM	No.		
Idiopathic-CM	ACTR2	11.2084	0.0801	60	10.1487	0.2159	35	1.1044	4.40E-07
Idiopathic-CM	CDC40	9.6907	0.2231	24	8.4066	0.2743	8	1.1528	4.63E-03
Idiopathic-CM	DBTdf	10.3068	0.1008	60	9.1643	0.2089	33	1.1247	2.75E-07
Idiopathic-CM	HNRNPC	11.9432	0.1139	50	10.7941	0.2503	29	1.1065	9.19E-06
Idiopathic-CM	MCL1	11.6712	0.2060	36	10.2201	0.3233	36	1.1420	3.21E-04
Idiopathic-CM	MOB1A	10.4879	0.0792	48	9.8165	0.1240	22	1.0684	1.50E-05
Idiopathic-CM	NCOA2	9.5979	0.1624	48	8.9721	0.1077	22	1.0698	1.53E-02
Idiopathic-CM	NRAS	9.8688	0.0430	24	9.1749	0.0725	8	1.0756	4.66E-09
Idiopathic-CM	NUDT21	10.2331	0.1222	36	9.2766	0.1482	15	1.1031	4.18E-05
Idiopathic-CM	PICALM	10.9074	0.1273	50	10.0119	0.1808	29	1.0894	8.91E-05
Idiopathic-CM	PIK3C2A	10.2398	0.1485	63	9.6907	0.1643	30	1.0567	2.65E-02
Idiopathic-CM	PRKACB	11.1593	0.2033	26	9.3188	0.3552	14	1.1975	2.13E-05
Idiopathic-CM	RHOA	13.2077	0.1244	24	9.9748	0.8547	14	1.3241	2.38E-05
Idiopathic-CM	SRSF10	10.0730	0.0811	49	9.2511	0.1343	29	1.0888	3.61E-07
Idiopathic-CM	SRSF11	10.6855	0.2445	26	8.6993	0.3674	20	1.2283	2.89E-05
Idiopathic-CM	UBE2D3	12.0910	0.1415	48	11.1499	0.2675	28	1.0844	1.03E-03
Idiopathic-CM	UEVLD	9.5126	0.0569	48	9.1576	0.0826	22	1.0388	7.84E-04
Idiopathic-CM	USO1	11.7514	0.0463	24	11.4743	0.0691	8	1.0242	4.34E-03
Familial-CM	PIK3C2A	10.8135	0.1760	20	10.2671	0.1454	49	1.0532	0.0352
Post.partum CM	MOB1A	10.4854	0.1408	16	10.0362	0.1267	28	1.0448	0.0291
Ischemic CM	BUB3	10.5611	0.0544	80	10.1011	0.1089	26	1.0455	1.02E-04
Ischemic CM	CCNC	11.0275	0.0298	20	10.7252	0.0456	7	1.0282	1.79E-05
Ischemic CM	LAMP2	10.3801	0.0553	80	9.9620	0.0684	28	1.0420	8.20E-05
Ischemic CM	PICALM	11.0499	0.1150	80	10.5756	0.1691	31	1.0448	2.78E-02
Ischemic CM	TMX3	10.7645	0.0284	20	10.5116	0.0312	7	1.0241	5.05E-05
Ischemic CM	DNAJB14	9.9064	0.0630	54	9.6485	0.0641	21	1.0267	2.05E-02
Ischemic CM	ZDHHC17	11.1324	0.0400	20	10.7254	0.0932	7	1.0380	8.13E-05
Viral CM	LSM5	8.8316	0.1141	14	8.2936	0.0702	22	1.0649	1.54E-04
Viral CM	UBA3	12.5941	0.0443	7	12.0367	0.1232	11	1.0463	3.07E-03
Viral CM	EIF3M	12.4758	0.0754	14	12.2416	0.0701	22	1.0191	3.49E-02
Viral CM	GTF2H3	9.4284	0.1021	14	8.8712	0.1514	22	1.0628	1.10E-02
Viral CM	UFL1	10.7951	0.2482	14	10.1670	0.1744	22	1.0618	4.04E-02

Table 3: Significant Expressed Hub genes between different traits and health.

