

# Exploration and Validation of Hub Genes and Pathways in the Progression of Hypoplastic Left Heart Syndrome via Weighted Gene Co-Expression Network Analysis

**Xuelan Liu**

Central China Fuwai Hospital of Zhengzhou University <https://orcid.org/0000-0002-9230-1438>

**Honglei Shang**

The Third Affiliated Hospital of Zhengzhou University

**Bin Li**

Central China Fuwai Hospital of Zhengzhou University

**Liyun Zhao**

Central China Fuwai Hospital of Zhengzhou University

**Ying Hua**

Central China Fuwai Hospital of Zhengzhou University

**Kaiyuan Wu**

Central China Fuwai Hospital of Zhengzhou University

**Manman Hu**

Central China Fuwai Hospital of Zhengzhou University

**Taibing Fan** (✉ [fantaibing@163.com](mailto:fantaibing@163.com))

Central China Fuwai Hospital of Zhengzhou University

---

## Research article

**Keywords:** hypoplastic left heart syndrome, weighted gene co-expression network analysis, protein-protein interaction network, hub genes, pathways

**Posted Date:** November 17th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-105782/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

**Version of Record:** A version of this preprint was published at BMC Cardiovascular Disorders on June 15th, 2021. See the published version at <https://doi.org/10.1186/s12872-021-02108-0>.

# Abstract

**Objective:** Despite significant progress in surgical treatment of hypoplastic left heart syndrome (HLHS), its mortality and morbidity are still high. Little is known about the molecular abnormalities of the syndrome. In this study, we aimed to probe into hub genes and key pathways in the progression of the syndrome.

**Methods:** Differentially expressed genes (DEGs) were identified in left ventricle (LV) or right ventricle (RV) tissues between HLHS and controls using the GSE77798 dataset. Then, weighted gene co-expression network analysis (WGCNA) was performed and key modules were constructed for HLHS. Based on the genes in the key modules, protein-protein interaction (PPI) networks were conducted, and hub genes and key pathways were screened. Finally, the GSE23959 dataset was used to validate hub genes between HLHS and controls.

**Results:** 88 and 41 DEGs were identified for LV and RV tissues between HLHS and controls, respectively. DEGs in LV tissues of HLHS were distinctly involved in heart development, apoptotic signaling pathway and ECM receptor interaction. DEGs in RV tissues of HLHS were mainly enriched in BMP signaling pathway, regulation of cell development and regulation of blood pressure. A total of 16 co-expression network were constructed. Among them, black module ( $r=0.79$  and  $p\text{-value}=2e-04$ ) and pink module ( $r=0.84$  and  $p\text{-value}=4e-05$ ) had the most significant correlation with HLHS, indicating that the two modules could be the most relevant for HLHS progression. We identified five hub genes in the black module (including *Fbn1*, *Itga8*, *Itga11*, *Itgb5* and *Thbs2*), and five hub genes (including *Cblb*, *Ccl2*, *Edn1*, *Itgb3* and *Map2k1*) in the pink module for HLHS. Their abnormal expression was verified in the GSE23959 dataset.

**Conclusion:** Our findings revealed hub genes and key pathways for HLHS through WGCNA, which could play key roles in the molecular mechanism of HLHS.

## Introduction

Hypoplastic left heart syndrome (HLHS) is a group of complex congenital heart malformations characterized by severe stenosis or atresia of the aortic valve and mitral valve, ascending aorta, and left ventricular dysplasia [1–3]. It was first proposed by Noonan and Nadas in 1958 [4]. HLHS accounts for 1.4–4.1% of congenital cardiovascular malformations, and the prevalence of live births is about 2:10000–3:10000 [5]. The dysplastic left ventricle cannot provide adequate systemic circulation perfusion and oxygen supply. Oxygenated blood bypasses the dysplastic left heart to supply the whole body through an atrial septal defect and an open arterial catheter, followed by symptoms such as strenuous breastfeeding, difficulty breathing, rapid heartbeat, pulse weakness, dull skin or cyanosis, and severe heart failure. HLHS accounts for 25% of all babies who die of congenital heart diseases. Without diagnosis and treatment, 95% of babies die within one month after birth [6]. The treatment of HLHS remains extremely challenging. Patients usually undertake aggressive palliative operations. Despite the diagnosis and treatment of the

disease have made great progress since the first implementation of Norwood surgery in 1983, the morbidity and mortality of HLHS patients are still high [7].

The exact cause of HLHS is still unclear. Increasing evidence suggests that genetic pathology is involved in the progression of HLHS. Approximately 30% children with HLHS have genetic syndrome or other extracardiac abnormalities [8]. Various syndromes caused by chromosomal abnormalities are also associated with HLHS, including Turner syndrome (X chromosome monomer), Edwards syndrome (Trisomy 18) and DiGeorge syndrome (22q11.21 deletion) [9]. It has been reported that several specific genes are related to HLHS such as HAND1, TBX5, FOXC2, GJA1, NKX2-5, NOTCH1, MYH6 and ERBB4 [10]. Theis et al. demonstrated that the compound heterozygous mutation of the allele NOTCH1 is the basis of impaired cardiac development in patients with HLHS [11]. During human embryonic development, TAB2 is abundantly expressed in the endocardial pad, which plays an important role in the outflow tract and valve formation. TAB2 haploinsufficiency is also a risk factor for HLHS [10]. Therefore, genomics analysis is of great significance for the precise management and treatment of HLHS.

WGCNA, as a systematic biological method, may describe the pattern of gene association between different samples [12]. It can be used to identify highly synergistically changing gene sets, and to identify potential biomarkers and therapeutic targets based on the interconnectivity of gene sets and the association between gene sets and phenotypes. Compared with only focusing on DEGs, WGCNA utilizes the information of thousands of genes with the most changes or all genes to identify gene sets of interest and to perform correlation analysis with phenotypes [13]. The advantages of WGCNA are as follows: one is to make full use of information, and the other is to convert the association of thousands of genes and phenotypes into the association of several gene sets and phenotypes, eliminating the problem of multiple hypothesis test correction [14]. In this study, we firstly constructed a co-expression network for HLHS. Furthermore, we identified hub genes and pathways for HLHS progression, which deserve further research in more basic experiments and clinical research.

## Materials And Methods

### Data acquisition and preprocessing

The workflow of this study is shown in Fig. 1. Two HLHS mRNA expression profiling datasets were downloaded from the Gene Expression Omnibus (GEO; <https://www.ncbi.nlm.nih.gov/geo>). GSE77798 RNA-seq expression profiling dataset was composed of 16 left ventricle (LV) and right ventricle (RV) from HLHS mice (n = 6) and littermate controls (n = 10). The dataset was based on GPL13112 Illumina HiSeq 2000 (*Mus musculus*) [15]. GSE23959 microarray expression profiling dataset included 6 RV samples from HLHS neonates and 10 LV and RV from controls on the GPL5188 [HuEx-1\_0-st] Affymetrix Human Exon 1.0 ST Array [probe set (exon) version] platform. Quality control was achieved via arrayQualityMetrics package in R [16]. Robust multi-array average (RMA) background correction was performed on the raw expression data. The processed signals were transformed into log<sub>2</sub>, followed by

quantile normalization. The robust K-Nearest Neighbor (KNN) algorithm was executed to process missing values [17].

## Differential expression analysis

Differentially expressed genes (DEGs) were screened between HLHS and control samples using Linear Models for Microarray Data (limma) package in R [18]. False discovery rate (FDR) was calculated via Benjamini-Hochberg method.  $\log_2|\text{fold change (FC)}| > 1$  and  $\text{FDR} < 0.05$  were set as the threshold value.

## Functional enrichment analysis

Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) were analyzed via MetaScape online database (<http://metascape.org/gp/index.html>) [19]. Metascape integrates more than 40 bioinformatics knowledge bases into a single user interface, which provides comprehensive gene list annotation and analysis resources.

## WGCNA

WGCNA package in R was utilized to construct co-expression network [20]. The gene co-expression similarity matrix was composed of the absolute values of the correlation coefficients between genes. For continuous variables of genes, Pearson correlation coefficients were used (the range of correlation coefficients is [0, 1]), and the correlation matrix was as follows:  $S = [S_{ij}]$  (i, j refers to the i-th, j-th gene). Soft threshold value was calculated through an exponential adjacency function, as follows:  $a_{ij} = \text{power}(S_{i,j}, \beta) = |S_{i,j}|^\beta$  ( $a_{ij}$  is the adjacency function between the i-th and j-th genes;  $\beta$  refers to soft threshold value). To make the co-expression network in line with the characteristics of scale-free network, soft threshold value was screened under the threshold of  $R^2 > 0.8$ . Then, correlation matrix  $S = [S_{ij}]$  was converted into adjacency matrix  $A = [A_{ij}]$  by pickSoftThreshold function. The topological overlap dissimilarity measure (TOM) was used to calculate the degree of correlation between genes, as follows:

$$TOM_{IJ} = \frac{\sum_u a_{iu} a_{uj} + a_{ij}}{\min(k_i, k_j) + 1 - a_{ij}}$$
 ( $a_{ij}$  is [0, 1]). Gene modules were assigned based on the degree of connection between modules. Therefore, TOM was converted into dissimilarity degree, as follows:  $\text{dissTOM}_{ij} = 1 - \text{TOM}_{ij}$ . Then, a hierarchical clustering tree diagram of genes was constructed and gene modules were assigned using dynamic cutting method. Afterwards, we calculated eigengene that refers to the first principal component of all gene expression level vectors in the modules. Correlation between eigengene of each module and clinical traits including HLHS and heart region was analyzed.

## PPI network

Interested genes were imported into STRING online database (version 11; <http://string-db.org/>) [21]. Protein-protein interactions were visualized using Cytoscape (version 3.8.0) [22].

## External dataset validation

The expression patterns of hub genes between HLHS and controls were validated using an external GSE23959 microarray expression profiling dataset.