Cardiomyopathy	Gene Symbol	Gene Title	Function	Associated Heart Diseases
Ischemic (10)	AKAP9	A-kinase anchoring protein 9	binding to the regulatory subunit of protein kinase A (PKA) and confining the holoenzyme to discrete locations within the cell	Dilated cardiomyopathy / Cardiac Arrhythmias
	CFLAR	CASP8 and FADD-like apoptosis regulator	Regulator of apoptosis and is structurally similar to caspase	Congestive heart failure / Myocardial Reperfusion Injury
	CTNNB1	Catenin beta 1	regulate cell growth and adhesion between cells	Dilated cardiomyopathy / Diabetic Cardiomyopathies
	DIP2A	disco interacting protein 2 omolog A	involved in axon patterning in the central nervous system	long QT syndrome
	MDM4	MDM4, p53 regulator	Inhibits p53 by binding its transcriptional activation domain	Dilated cardiomyopathy
	MEF2A	myocyte enhancer factor 2A	DNA-binding transcription factor that activates many musclespecific, growth factor-induced, and stress-induced genes	Dilated cardiomyopathy myocardial infarction
	PMP22	peripheral myelin protein 22	major component of myelin in the peripheral nervous system	Early-Onset Myopathy with Fatal Cardiomyopathy
	RAF1	Raf-1 proto-oncogene, serine/ threonine kinase	functions downstream of the Ras family of membrane associated GTPases	Cardiomegaly
	SMAD4	SMAD family member 4	acts as a tumor suppressor and inhibits epithelial cell proliferation	Myocardial Reperfusion Injury / myocardial infarction
	ZNF862	zinc finger protein 862	Regulation of transcription, DNA-templated	Long QT syndrome
	ADAM10	ADAM metallopeptidase domain 10	cleaves many proteins including TNF-alpha and E-cadherin	Cardiomyopathy
	ABCC9	ATP binding cassette subfamily C member 9	ATP-sensitive potassium channels in cardiac, skeletal, and vascular and non-vascular smooth muscle	Dilated cardiomyopathy
	ARL13B	ADP-ribosylation factor-like 13B	Maintenance of cilia	Long QT syndrome
	ATL3	atlastin GTPase 3	proper formation of the network of interconnected tubules of the endoplasmic reticulum	Long QT syndrome
	CACNA2D1	calcium voltage-gated channel auxiliary subunit alpha2delta 1	Calcium channels mediate the influx of calcium ions into the cell upon membrane polarization	Blugada syndrome and short QT syndrome
	CREB1	cAMP responsive element binding protein 1	induces transcription of genes in response to hormonal stimulation of the cAMP pathway	Dilated cardiomyopathy
	CTNNA3	catenin (cadherin-associated protein), alpha 3	plays a role in cell-cell adhesion in muscle cells	Dilated cardiomyopathy / Long QT syndrome
	DNM1L	dynamin 1-like	regulated apoptosis and programmed necrosis.	Dilated cardiomyopathy
	FGF2	fibroblast growth factor 2 (basic)	possess broad mitogenic and angiogenic activities	Cardiomegaly
	ELMOD2	ELMO domain containing 2	play a role in antiviral responses	long QT syndrome
	FHL1	four and a half LIM domains 1	Involved in many cellular processes.	Hypertrophic cardiomyopathy
	FHL2	four and a half LIM domains 2	have a role in the assembly of extracellular membranes	Cardiomegaly / Dilated Cardiomyopathy
	FPGT-TNNI3K	FPGT-TNNI3K readthrough	neighboring fucose-1-phosphate guanylyltransferase (FPGT) and TNNI3 interacting kinase	Cardiac Conduction Disease with or without Dilated Cardiomyopathy
	GNB1L	guanine nucleotide binding protein , beta polypeptide 1-like	cell cycle progression, signal transduction, apoptosis, and gene regulation	Dilated cardiomyopathy
	HSPD1	heat shock 60kDa protein 1 (chaperonin)	signaling molecule in the innate immune system, essential for the folding and assembly of newly imported proteins in the mitochondria.	Cardiomyopathy / Myocardial Ischemia
	ITGB1	integrin, beta 1	involved in cell adhesion and recognition in a variety of processes including embryogenesis, hemostasis, tissue repair, immune response	Cardiomyopathy
	ITGB1BP2	integrin beta 1 binding protein (melusin) 2	Act as a chaperone protein	Dilated cardiomyopathy
	LMNA	lamin A/C	matrix of proteins located next to the inner nuclear membrane	Dilated cardiomyopathy
	LRBA	LPS responsive beige-like anchor protein	associates with protein kinase A and may be involved in leading intracellular vesicles to activated receptor complexes	long QT syndrome
	MAPK1	mitogen-activated protein kinase 1	extracellular signal-regulated kinases (ERKs), act as an integration point for multiple biochemical signals	Cardiomyopathy
	MAPK11	mitogen-activated protein kinase 11	activated by proinflammatory cytokines and environme stresses through phosphorylation by MKKs	Cardiomyopathy
	MAPK14	mitogen-activated protein kinase 14	including cell proliferation, differentiation, transcription regulation	Cardiomyopathy
	NEBL	nebulette	binds actin and interacts with thin filaments and Z-li associated proteins in striated muscle.	Dilated cardiomyopath /Cardiomyopathy
	PDLIM5	PDZ and LIM domain 5	functions as a scaffold protein that tethers protein kinases to Z-disk in striated muscles	Dilated cardiomyopathy/ Myocardi Ischemia
	PECAM1	platelet and endothelial cell adhesion molecule	involved in leukocyte migration, angiogenesis, and integ activation	Myocardial infarction / Myocardia Reperfusion Injury
	PRKAA2	protein kinase, AMP-activated, alpha 2 cataly subunit	catalytic subunit of the AMP-activated protein kinase	Cardiomegaly / Cardiomyopathy
	PTPN11	protein tyrosine phosphatase, non-receptor ty 11	regulate a variety of cellular processes including cell growth differentiation, mitotic cycle, and oncogenic transformation	Cardiovascular Abnormalities
	RAB1A	RAB1A, member RAS oncogene family	small GTPase controls vesicle traffic from the endoplasm reticulum to the Golgi apparatus	Dilated cardiomyopathy
	RAC1	ras-related C3 botulinum toxin substrate 1	regulate a diverse array of cellular events, including the cont of cell growth, cytoskeletal reorganization, and the activation of protein kinases	Dilated cardiomyopathy
	RHEB	Ras homolog enriched in brain	Regulation of growth and cell cycle progression due to its ro in the insulin/TOR/S6K signaling pathway	Long QT syndrome
	RHOA	ras homolog family member A	reorganization of the actin cytoskeleton and regulate cell shape, attachment, and motility	Cardiomegaly
	RHOJ	ras homolog family member J	regulate endothelial cells focal adhesion and angiogenesis	Cardiomyopathy
	SGCB	sarcoglycan, beta (43kDa dystrophin-associated glycoprotein)	help stabilize the muscle fiber membranes and link the muscle cytoskeleton to the extracellular matrix	Cardiomyopathy / Cardiac Arrhythmias
	TECRL	trans-2,3-enoyl-CoA reductase-like	steroid 5-alpha reductase family.	ventricular tachycardia
	WDR26	WD repeat domain 26	Facilitate formation of heterotrimeric or multiprotein complexes	Long QT syndrome
	XIAP	X-linked inhibitor of apoptosis	anti - apoptotic function	Myocardial Reperfusion Injury
	YME1L1	YME1-like 1 ATPase	plays a role in mitochondrial protein metabolism	Dilated cardiomyopathy / Long Q syndrome