## Results

### Identification of DEGs and pathways for LV and RV HLHS

To eliminate the changes in intensity caused by the experimental technique, and to make the data of each sample and parallel experiment at the same level, GSE77798 RNA-seq expression profiling data were pre-processed, including data filtering, normalization, logarithmization, and estimation of missing values. GSE77798 RNA-seq expression profiling was processed by quantile normalization (Fig. 2A). Following normalization, under the threshold of  $FDR < 0.05$  and  $\log_2|FC| > 1$ , we identified 44 up-regulated and 44 down-regulated genes between LV HLHS and controls (Fig. 2B). Furthermore, 21 up-regulated and 20 down-regulated genes were screened between RV HLHS and controls. KEGG enrichment analysis results showed that DEGs in LV HLHS were distinctly enriched in several key pathways including connective tissue development, BMP signaling pathway, skeletal muscle contraction, intermediate filament-based process, apoptotic signaling pathway, extracellular matrix (ECM)-receptor interaction, cellular component assemble, mechanical stimulus, regulation of ion transport, organelle localization, response to hypoxia, antigen processing and presentation and cell adhesion molecules (CAMs) in Fig. 2C. Moreover, DEGs in RV HLHS were significantly involved in regulation of catalytic activity, extracellular matrix organization, BMP signaling pathway, positive regulation of cell development, hematopoietic progenitor cell differentiation, regulation of blood pressure and cell surface receptor signaling (Fig. 2D). GO enrichment analysis including biological process (BP), cellular component (CC) and molecular component (MF) was performed. As shown in Fig. 2E, DEGs in LV HLHS were most significantly enriched in regulation of ion transmembrane transport, extracellular matrix and actin binding. Furthermore, DEGs in RV HLHS were most significantly enriched in protein serine/threonine kinase signaling pathway, extracellular matrix and glycosaminoglycan binding (Fig. 2F).

### Construction of a co-expression network and key modules for HLHS

As shown in Fig. 3A, there were no outlier samples. To ensure a sparse-free network, soft threshold value ( $\beta$ ) was set as 8 ( $R^2 = 0.8$ ) in Fig. 3B. DEGs with similar expression patterns were assigned into modules by average link clustering and dynamic tree cutting methods (Fig. 3C). Finally, 16 co-expression modules were conducted for HLHS. Among them, black module ( $r = 0.79$  and  $p\text{-value} = 2e-04$ ) and pink module ( $r = 0.84$  and  $p\text{-value} = 4e-05$ ) had the most significant correlation with HLHS in Fig. 3D. Thus, the two modules were considered the most relevant for HLHS progression. There were 249 genes in the black module and 222 genes in the pink module. Figure 3E shows the relationships among different modules through an eigengene adjacency heatmap. To further validate the roles of the two modules in HLHS progression, we carried out KEGG pathway enrichment and PPI network analyses. In Fig. 3F, the genes in the black module were most significantly enriched in  $TGF\beta$  in extracellular matrix, collagen metabolic

process, ECM receptor interaction, PI3K-Akt pathway and dilated cardiomyopathy, especially extracellular matrix organization. The genes in the extracellular matrix organization pathways were used to conduct a PPI network. We found that *Fbn1* (degree = 13), *Thbs2* (degree = 11), *Itga8* (degree = 9), *Itga11* (degree = 9) and *Itgb5* (degree = 9) could be hub genes for HLHS. For the genes in the pink module, heart development, heart valve morphogenesis and cardiac chamber development were the most significantly enriched pathways (Fig. 3G). Among the genes in the three pathways, five hub genes were identified including *Map2k1* (degree = 8), *Ccl2* (degree = 7), *Itgb3* (degree = 5), *Edn1* (degree = 4) and *Cblb* (degree = 4) through PPI network.

## Identification and validation of hub genes for HLHS

We examined the expression patterns of these hub genes in the black and pink modules from PPI networks between HLHS and controls in the GSE77798 dataset. For hub genes in the black module, compared to LV controls, the expression of *Fbn1*, *Itga11* and *Itgb5* were significantly higher in RV controls, LV HLHS and RV HLHS (all p-values < 0.05) in the GSE77798 dataset (Fig. 4A). Furthermore, *Itga8* and *Thbs2* had higher expression levels in LV HLHS and RV HLHS than in LV controls (all p-values < 0.05). For hub genes in the pink module, compared to LV controls, the expression *Cblb* was distinctly higher in RV HLHS (p-value < 0.05). *Ccl2* had a significantly higher expression in LV HLHS compared to LV control (p-value < 0.05). In comparison to LV controls, the expression of *Itgb3* was significantly higher in LV HLHS (p-value < 0.05). However, its expression was distinctly lower in RV controls and RV HLHS compared to LV controls (both p-value < 0.05). *Edn1* and *Map2k1* exhibited higher expression levels in LV and RV HLHS than in LV controls (all p-values < 0.05).

After normalization (Fig. 4B), the expression patterns of these hub genes were further validated in the GSE23959 dataset. Compared to LV controls, the expression of *Fbn1*, *Itga8*, *Itga11* and *Itgb5* had higher expression levels in RV controls and RV HLHS (all p-values < 0.05) in Fig. 4C. Furthermore, *Thbs2* expression was significantly higher in RV HLHS in comparison to LV controls (p-value < 0.05). Compared with LV controls, the expression of *Cblb* and *Edn1* in RV HLHS (both p-values < 0.05). *Itgb3* had lower expression levels both in RV controls and RVHLHS in comparison to LV controls (all p-values < 0.05). Moreover, *Map2k1* had a lower expression level in RV HLHS than in LV controls (p-value < 0.05). KEGG pathway enrichment analysis of all DEGs was presented using the GSE23959 dataset. As shown in Fig. 4D, LV DEGs were most significantly enriched in the heart development and extracellular structure organization. Additionally, cell-matrix adhesion and extracellular structure organization were the most significant pathways for RV DEGs (Fig. 4E).

## Discussion

HLHS is characterized by LV hypoplasia and increased biomechanical pressure on RV by single ventricular physiology [23]. By analyzing HLHS gene RNA-seq profile data, we obtained 14889 genes for HLHS. The expression profiles of these genes were used as data sources to perform WGCNA on HLHS and a total of 16 co-expressed gene modules were identified in this study. Among them, black module and pink module were the two modules most relevant to HLHS. Based on the genes in these two modules,

we constructed PPI networks. Finally, 10 hub genes were confirmed for HLHS, which were differentially expressed in LV and RV tissues between controls and HLHS. By verifying these hub genes in the HLHS expression profile data of an independent cohort, we found that most of the genes were consistent in the HLHS cohort of different data sources, indicating that our analysis method was accurate and reproducible.

We identified five hub genes for HLHS in the black module, including *Fbn1*, *Itga8*, *Itga11*, *Itgb5* and *Thbs2*. *Fbn1* (Fibrillin-1) has been found to be associated with heart development [24]. Its mutation could increase genetic susceptibility to thoracic aortic aneurysms [25]. Furthermore, its mutation leads to Marfan syndrome (MFS) that is the most common hereditary connective tissue disease [26]. In this study, *Fbn1* expression was distinctly up-regulated in LV/RV HLHS compared to controls. Furthermore, there was a significant difference in *Fbn1* expression between LV and RV for control mice. *Itga8* (Integrin Subunit Alpha 8) inhibits NF $\kappa$ B and JAK-STAT signaling and cardiac injury in myocardium without stress [27]. *Itga8* expression was significantly higher in LV/RV HLHS in comparison to controls both in the GSE77798 and GSE23959 datasets. *Itga11* (Integrin Subunit Alpha 11) expression has been detected to be increased in methylglyoxal-induced collagen-treated human cardiac fibroblasts and streptozotocin-treated Sprague-Dawley rat cardiac fibroblasts, which may promote the formation of pre-fibrotic fibroblasts and fibrotic stroma in diabetic cardiomyopathy [28]. *Itgb5* (Integrin Subunit Beta 5) has been identified to be in significant correlation with coronary artery disease and age-dependent organ fibrosis [29, 30]. We found that *Fbn1* had distinctly higher expression in LV/RV HLHS in comparison to controls. Moreover, a significant difference in *Fbn1* expression was detected between LV and RV for control mice and neonates. *Thbs2* (Thrombospondin 2) mediates cell-matrix interactions, vascular integrity and thrombosis [31]. In our study, its expression was lower in LV HLHS than controls, which was higher in RV HLHS compared to controls.

Five hub genes including *Cblb*, *Ccl2*, *Edn1*, *Itgb3* and *Map2k1* were screened for HLHS in the pink module. *Cblb* (Casitas B-cell lymphoma-B) is lowly expressed in plaques for human atherosclerosis, thereby leading to CD8<sup>+</sup> T cell-induced macrophage death and accelerating atherosclerosis [32]. Our findings revealed that *Cblb* expression was significantly down-regulated in RV HLHS not LV HLHS in comparison to controls. It has been reported that targeting *Ccl2* (C-C Motif Chemokine Ligand 2) could ameliorate atherosclerosis [33]. Moreover, Dectin-2-mediated *Ccl2* in resident tissue macrophages can facilitate cardiac arteritis [34]. In mice, *Ccl2* expression was up-regulated in LV HLHS than controls. However, no significant difference in *Ccl2* expression was detected between HLHS and controls. *Edn1* (Endothelin 1) genetic locus is correlated to spontaneous coronary artery dissection [35]. It was highly expressed in LV/RV HLHS compared to controls in mice, and was highly expressed in RV HLHS than control neonates. *Itgb3* (Integrin Subunit Beta 3) is related to myocardial infarction risk [36]. *Map2k1* (Mitogen-Activated Protein Kinase Kinase 1) mutation is often in association with the clinical phenotype of the cardiovascular system skin syndrome [37]. Our study found that *Map2k1* was highly expressed in LV HLHS and lowly expressed in RV HLHS compared to controls in mice. Furthermore, its low expression was found in human RV HLHS in comparison to controls.