	LMNA	lamin A/C	matrix of proteins located next to the inner nuclear membrane	Dilated cardiomyopathy
Familiar Cardiomyopathy & Cardiomyopathy & Idiopathic Cardiomyopathy (40)	MAPK1	mitogen-activated protein kinase 1	extracellular signal-regulated kinases (ERKs), act as a integration point for multiple biochemical signals	Cardiomyopathy
	MAPK11	mitogen-activated protein kinase 11	activated by pro inflammatory cytokines and environmental stresses through phosphorylation by MKKs	Cardiomyopathy
	MAPK14	mitogen-activated protein kinase 14	involved in the integration of a wide variety of cellular processes, including cell proliferation, differentiation, transcriptional regulation	Cardiomyopathy
	RAC1	ras-related C3 botulinum toxin substrate 1	regulate a diverse array of cellular events, including the control of cell growth, cytoskeletal reorganization, and the activation of protein kinases	Dilated cardiomyopathy
	PECAM1	platelet and endothelial cell adhesion molecule	involved in leukocyte migration, angiogenesis, and integrin activation	Myocardial infarction / Myocardial Reperfusion Injury
	XIAP	X-linked inhibitor of apoptosis	anti - apoptotic function	Myocardial Reperfusion Injury
	CREB1	cAMP responsive element binding protein 1	induces transcription of genes in response to hormonal stimulation of the cAMP pathway	Dilated cardiomyopathy
	AKAP8	A-kinase anchoring protein 8	and recruit PKA and other signaling molecules to specific subcellular location	Long QT syndrome
	AVP	arginine vasopressin	supraoptic nucleus and paraventricular nucleus of the hypothalamus	Cardiomyopathy / Congestive heart failure
	AVPR2	arginine vasopressin receptor 2	GPCR concentrate the urine and maintain water homeostasis in the organism.	Cardiomyopathy / Congestive heart failure
	BAD	BCL2 associated agonist of cell death	positively regulates cell apoptosis by forming heterodimers w BCL-xL and BCL-2	Myocardial Reperfusion Injury
	CACNB2	calcium voltage-gated channel auxiliary subu beta 2	identified as an antigen target in Lambert-Eaton my as the syndrome,	Cardiac arrest /hypertrophic cardiomyopathy
	CASP9	caspase 9	Sequential activation of caspases plays a central role in the execution-phase of cell apoptosis	Dilated cardiomyopathy / Diabetic Cardiomyopathies
	CD40	CD40 molecule	a member of the TNF-receptor superfamily mediating a bro variety of immune and inflammatory responses	Dilated cardiomyopathy
	CD40LG	CD40 ligand	regulates B cell function by engaging CD40 on the B cell surface	Myocardial Ischemia / Myocardial infarction
	CD55	CD55 molecule	a glycoprotein involved in the regulation of the complement cascade	Viral Myocarditis
	CFLAR	CASP8 and FADD like apoptosis regulator	a regulator of apoptosis	Congestive heart failure / Myocardial Reperfusion Injury
	DES	desmin	Mutations in this gene are associated with desmin-related myopathy, a familial cardiac and skeletal myopathy	Cardiomyopathy / myofibrillar myopathy
	DICER1	dicer 1, ribonuclease III	Functions as a ribonuclease	Dilated cardiomyopathy
	DTNA	dystrobrevin alpha	a component of the dystrophin-associated protein complex	Cardiomyopathy
	EDNRB	endothelin receptor type B	GPCR that activates a phosphatidylinositol-calcium second messenger system.	Myocardial Ischemia
	FAM135A	family with sequence similarity 135 member A	Associated with infantile histiocytoid cardiomyopathy	infantile histiocytoid cardiomyopathy
	FKTN	fukutin	involved in the glycosylation of alpha-dystroglycan in skeletal muscle	cardiomyopathy / dilate cardiomyopathy
	GATA4	GATA binding protein 4	a member of the GATA family of zinc-finger transcription factors	Congenital heart disease
	GNAQ	G protein subunit alpha q	a guanine nucleotide-binding protein	congestive heart failure
	GSK3A	glycogen synthase kinase 3 alpha	multifunctional Ser/Thr protein kinase that regulates WNT and PI3K signaling pathways	dilated cardiomyopathy
	HAND2-AS1	HAND2 antisense RNA 1	Associated with dilated cardiomyopathy 1A	dilated cardiomyopathy 1A
	HMGB1	high mobility group box 1	non-histone, nuclear DNA-binding protein regulate transcription and involves in organization of DNA	Myocardial Ischemia / Myocardia Reperfusion Injury
	HNRNPM	heterogeneous nuclear ribonucleoprotein M	RNA binding proteins and they complex with heterogeneous nuclear RNA	Long QT syndrome
	HSPB1	heat shock protein family B (small) member 1	plays an important role in the differentiation of a wide variety of cell types.	congestive heart failure
	ITPR1	inositol 1,4,5-trisphosphate receptor type 1	mediates calcium release from the endoplasmic reticulum	dilated cardiomyopathy
	JAG1	jagged 1	the ligand for the receptor notch 1	dilated cardiomyopathy 1G
	KAT8	lysine acetyltransferase 8	a member of the MYST histone acetylase protein family	hypertrophic cardiomyopathy / congestive heart failure
	KCNE1	potassium voltage-gated channel subfamily regulatory subunit 1	potassium channel KCNE family	long QT syndrome / hypertroph cardiomyopathy
	LAMA3	laminin subunit alpha 3	alpha subunit is responsive to several epithelial-mesenchymal regulators	dilated cardiomyopathy 1JJ
	LAMP2	lysosomal associated membrane protein 2	glycoprotein provides selectins with carbohydrate ligands	hypertrophic cardiomyopathy
	LRP4	LDL receptor related protein	a regulator of Wnt signaling	hypertrophic cardiomyopathy
	MAP1LC3A	microtubule associated protein 1 light cha 3 alpha	light chain subunits and can associate with either MAP1A MAP1B	hypertrophic cardiomyopathy
	MDM4	MDM4, p53 regulator	nuclear protein that contains a p53 binding domain at the terminus and a RING finger domain at the C-terminus	dilated cardiomyopathy
	MYO6	myosin VI	plays a role in intracellular vesicle and organelle transport	Sensorineural Deafness with Hypertrophic Cardiomyopathy
	NFE2L2	nuclear factor, erythroid 2 like 2	transcription factor regulates genes which contain antioxidants response elements (ARE)	Congestive heart failure / Diabetic Cardiomyopathies
	NOS2	nitric oxide synthase 2	a nitric oxide synthase is expressed in liver and is inducible by a combination of lipopolysaccharide and certain cytokines.	congestive heart failure
	NOS3	nitric oxide synthase 3	nitric oxide synthase 3 transcript	cardiomyopathy / congestive heart failure
	NR2C2	nuclear receptor subfamily 2 group C member	plays a role in protecting cells from oxidative stress and damage induced by ionizing radiation	Isolated Noncompaction of the Ventricular Myocardium
	NRIP1	nuclear receptor interacting protein 1	modulates transcriptional activity of the estrogen receptor	Congestive heart failure / long QT syndrome
	PCNA	proliferating cell nuclear antigen	cofactor of DNA polymerase delta	Chagas Cardiomyopathy
	PKD1L2	polycystin 1 like	function as a component of cation channel po	Long QT syndrome

	PLCB4	phospholipase C beta 4	plays an important role in the intracellular transduction of m extracellular signals in the retina	Long QT syndrome
	POLRMT	RNA polymerase mitochondrial	a mitochondrial DNA-directed RNA polymerase.	Long QT syndrome
	PRKCA	protein kinase C alpha	serine- and threonine-specific protein kinases	dilated cardiomyopathy
	PSEN1	presenilin 1	regulate APP processing through their effects on gamma-secretase	cardiomyopathy
	RALGAPA1	Ral GTPase activating protein catalytic alpha subunits	major subunit of the RAL-GTPase activating protein	Long QT syndrome
	RPS6KB1	ribosomal protein S6 kinase B	responds to mTOR (mammalian target of rapamycin) signalingto promote protein synthesis, cell growth, and cell proliferation	dilated cardiomyopathy / Left Ventricular Hypertrophy
	SCN4B	sodium voltage-gated channel beta subunit 4	interact with voltage-gated alpha subunits to change sodium channel kinetics.	Dilated cardiomyopathy / long QT syndrome
	SLC2A5	solute carrier family 2 member 5	a fructose transporter responsible for fructose uptake by the small intestine	Long QT syndrome
	SMAD1	SMAD family member 1	mediates the signals of the bone morphogenetic proteins (BMPs),	Myocardial Reperfusion Injury
	SNTA1	syntrophin alpha 1	a member of the syntrophin gene family and encodes the most common syntrophin isoform found in cardiac tissues	arrhythmogenic right ventricular cardiomyopathy / Long QT syndrome
	SOD2	superoxide dismutase 2	a member of the iron/manganese superoxide dismutase family	Cardiomyopathy
	TCAP	titin-cap	a giant elastic protein with kinase activity that extends half the length of a sarcomere	Dilated cardiomyopathy
	TMEM70	transmembrane protein 70	a mitochondrial membrane protein that play a role in biogenesis of mitochondrial ATP synthase	cardiomyopathy
	TMUB1	transmembrane and ubiquitin like doma containing	involved in ubiquitin-dependent ERAD pathway	Long QT syndrome
	TXN	thioredoxin	active in the reversible S-nitrosylation of cysteines in cert proteins	Dilated cardiomyopathy
	TXN2	thioredoxin 2	play important roles in the regulation of the mitochondria membrane potential and in protection against oxidant-induced apoptosis	Myocardial Reperfusion Injury
	TXNRD2	thioredoxin reductase 2	a member of the thioredoxin (Trx) system and and play a key role in redox homoeostasis	Dilated cardiomyopathy / hypertrophic cardiomyopathy
	UBR4	ubiquitin protein ligase E3 component recognin 4	a cytoskeletal component in the cytoplasm and part o chromatin scaffold in the nucleus	Long QT syndrome
	UBR5	ubiquitin protein ligase E3 component nrecognin 5	a progestin-induced protein plays a role in regulation of cell proliferation or differentiation	Long QT syndrome
	UPF3B	regulator of nonsense mediated mRNA decay	This protein binds to the mRNA and remains bound after nuc export, acting as a nucleocytoplasmic shuttling protein.	Danon disease
	ZBTB33	zinc finger and BTB domain containing 33	contribute to the repression of target genes of the W signaling pathway	Danon disease

Table 4: Identified signature genes of different cardiomyopathies through Cardiovascular Diseases dataport blast.

upregulated in Idiopathic cardiomyopathy group, and PIK3C2A was up-regulated in Familiar cardiomyopathy group (Table 3). It hints that these hub genes play different regulatory pattern in the progress of these subtype's cardiomyopathies.