Complex life phenomena are the result of the interaction of a large number of biological components. Biological research has shifted from collecting gene and protein information to systematically using this information to clarify the synergy between them. In this study, we tried to probe into the molecular mechanism of HLHS through functional enrichment analysis of DEGs-related HLHS. DEGs in LV HLHS were distinctly involved in heart development, apoptotic signaling pathway and ECM receptor interaction. Abnormally expressed genes related to heart development could contribute to the progression of HLHS. Furthermore, imbalance of apoptotic signaling pathway in cardiomyocytes may be an important factor of HLHS. As a previous study, RV tissues in HLHS exhibit immature ECM and increased cardiomyocyte apoptosis [38]. Thus, the roles of these DEGs in HLHS need further exploration. DEGs in RV HLHS were mainly enriched in BMP signaling pathway, regulation of cell development and regulation of blood pressure. Dysregulation of the BMP pathway is the basis of many diseases of different organ systems in humans [39]. As a previous study, changes in gene expression in the BMP pathway has been found in RV tissues for neonates with HLHS [23]. Cardiomyocytes from neonates with HLHS exhibit multiple expression and function differences [40], which could be mediated by a variety of DEGs at a transcriptional level [41]. Our study found that DEGs in RV HLHS were involved in the regulation of blood pressure, as previous studies [42].

Facing the increasing amount of high-throughput data, it is a difficult problem about how to effectively extract useful information to obtain the regulatory relationship between genes in the research of systems biology. The regulatory relationship between genes has spatiotemporal specificity. In different organs, different physiological conditions and pathological states, and at different time points, this regulatory relationship will change accordingly. It is these changes that determine cell proliferation, differentiation, as well as occurrence, development of HLHS. The modularity of the biological network is the result of living organisms to achieve specific biological functions. Modularity provides us with a simple and effective method to understand the regulatory relationship between genes, which is an indispensable method in the research of systems biology. This study is the first to analyze HLHS data by WGCNA. Our results showed that WGCNA can discover biologically meaningful gene modules, and the hub genes related to the clinical information found are consistent with literature reports, which also proves the accuracy and effectiveness of WGCNA of gene expression data. Further excavation of information on gene modules will help us better understand the role and significance of hub genes, key signaling pathways, as well as the regulatory mechanisms between genes on the development of HLHS.

## Conclusion

In this study, we identified hub genes and key pathways for HLHS via WGCNA. These findings showed that these hub genes could play an important role in HLHS and cardiovascular diseases, which provided important clues for further revealing the molecular mechanism of HLHS.

## Abbreviations

HLHS

hypoplastic left heart syndrome; DEGs:differentially expressed genes; LV:left ventricle; RV:right ventricle; WGCNA:weighted gene co-expression network analysis; PPI:protein-protein interaction; GEO:Gene Expression Omnibus; FDR:false discovery rate; FC:fold change; GO:Gene Ontology; KEGG:Kyoto Encyclopedia of Genes and Genomes; TOM:topological overlap dissimilarity measure; BP:biological process; CC:cellular component; MF:molecular component.

## Declarations

### Authors' contributions

XL, HS and BL: takes responsibility for all aspects of the reliability and freedom from bias of the data presented and their discussed interpretation, drafting the article.

LZ, YH, KW and MH: takes responsibility for statistical analyses, and interpretation of data. TF: takes responsibility for full text evaluation and guidance, final approval of the version to be submitted. All authors read and approved the final manuscript.

### Conflict of Interests:

The authors have no conflict of interest to disclose.

### Acknowledgments

This work was supported by the Zhengzhou high-level Clinical Pivot Subjects Construction project, China (Grant No. 201601023 and SBGJ2018074).

### Ethics approval and consent to participate

Not applicable.

### Consent for publication

Not applicable.

### Availability of data and materials

All data generated or analysed during this study are included in this article.

### Corresponding Authors:

Taibing Fan, Department of Children's Heart Center, Henan Provincial People's Hospital, Department of Children's Heart Center of Fuwai Central China Cardiovascular Hospital, Central China Fuwai Hospital of Zhengzhou University. Email: [fantaibing@163.com](mailto:fantaibing@163.com)

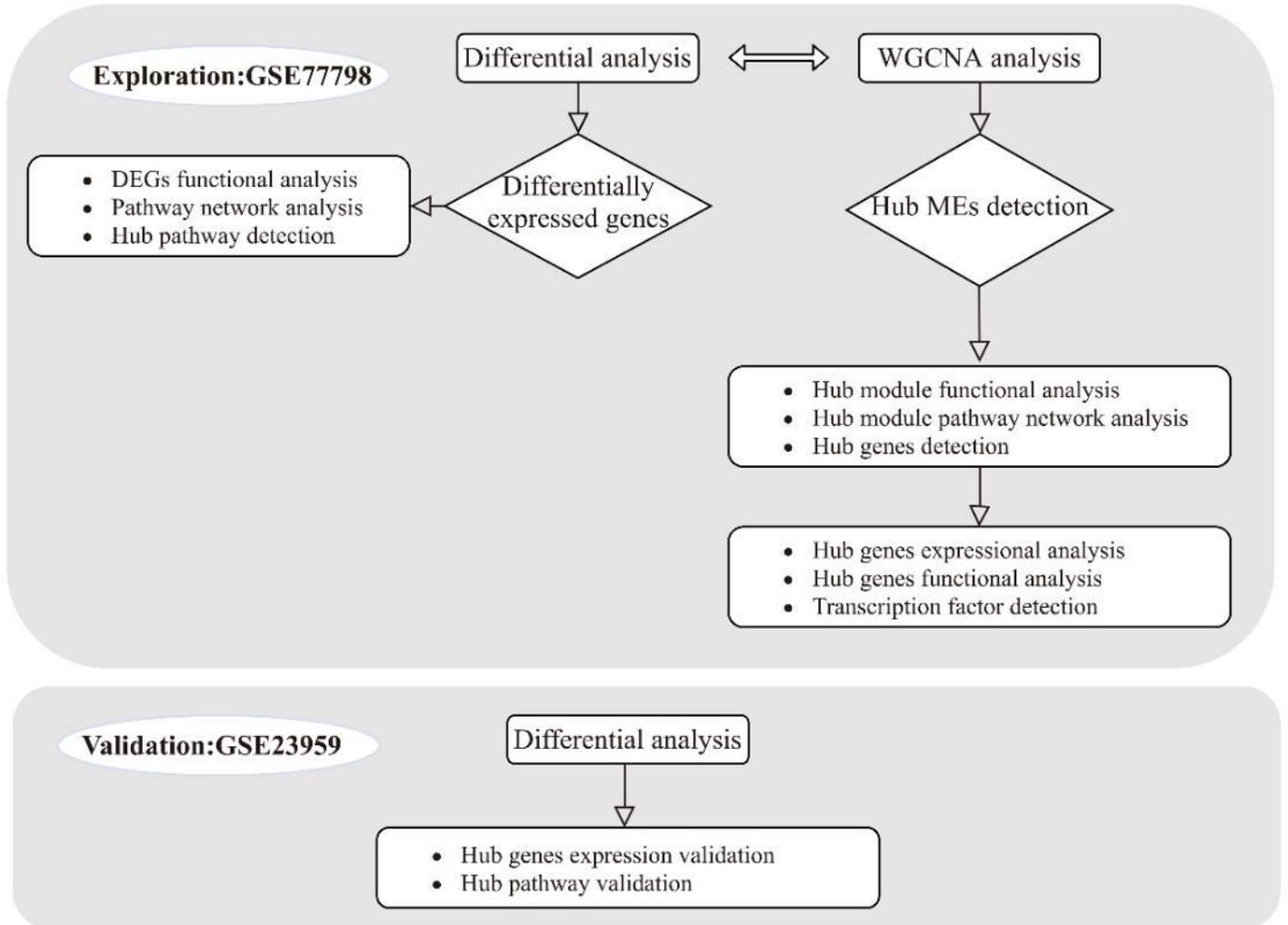
## References

1. D'Souza R, Wang Y, Calderon-Anyosa RJC, Montero AE, Banerjee MM, Ekhomu O, Matsubara D, Mercer-Rosa L, Agger P, Sato T et al: Decreased right ventricular longitudinal strain in children with hypoplastic left heart syndrome during staged repair and follow-up: does it have implications in clinically stable patients? *Int J Cardiovasc Imaging* 2020.
2. Garcia AM, Beatty JT, Nakano SJ: Heart failure in single right ventricle congenital heart disease: physiological and molecular considerations. *Am J Physiol Heart Circ Physiol* 2020, 318(4):H947-h965.
3. Kelly TJ, Zannino D, Brink J, Konstantinov IE, Cheung MM, d'Udekem Y, Brizard CP: A shunt decision-making protocol in the surgical palliation of hypoplastic left heart syndrome from 2004 to 2016. *Eur J Cardiothorac Surg* 2020, 58(1):153-162.
4. Noonan JA, Nadas AS: The hypoplastic left heart syndrome; an analysis of 101 cases. *Pediatr Clin North Am* 1958, 5(4):1029-1056.
5. Javed R, Cetta F, Said SM, Olson TM, O'Leary PW, Qureshi MY: Hypoplastic Left Heart Syndrome: An Overview for Primary Care Providers. *Pediatr Rev* 2019, 40(7):344-353.
6. Feinstein JA, Benson DW, Dubin AM, Cohen MS, Maxey DM, Mahle WT, Pahl E, Villafañe J, Bhatt AB, Peng LF et al: Hypoplastic left heart syndrome: current considerations and expectations. *J Am Coll Cardiol* 2012, 59(1 Suppl):S1-42.
7. Lowenstein S, Macauley R, Perko K, Ronai C: Provider perspective on the role of palliative care in hypoplastic left heart syndrome. *Cardiol Young* 2020, 30(3):377-382.
8. Natowicz M, Chatten J, Clancy R, Conard K, Glauser T, Huff D, Lin A, Norwood W, Rorke LB, Uri A et al: Genetic disorders and major extracardiac anomalies associated with the hypoplastic left heart syndrome. *Pediatrics* 1988, 82(5):698-706.
9. Warburton D, Ronemus M, Kline J, Jobanputra V, Williams I, Anyane-Yeboah K, Chung W, Yu L, Wong N, Awad D et al: The contribution of de novo and rare inherited copy number changes to congenital heart disease in an unselected sample of children with conotruncal defects or hypoplastic left heart disease. *Hum Genet* 2014, 133(1):11-27.
10. Cheng A, Neufeld-Kaiser W, Byers PH, Liu YJ: 6q25.1 (TAB2) microdeletion is a risk factor for hypoplastic left heart: a case report that expands the phenotype. *BMC Cardiovasc Disord* 2020, 20(1):137.
11. Theis JL, Hrstka SC, Evans JM, O'Byrne MM, de Andrade M, O'Leary PW, Nelson TJ, Olson TM: Compound heterozygous NOTCH1 mutations underlie impaired cardiogenesis in a patient with hypoplastic left heart syndrome. *Hum Genet* 2015, 134(9):1003-1011.
12. Niu X, Zhang J, Zhang L, Hou Y, Pu S, Chu A, Bai M, Zhang Z: Weighted Gene Co-Expression Network Analysis Identifies Critical Genes in the Development of Heart Failure After Acute Myocardial Infarction. *Front Genet* 2019, 10:1214.
13. Chen R, Ge T, Jiang W, Huo J, Chang Q, Geng J, Shan Q: Identification of biomarkers correlated with hypertrophic cardiomyopathy with co-expression analysis. *J Cell Physiol* 2019, 234(12):21999-22008.