The fourth axis of hub genes (HNRNPC, UEVLD) were shared by Familiar Cardiomyopathy and Idiopathic Cardiomyopathy groups, and significantly overexpressed in Idiopathic cardiomyopathy group (Table 3). HNRNPC (heterogeneous nuclear ribonucleoprotein C) is RNA binding protein that belong to ubiquitously expressed heterogeneous nuclear ribonucleoproteins subfamily, and mediates pre-mRNAs transport and metabolism between cytoplasm and nucleus and overexpression caused cells multi-nucleation [32,33]. UEVLD (EV and lactate/malate dehydrogenase domain-containing protein) involves the protein degradation and dysregulated linked with metabolic disease [34]. In this study, the expression level of HNRNPC and UEVLD were significantly up-regulated in Idiopathic cardiomyopathy group (Table 3). Furthermore, through different significant expression analysis, the significant changed hub genes were summarized (Table 3, $p<0.05$). Combined these results together, it hints that these significantly expressed Hub genes play dominant role and work as common key regulatory nodes in progress of cardiomyopathies.

Disease signature genes identification and expression analysis

The filtered disease signature genes were summarized with the functional annotation of genetic dysregulation correlated to heart diseases phenotypes, including ten signature genes in the ischemic group and viral cardiomyopathy group, forty signature genes among the group of familiar cardiomyopathy, Post-partum cardiomyopathy and Idiopathic cardiomyopathy, and 69 signature genes in the ischemic cardiomyopathy group (Table 4). Through Venn Diagram analysis, the common signature genes were determined among different groups (Figure 5A). Four signature genes (MDM4, CFLAR, RPS6KB1, PKD1L2) were shared by Ischemic and Ischemic Cardiomyopathy group (Table 4, Figure 5A). Ischemic cardiomyopathy group did share eight disease signature genes (MAPK1, MAPK11, MAPK14, LMNA, RAC1, PECAM1, XIAP, CREB1) with Post. Partum/Familiar/Idiopathic Cardiomyopathy groups, which dysregulated in cardiomyopathies [24,35], and genes expression level of MAPK1, MAPK11, LMNA, RAC1 were significantly up-regulated in these cardiomyopathies groups (Table 5, Figure 5A). Two signature genes (TFAM, RHEB) were shared between Viral Cardiomyopathy and Post. Partum / Familiar /Idiopathic Cardiomyopathy groups, which involved in development of cardiac hypertrophy and Mitochondrial Cardiomyopathy [36].

Furthermore, through different significant expression analysis, the significantly changed signature genes were summarized (Figure 5D-F, Table 5). In ischemic group, MDM4 gene was significantly upregulated (FC=1.0495, p=0.0037) (Table 5, Figure 5B), which genetic deletion associated with cardiomyopathy [37]. In viral cardiomyopathy group, COA5 was overexpressed (FC=1.087485, $p<0.0001$) (Table 5,

Traits Group	GeneSymbol	Disease			Health			FC (D/H)	p Value
		Mean	SEM	No.	Mean	SEM	No.		
Ischemic	MDM4	10.0726	0.0889	44	9.5977	0.1409	28	1.0495	0.0037
Idiopathic-CM	ABCC9	9.4143	0.0773	48	9.0757	0.1231	28	1.0373	0.0165
Idiopathic-CM	ADAM10	10.1708	0.2197	36	9.0434	0.0984	14	1.1247	0.0029
Idiopathic-CM	ELMOD2	10.1111	0.0369	24	9.6203	0.1381	14	1.0510	0.0001
Idiopathic-CM	FGF2	8.6292	0.0706	21	8.2251	0.1243	14	1.0491	0.0046
Idiopathic-CM	GUF1	9.3290	0.0521	24	8.9318	0.0928	14	1.0445	0.0003
Idiopathic-CM	RAB1A	11.5943	0.2298	36	10.6694	0.1379	14	1.0867	0.0188
Idiopathic-CM	TFAM	9.0046	0.1526	36	8.4651	0.1932	25	1.0637	0.0308
Idiopathic-CM	FHL1	14.2426	0.0392	60	14.4595	0.0448	35	0.9850	0.0164
Idiopathic-CM	FHL2	14.1101	0.0609	12	14.4395	0.1253	7	0.9772	0.0007
Idiopathic-CM	GNB1L	5.9289	0.1124	32	6.5029	0.1702	15	0.9117	0.0065
Idiopathic-CM	LMNA	9.7879	0.1154	60	10.3262	0.1695	35	0.9479	0.0081
Idiopathic-CM	PDLIM5	12.9541	0.1015	72	13.3927	0.1072	35	0.9673	0.0086
Idiopathic-CM	RHEB	11.3440	0.2426	48	12.2831	0.0677	14	0.9235	0.0426
Idiopathic-CM	SIRT4	7.2843	0.0994	24	7.6728	0.0864	14	0.9494	0.0117
Idiopathic-CM	YME1L1	11.2890	0.2613	49	12.1832	0.0478	21	0.9266	0.0295
Familial-CM	ABCC9	9.4432	0.1260	20	9.0757	0.1231	28	1.0405	0.0475
Familial-CM	ELMOD2	10.1192	0.0658	10	9.6203	0.1381	14	1.0519	0.0087
Familial-CM	FGF2	8.6766	0.0682	10	8.2251	0.1243	14	1.0549	0.0094
Familial-CM	FHL1	14.1261	0.0577	25	14.4595	0.0448	35	0.9769	< 0.0001
Familial-CM	FHL2	14.0531	0.0580	5	14.4395	0.1253	7	0.9732	0.0344
Familial-CM	GUF1	9.4334	0.0836	10	8.9318	0.0928	14	1.0562	0.0009
Familial-CM	NEBL	10.6981	0.2109	20	12.9782	0.1751	35	0.8243	< 0.0001
Familial-CM	SIRT4	7.2472	0.1425	10	7.6728	0.0864	14	0.9445	0.0131
Familial-CM	YME1L1	12.3408	0.0365	15	12.1832	0.0478	21	1.0129	0.02
Post.partum CM	ADAM10	10.1054	0.3804	12	9.0434	0.0984	14	1.1174	0.0079
Post.partum CM	ATL3	8.8791	0.4428	11	7.3675	0.2004	11	1.2052	0.0055
Post.partum CM	CTNNA3	11.1428	0.2581	11	11.8049	0.1564	15	0.9439	0.0295
Post.partum CM	ELMOD2	10.2501	0.0962	8	9.6203	0.1381	14	1.0655	0.0047
Post.partum CM	FGF2	8.7142	0.1084	8	8.2251	0.1243	14	1.0595	0.0154
Post.partum CM	FHL1	14.2645	0.0738	20	14.4595	0.0448	35	0.9865	0.02
Post.partum CM	GNB1L	5.7871	0.1726	11	6.4053	0.1497	14	0.9035	0.0124
Post.partum CM	GUF1	9.3977	0.1125	8	8.9318	0.0928	14	1.0522	0.0054
Post.partum CM	PTPN11	11.1674	0.3287	17	11.8783	0.1393	21	0.9401	0.0397
Post.partum CM	SIRT4	6.9157	0.1035	8	7.6728	0.0864	14	0.9013	< 0.0001
Post.partum CM	YME1L1	12.3456	0.0453	12	12.1832	0.0478	21	1.0133	0.0315
Ischemic CM	EDNRB	10.6176	0.0534	60	10.2983	0.1701	14	1.0310	0.0226
Ischemic CM	FKTN	9.4314	0.0571	20	9.2066	0.0650	7	1.0244	0.0415
Ischemic CM	LAMP2	10.3801	0.0553	80	9.8800	0.0995	14	1.0506	0.0005
Ischemic CM	MYO6	5.8864	0.1716	40	4.8988	0.2018	14	1.2016	0.0028
Ischemic CM	PSEN1	9.3882	0.0759	80	8.9537	0.1473	14	1.0485	0.0259
Ischemic CM	RAC1	11.8742	0.0445	40	11.6806	0.0448	14	1.0166	0.0192
Ischemic CM	RALGAPA1	9.8937	0.0576	40	9.5576	0.0895	14	1.0352	0.0038

Ischemic CM	SCN4B	9.6672	0.0679	20	8.9785	0.2344	7	1.0767	0.0007
Ischemic CM	SMAD1	10.1734	0.0527	40	9.3196	0.1080	14	1.0916	< 0.0001
Ischemic CM	TXN	11.2957	0.0583	20	10.9175	0.0991	7	1.0346	0.0029
Ischemic CM	UBR5	10.0164	0.1104	60	9.5019	0.1791	14	1.0541	0.0391
Ischemic CM	ZBTB33	10.2999	0.0476	20	9.9715	0.0490	7	1.0329	0.0008
Ischemic CM	AKAP8	9.1814	0.0535	40	9.6180	0.0903	14	0.9546	0.0001
Ischemic CM	AVP	4.4059	0.2037	20	5.5228	0.3518	7	0.7978	0.0102
Ischemic CM	BAD	8.4669	0.0464	40	8.9077	0.0766	14	0.9505	< 0.0001
Ischemic CM	CACNB2	8.5018	0.1058	100	9.5708	0.2836	14	0.8883	0.0006
Ischemic CM	CASP9	9.0736	0.0486	40	9.4403	0.0609	14	0.9612	0.0002
Ischemic CM	DES	13.9494	0.0492	40	14.6541	0.0926	14	0.9519	< 0.0001
Ischemic CM	GATA4	8.9982	0.2629	60	10.2513	0.2197	14	0.8778	0.0275
Ischemic CM	HSPB1	13.7888	0.0498	20	14.2947	0.1083	7	0.9646	< 0.0001
Ischemic CM	LAMA3	7.2065	0.1142	60	7.7601	0.1019	14	0.9287	0.0253
Ischemic CM	LMNA	10.0702	0.0810	80	10.8284	0.1297	14	0.9300	0.0003
Ischemic CM	LRP4	8.2463	0.0583	20	8.6902	0.1014	7	0.9489	0.0007
Ischemic CM	MAP1LC3A	9.8156	0.0813	60	10.5112	0.0983	14	0.9338	0.0002
Ischemic CM	MAPK1	10.5585	0.1339	80	11.5055	0.0424	14	0.9177	0.0041
Ischemic CM	MAPK14	9.0078	0.1004	60	9.6265	0.1016	14	0.9357	0.0052
Ischemic CM	NOS3	7.6231	0.1380	20	8.2386	0.1546	7	0.9253	0.0221
Ischemic CM	POLRMT	4.6718	0.2242	40	6.3274	0.1253	14	0.7383	< 0.0001
Ischemic CM	SNTA1	11.5196	0.0639	20	11.9146	0.1288	7	0.9668	0.006
Ischemic CM	SOD2	10.0271	0.0801	20	10.7952	0.0394	7	0.9288	< 0.0001
Ischemic CM	TCAP	12.8665	0.0692	20	13.6153	0.1449	7	0.9450	< 0.0001
Virtal CM	COA5	10.6324	0.148884	14	9.7770	0.101869	14	1.087485	< 0.0001

Table 5: Significant expressed Signature genes between different traits and health.

Figure 5C), which was upregulated in Ischemia/Reperfusion Injury caused cardiomyopathy [38]. In idiopathic cardiomyopathy group, seven genes (ADAM10, RAB1A, TFAM, FGF2, ELMOD2, GUF1, ABCC9) were significantly up-regulated, while 8 genes (FHL1, CTNNA3, PDLIM5, LMNA, SIRT4, YME1L1, RHEB, GNB1L) were down-regulated (Figure 5D, Table 5). In idiopathic cardiomyopathy group, four down-regulated genes (NEBL, FHL1, FHL2, SIRT4) and 5 up-regulated genes (GUF1, ELMOD2, ABCC9, FGF2, YME1L1) were significantly changed (Figure 5E, Table 5). In post-partum cardiomyopathy group, 11 genes (ATL3, ADAM10, ELMOD2, FGF2, GUF1, YMEIL1, up-regulated; FHL1, CTNNA3, PTPN11, GNB1L, SIRT4, down-regulated) were significantly changed (Figure 5F, Table 5). In Ischemic cardiomyopathy group, 31 genes are significantly changed expression, including 12 genes (RAC1, FKTN, EDNRB, ZBTB33, TXN, RALGAPA1, PSEN1, LAMP2, UBR5, SCN4B, SMAD1, MYO6) down-regulated and 19 genes up-regulated expression (POLRMT, AVP, GATA4, CACNB2, MAPK1, NOS3, LAMA3, SOD2, LMNA, MAP1LC3A, MAPK14, TCAP, LRP4, BAD, DES, AKAP8, CASP9, HSPB1, SNTA1) (Figure 5G, Table 5). These results suggest that these common disease signature genes work as novel biomarker and be potential key regulators of the cardiomyopathy progress.

Discussion

In this study, to discover novel signature genes or biomarkers to accelerate the precise clinical diagnostics and interference for different subtype of cardiomyopathies, the WGCNA pipeline was applied to analyze the gene expression profiling of 86 clinical left ventricle biopsy samples, which represents 8 subtype's cardiomyopathies. The WGCNA pipeline was widely used for performing various functions in weighted correlation network analysis, including constructing network, detecting module, calculating topological properties, simulating data, visualization, and interfacing with external software [9]. The whole transcriptome profile contained 20,283 target genes for promise diagnostic assessment and mainly covered variously biological and cellular processes. It is representative of real pathological satiation and valuable to discover the signature gene of cardiomyopathies. First of all, it was reasonable to build the co-expression networks with different clinical cardiomyopathies traits using the Pearson correlation (Figure 2A). To discover the related modules to cardiomyopathies phenotype, the genes significance of the modules was calculated by the linear mixed effects model for testing the association of node to the pathological