14. Wang T, Zheng X, Li R, Liu X, Wu J, Zhong X, Zhang W, Liu Y, He X, Liu W et al: Integrated bioinformatic analysis reveals YWHAB as a novel diagnostic biomarker for idiopathic pulmonary arterial hypertension. *J Cell Physiol* 2019, 234(5):6449-6462.
15. Liu X, Yagi H, Saeed S, Bais AS, Gabriel GC, Chen Z, Peterson KA, Li Y, Schwartz MC, Reynolds WT et al: The complex genetics of hypoplastic left heart syndrome. *Nat Genet* 2017, 49(7):1152-1159.
16. Kauffmann A, Gentleman R, Huber W: arrayQualityMetrics—a bioconductor package for quality assessment of microarray data. *Bioinformatics* 2009, 25(3):415-416.
17. Huttenhower C, Flamholz AI, Landis JN, Sahi S, Myers CL, Olszewski KL, Hibbs MA, Siemers NO, Troyanskaya OG, Collier HA: Nearest Neighbor Networks: clustering expression data based on gene neighborhoods. *BMC Bioinformatics* 2007, 8:250.
18. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK: limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic acids research* 2015, 43(7):e47.
19. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, Benner C, Chanda SK: Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nature communications* 2019, 10(1):1523.
20. Langfelder P, Horvath S: WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 2008, 9:559.
21. Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, Santos A, Doncheva NT, Roth A, Bork P et al: The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res* 2017, 45(D1):D362-d368.
22. Doncheva NT, Morris JH, Gorodkin J, Jensen LJ: Cytoscape StringApp: Network Analysis and Visualization of Proteomics Data. *J Proteome Res* 2019, 18(2):623-632.
23. Ricci M, Mohapatra B, Urbiztondo A, Birusingh RJ, Morgado M, Rodriguez MM, Lincoln J, Vatta M: Differential changes in TGF- $\beta$ /BMP signaling pathway in the right ventricular myocardium of newborns with hypoplastic left heart syndrome. *J Card Fail* 2010, 16(8):628-634.
24. Molck MC, Simioni M, Paiva Vieira T, Sgardioli IC, Paoli Monteiro F, Souza J, Fett-Conte AC, Félix TM, Lopes Monlléo I, Gil-da-Silva-Lopes VL: Genomic imbalances in syndromic congenital heart disease. *J Pediatr (Rio J)* 2017, 93(5):497-507.
25. Isselbacher EM, Lino Cardenas CL, Lindsay ME: Hereditary Influence in Thoracic Aortic Aneurysm and Dissection. *Circulation* 2016, 133(24):2516-2528.
26. Ramirez F, Caescu C, Wondimu E, Galatioto J: Marfan syndrome; A connective tissue disease at the crossroads of mechanotransduction, TGF $\beta$  signaling and cell stemness. *Matrix Biol* 2018, 71-72:82-89.
27. Ashton KJ, Reichelt ME, Mustafa SJ, Teng B, Ledent C, Delbridge LM, Hofmann PA, Morrison RR, Headrick JP: Transcriptomic effects of adenosine 2A receptor deletion in healthy and endotoxemic murine myocardium. *Purinergic Signal* 2017, 13(1):27-49.
28. Talior-Volodarsky I, Connelly KA, Arora PD, Gullberg D, McCulloch CA:  $\alpha$ 11 integrin stimulates myofibroblast differentiation in diabetic cardiomyopathy. *Cardiovasc Res* 2012, 96(2):265-275.