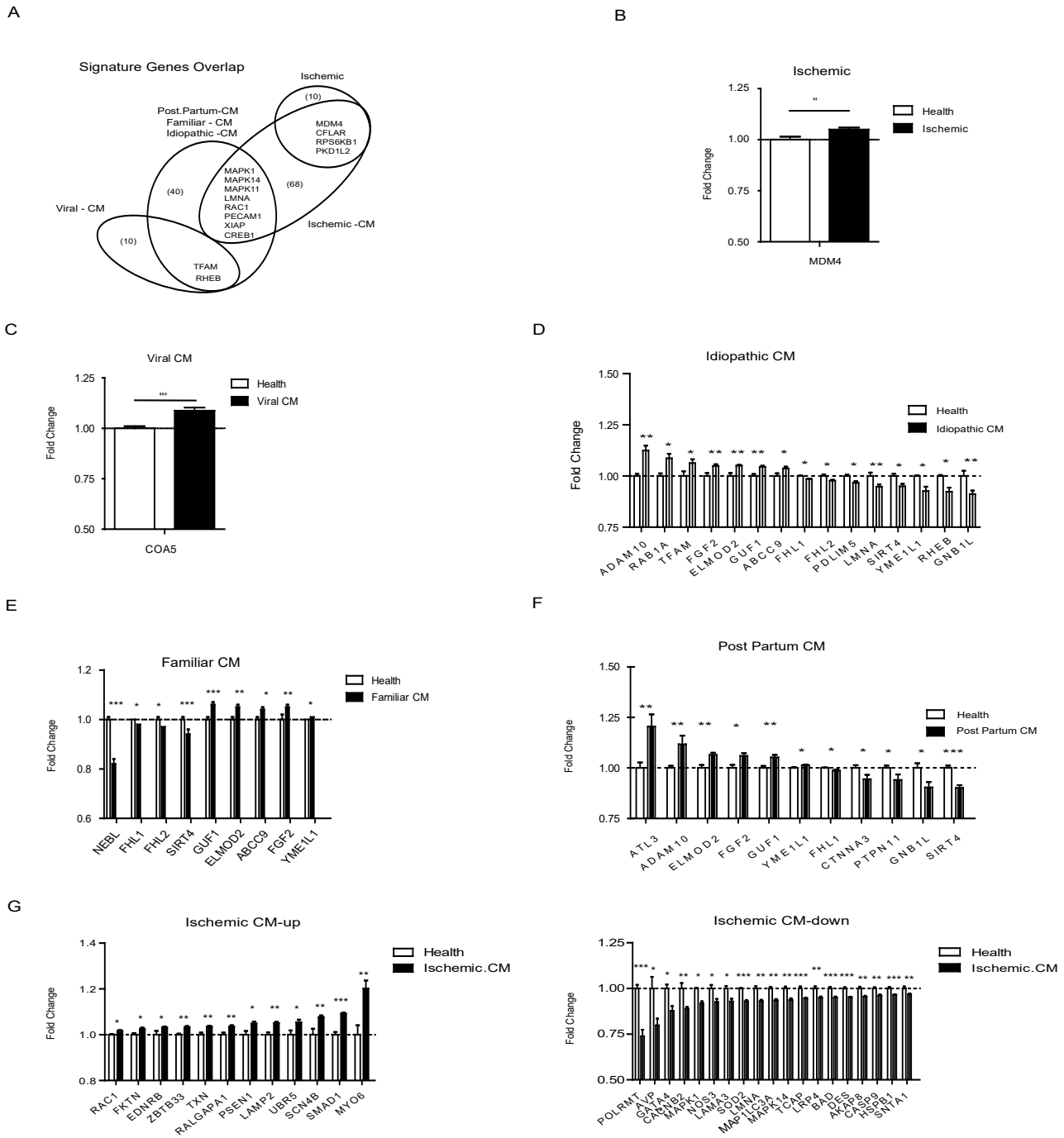

Figure 5: Validation of Disease Signature genes expression. (A). The Venn diagram showed the overlapped disease signature genes for different cardiomyopathies. (B)~(G). Expression levels of signature genes between subtype of cardiomyopathies and healthy control. (***, represent p value <0.0001; **, represent 0.001<p value <0.0001; *, represents p value <0.05).

phenotypes. It was identified that the association significance between individual modules of gene expression profile and different cardiomyopathies feature (Figure 2B). Through the Eigengene dendrogram analysis, the most significantly module was pick out for next analysis (Figure 3B, Figure S4A-G). For the next step, the real hub genes among each significant module were screened by module membership (MM) - Gene Significance and Protein-Protein Interaction Network analysis, and comprised as key interconnected nodes within a functionally network and played important roles in biological functions [14]. Without any real hub genes identified in the Idiopathic Dilated group, it was discarded for next analysis. In addition, the Idiopathic Dilated case was treated as unique physiological state without impaired the normal cardiac function and ignored for analysis. Briefly, the next analysis was mainly concentrated on these five subtype's groups, including idiopathic cardiomyopathy (IdCM), familial cardiomyopathy (FCM), post-partum cardiomyopathy (PCM), Ischemic cardiomyopathy (IsCM) and viral cardiomyopathy (VCM). There were four axes of hub genes shared among these cardiomyopathic groups. It was suggesting that these Hub genes work as common key regulator. It was exception that viral cardiomyopathy group did not share hub gene with the others groups. It was possible

unique that the dysregulation expression pattern of viral cardiomyopathy. Furthermore, to deeply dig the correlation of significance genes and different cardiomyopathies, the significance genes were blast through GenCliP2 to mine gene networks and functions connection, biological process, molecular functions, the cellular components and functional pathways. Although the enriched pathways were varied from different subtype's cardiomyopathies, the key pattern was similar (Supplementary Figure S17-22). The biological processes were significantly concentrated on cellular metabolic, protein metabolic, organic substance metabolic and macromolecule modification. It suggests that the metabolic process disorder be associated with progress of cardiomyopathies. The molecular functions were mainly involving in protein binding, heterocyclic compound binding, purine ribonucleotide binding, iron binding and nucleotide binding, etc. The cellular components were including membrane-bounded organelle, intracellular organelle part, etc. The pathways were concentrated on MAPK signaling pathway, protein kinase C protein processing in endoplasmic reticulum, regulation of actin cytoskeleton, etc (Supplementary Figure S23-28). These related genes were summarized with functions and pathways (Supplementary Table 4). Furthermore, most of significance genes were labelled as gene-term association not reported. It represents that the regulatory mechanism of these genes is illusive in progress of cardiomyopathies. Through Literature Gene Networks Mining, the genes were identified in the regulatory network with less published literatures, which were associated with different cardiomyopathies except Post-Partum cardiomyopathy (Supplementary Figure S20-B). It is possible that less researches and few reports concentrated on Post-Partum cardiomyopathy. For subtype of idiopathic cardiomyopathy (IdCM), familial cardiomyopathy (FCM), and Ischemic cardiomyopathy (IsCM), the highlighted genes were mainly concentrated in MAPK signaling pathway, including MAPK1, MAPK14, CREB1, RAC1. The genes YY1, RAPGEF1, SMAD2, JUND, ATF1, and SRA1 linked with viral cardiomyopathy (Supplementary Figure S22-B), which are assemble in SMAD signaling pathway and YY1 was overexpressed in heart failure [39]. These results partially matched the key axes of hub genes linked the functions and pathways. These new discovered genes linked to viral cardiomyopathy may give a new view for clinical diagnostics and treatment. It suggests that the dysfunctions of these significance genes would be associated with metabolism disorder and progress of cardiomyopathies.

Limited by hardly accessible to clinical samples, blast through the cardiovascular disease BioPortal database was employed as validation strategy to explore the disease signature genes that associated with different subtype cardiomyopathies. The Cardiovascular Disease Portal provides easy access to multiple genetic data associated with specific cardiomyopathy types and. The Cardiovascular Disease Portal integrates data for genes, QTLs and strains associated with the disease(s) highlighted and translational research annotation and associated disease information. The filtered genes will be defined as disease signature genes for cardiomyopathies [12,13]. The varied number of disease signature genes were filtered for different groups (Table 4). There were three axes of common signature genes identified among five subtype's cardiomyopathies groups (Figure 5A). The first axis contained 4 disease signature genes (MDM4, CFLAR, RPS6KB1, PKD1L2) shared by ischemic and ischemic cardiomyopathy group. Compared with health control, only MDM4 was significantly overexpressed (FC=1.0495, p=0.0037) in ischemic group, while the four signature genes did not significantly change in ischemic cardiomyopathy group. It matches the previously reports that the upregulated MDM4 plays cardio-protective effect in ischemia-refuse injury [40]. Overexpression of MDM4 reflects the self-correction of physiological system in abnormal physiological condition of ischemic. It will generate new strategy to interfere the ischemic lapse into cardiomyopathies by artificial upregulated MDM4 expression. The secondary axis, consisted of eight signature genes (MAPK1, MAPK11, MAPK14, LMNA, RAC1, PECAM1, XIAP, CREB1), was shared by Ischemic Cardiomyopathy with Post. Partum/Familiar/Idiopathic Cardiomyopathy groups. These signature genes (MAPK1, MAPK11 and LMNA down-regulated; RAC1 up-regulated) were significantly changed in ischemic cardiomyopathy group, which played dominant role in progress of cardiomyopathy, and only LMNA was significantly down-regulated in Idiopathic Cardiomyopathy group (Table 5, Figure 5G). It suggests that the disorder expression pattern of three subtypes cardiomyopathies could be more complex. The signature genes (TFAM, RHEB) were shared by Viral Cardiomyopathy and Post. Partum/Familiar/ Idiopathic Cardiomyopathy groups, and only RHEB was significantly down-regulated in Idiopathic cardiomyopathy group. It suggests that signature genes (TFAM, RHEB) play less contribution in pathological progress of viral cardiomyopathy and Post. Partum/Familiar /Idiopathic Cardiomyopathy. Combined these results together, it strongly suggests that these genes could be used as promising biomarkers or therapy targets for cardiomyopathies. This progress will be helpful to integrate precise clinical application for different subtype cardiomyopathies.

There are some limitations in this study. Firstly, lacking of larger number of clinical samples and more detail of disease state information, it was hardly to track the patient's cases with expression profiles and verify these potential biomarkers with original patient's pathological feature. Secondly, due to the nature of bioinformatics analysis, the discovered specific GO pathways and biomarkers do need further investigation. Although these genes were validated significantly associated with cardiomyopathies feature through cardiovascular disease BioPortal database, and it is necessary to verify these potential biomarkers with more clinic patient's biopsies by immunohistochemistry (IHC) or other genetic detection method, like qPCR or sequencing in coming research work. It is mandatory and rational to investigate the contribution and mechanism of these potential disease signature genes to the progress of cardiomyopathies by animal model and validate with clinical data in the future research. On the other side, this study has several novelties. Firstly, it has applied reverse strategy by using WGCNA approach to discover the genes significantly associated with cardiomyopathies pathological feature in clinical samples. In parallels, compared these signature genes

expression level among health control and disease groups, the results indicated that signature genes have significant expression change in disease groups (Table 5). Secondly, the key functional & GO pathways that genes significance in module involving in the progress of cardiomyopathies were inquired by Genclip enrichment analysis. The results will be a clause for the next step research. Thirdly, exploring through bioportal database, some novel potential signature genes were identified as potential biomarkers of cardiomyopathies in previously independent researches. These results are partially as evidences to support our results and research strategy.

Conclusion

In summary, this study provides new insight to identify the potential novel key regulatory biomarkers or therapy targets for varied cardiomyopathies induced by different etiologies. The disease signature genes associated with cardiomyopathies were identified and listed as the potential therapy targets for clinical application. In the future research, the detail of regulatory mechanism of these disease signature genes will be deeply investigated and develop novel therapy strategy for cardiomyopathies.

Competing Interests

All authors declare that they have no competing interests.

Funding

This work was supported by Chinese Academy of Science "CAS Pioneer Hundred Talents Program" (E0241211H1) and startup program (Y8677211K1, Y8690211Z1) from State Key Laboratory of Phytochemistry and Plant Resources in West China, Kunming Institute of Botany, the Chinese Academy of Sciences to Dr. Shubai Liu. The roles of these grants were to support the activities of study design and data collection, analysis and interpretation, manuscript writing and publication cost.

Availability of Data and Materials

All data generated or analyzed during this study are included in this published article and its supplementary information files. The datasets generated and/or analyzed during this current study are available online (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE114 5).

Author's Contributions

SL and YH designed the overall project study; YH collected data, performed data analysis, and drafted the manuscript; SL, YH, YZ, JT and ZW interpreted and summarized the results; SL and YH wrote and revised the manuscript; all authors have read and approved the final version of manuscript. JT and ZW contributed equally.

Acknowledgement

Not applicable.

Ethics Approval and Consent to Participate

Not applicable.

Patient Consent for Publication

Not applicable.

References

1. Braunwald E (2015) The war against heart failure: the Lancet lecture. Lancet 385: 812-824.
2. Savarese G, Lund LH (2017) Global Public Health Burden of Heart Failure. Card Fail Rev 3: 7-11.
3. Maron BJ, Towbin JA, Thiene G, Antzelevitch C, Corrado D, et al. (2006) Contemporary definitions and classification of the cardiomyopathies: an American Heart Association Scientific Statement from the Council on Clinical Cardiology, Heart Failure and Transplantation Committee; Quality of Care and Outcomes Research and Functional Genomics and Translational Biology Interdisciplinary Working Groups; and Council on Epidemiology and Prevention. Circulation 113: 1807-1816.
4. Hwang JJ, Allen PD, Tseng GC, Lam CW, Fananapazir L, et al. (2002) Microarray gene expression profiles in dilated and hypertrophic cardiomyopathic end-stage heart failure. Physiol Genomics 10: 31-44.
5. Zhang B, Horvath S (2005) A general framework for weighted gene co-expression network analysis. Stat Appl Genet Mol Biol 4: 17.
6. Zhang J, Baddoo M, Han C, Strong MJ, Cvitanovic J, et al. (2016) Gene network analysis reveals a novel 22-gene signature of carbon metabolism in hepatocellular carcinoma. Oncotarget 7: 49232-49245.
7. Guo Y, Xing Y (2016) Weighted gene co-expression network analysis of pneumocytes under exposure to a carcinogenic dose of chloroprene. Life Sci 151: 339-347.
8. Wu D, Hu S, Hou Y, He Y, Liu S (2020) Identification of potential novel biomarkers to differentiate malignant thyroid nodules with cytological indeterminate. BMC cancer 20: 199.
9. Langfelder P, Horvath S (2008) WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics 9: 559.
10. Li A, Horvath S (2007) Network neighborhood analysis with the multi-node topological overlap measure. Bioinformatics 23: 222-231.
11. Wang JH, Zhao LF, Lin P, Su XR, Chen SJ, et al. (2014) GenCLiP 2.0: a web server for functional clustering of genes and construction of molecular networks based on free terms. Bioinformatics 30: 2534-2536.
12. Hayman GT, Laulederkind SJ, Smith JR, Wang SJ, Petri V, et al. (2016) The Disease Portals, disease-gene annotation and the RGD disease ontology at the Rat Genome Database. Database (Oxford) 2016: baw034.
13. Shimoyama M, De Pons J, Hayman GT, Laulederkind SJ, Liu W, et al. (2015) The Rat Genome Database 2015: genomic, phenotypic and environmental variations and disease. Nucleic Acids Res 43: D743-D750.
14. Chen P, Wang F, Feng J, Zhou R, Chang Y, et al. (2017) Co-expression network analysis identified six hub genes in association with metastasis risk and prognosis in hepatocellular carcinoma. Oncotarget 8: 48948-48958.
15. Ford MG, Pearse BM, Higgins MK, Vallis Y, Owen DJ, et al. (2001) Simultaneous binding of PtdIns(4,5)P2 and clathrin by AP180 in the nucleation of clathrin lattices on membranes. Science 291: 1051-1055.
16. Scotland PB, Heath JL, Conway AE, Porter NB, Armstrong MB, et al. (2012) The PICALM protein plays a key role in iron homeostasis and cell proliferation. PLoS One 7: e44252.
17. Tebar F, Bohlander SK, Sorkin A (1999) Clathrin assembly lymphoid myeloid leukemia (CALM) protein: localization in endocytic-coated pits, interactions with clathrin, and the impact of overexpression on clathrin-mediated traffic. Mol Biol Cell 10: 2687-2702.
18. Wang J, Liu S, Heallen T, Martin JF (2018) The Hippo pathway in the heart: pivotal roles in development, disease, and regeneration. Nat Rev Cardiol 15: 672-684.
19. Couzens AL, Xiong S, Knight JDR, Mao DY, Guettler S, et al. (2017) MOB1 Mediated Phospho-recognition in the Core Mammalian Hippo Pathway. Mol Cell Proteomics 16: 1098-1110.
20. Curti A, Lapucci C, Berto S, Prandstraller D, Perolo A, et al. (2016) Maternal plasma mRNA species in fetal heart defects: a potential for molecular screening. Prenat Diagn 36: 738-743.

21. Florindo C, Perdigao J, Fesquet D, Schiebel E, Pines J, et al. (2012) Human Mob1 proteins are required for cytokinesis by controlling microtubule stability. J Cell Sci 125(Pt 13): 3085-3090.

22. Ben Yehuda S, Dix I, Russell CS, Levy S, Beggs JD, et al. (1998) Identification and functional analysis of hPRP17, the human homologue of the PRP17/CDC40 yeast gene involved in splicing and cell cycle control. RNA 4: 1304-1312.

23. Li BX, Gardner R, Xue C, Qian DZ, Xie F, et al. (2016) Systemic Inhibition of CREB is Well-tolerated in vivo. Sci Rep 6: 34513.

24. Chien PT, Lin CC, Hsiao LD, Yang CM (2015) c-Src/Pyk2/EGFR/PI3K/Akt/CREB-activated pathway contributes to human cardiomyocyte hypertrophy: Role of COX-2 induction. Mol Cell Endocrinol 409: 59-72.

25. Chang CF, Chou HT, Lin YJ, Lee SJ, Chuang JL, et al. (2006) Structure of the subunit binding domain and dynamics of the di-domain region from the core of human branched chain alpha-ketoacid dehydrogenase complex. J Biol Chem 281: 28345-28353.

26. Ansari AA, Neckelmann N, Villinger F, Leung P, Danner DJ, et al. (1994) Epitope mapping of the branched chain alpha-ketoacid dehydrogenase dihydrolipoyl transacylase (BCKD-E2) protein that reacts with sera from patients with idiopathic dilated cardiomyopathy. J Immunol 153: 4754-4765.

27. Chen SL, Dowhan DH, Hosking BM, Muscat GE (2000) The steroid receptor coactivator, GRIP-1, is necessary for MEF-2C-dependent gene expression and skeletal muscle differentiation. Genes Dev 14: 1209-1228.

28. Yu B, Barbalic M, Brautbar A, Nambi V, Hoogeveen RC, et al. (2013) Association of genome-wide variation with highly sensitive cardiac troponin-T levels in European Americans and Blacks: a meta-analysis from atherosclerosis risk in communities and cardiovascular health studies. Circ Cardiovasc Genet 6: 82-88.

29. Zhang L, Zhang W (2018) Knockdown of NUDT21 inhibits proliferation and promotes apoptosis of human K562 leukemia cells through ERK pathway. Cancer Manag Res 10: 4311-4323.

30. Maffei A, Lembo G, Carnevale D (2018) PI3Kinases in Diabetes Mellitus and Its Related Complications. Int J Mol Sci 19: 4098.

31. Tan B, Liu M, Yang Y, Liu L, Meng F (2019) Low expression of PIK3C2A gene: A potential biomarker to predict the risk of acute myocardial infarction. Medicine (Baltimore) 98: e15061.

32. Nakielny S, Dreyfuss G (1996) The hnRNP C proteins contain a nuclear retention sequence that can override nuclear export signals. J Cell Biol 134: 1365-1373.

33. Williamson DJ, Banik-Maiti S, DeGregori J, Ruley HE (2000) hnRNP C is required for postimplantation mouse development but Is dispensable for cell viability. Mol Cell Biol 20: 4094-4105.

34. Markson G, Kiel C, Hyde R, Brown S, Charalabous P, et al. (2009) Analysis of the human E2 ubiquitin conjugating enzyme protein interaction network. Genome Res 19: 1905-1911.

35. Zhu Y, Yang T, Duan J, Mu N, Zhang T (2019) MALAT1/miR-15b-5p/MAPK1 mediates endothelial progenitor cells autophagy and affects coronary atherosclerotic heart disease via mTOR signaling pathway. Aging (Albany NY) 11: 1089-1109.

36. Zhang D, Li Y, Heims-Waldron D, Bezzerides V, Guatimosim S, et al. (2018) Mitochondrial Cardiomyopathy Caused by Elevated Reactive Oxygen Species and Impaired Cardiomyocyte Proliferation. Circ Res 122: 74-87.

37. Mandl A, Huong Pham L, Toth K, Zambetti G, Erhardt P (2011) Puma deletion delays cardiac dysfunction in murine heart failure models through attenuation of apoptosis. Circulation 124: 31-39.

38. Li Q, Cui HH, Yang YJ, Li XD, Chen GH, et al. (2017) Quantitative Proteomics Analysis of Ischemia/Reperfusion Injury-Modulated Proteins in Cardiac Microvascular Endothelial Cells and the Protective Role of Tongxinluo. Cell Physiol Biochem 41: 1503-1518.

39. Sucharov CC, Mariner P, Long C, Bristow M, Leinwand L (2003) Yin Yang 1 is increased in human heart failure and represses the activity of the human alpha-myosin heavy chain promoter. J Biol Chem 278: 31233-31239.

40. Zhao Z, Guan JZ, Wu M, Lai GH, Zhu ZL (2019) Downregulation of microRNA-23b protects against ischemia-reperfusion injury via p53 signaling pathway by upregulating MDM4 in rats. J Cell Biochem 120: 4599-4612.

Citation: Liu S, He YY, Wu ZY, Tian J (2021) Novel Signature Genes and Pathways Identified for Human Left Ventricle Cardiomyopathies Rise from different Etiologies. J Biomed Res Rev Vol: 4, Issu: 1. (31-50).