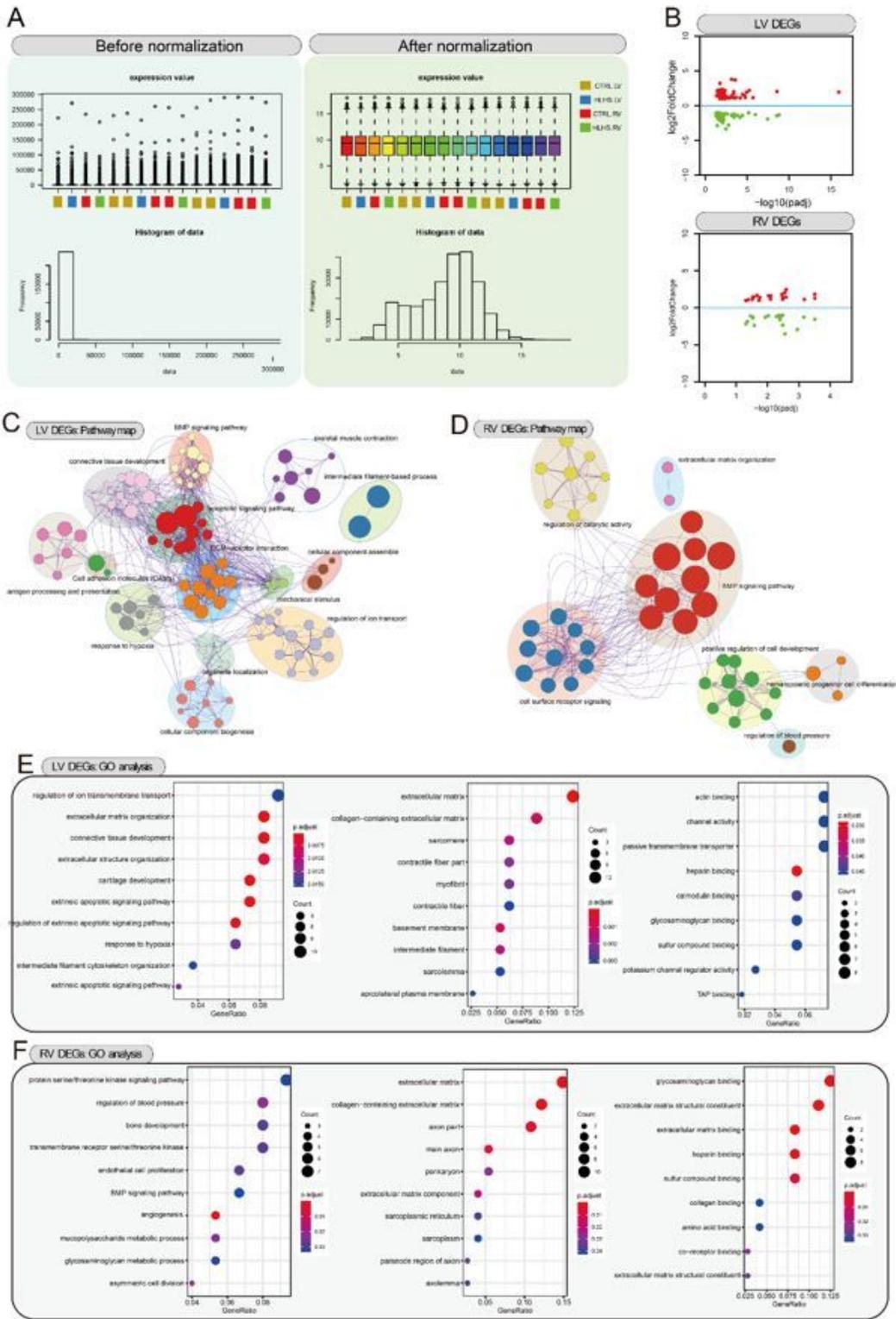
29. Cairns JT, Habgood A, Edwards-Pritchard RC, Joseph C, John AE, Wilkinson C, Stewart ID, Leslie J, Blaxall BC, Susztak K et al: Loss of ELK1 has differential effects on age-dependent organ fibrosis. *Int J Biochem Cell Biol* 2020, 120:105668.
30. Verweij N, Eppinga RN, Hagemeyer Y, van der Harst P: Identification of 15 novel risk loci for coronary artery disease and genetic risk of recurrent events, atrial fibrillation and heart failure. *Sci Rep* 2017, 7(1):2761.
31. Guo Y, Nan X, Zhang X, Wang G, Ren Y, Wang Y, Fu Y, Hou J: Molecular characterization and functional analysis of Japanese flounder (*Paralichthys olivaceus*) thbs2 in response to lymphocystis disease virus. *Fish Shellfish Immunol* 2019, 93:183-190.
32. Seijkens TTP, Poels K, Meiler S, van Tiel CM, Kusters PJH, Reiche M, Atzler D, Winkels H, Tjwa M, Poelman H et al: Deficiency of the T cell regulator Casitas B-cell lymphoma-B aggravates atherosclerosis by inducing CD8+ T cell-mediated macrophage death. *Eur Heart J* 2019, 40(4):372-382.
33. Winter C, Silvestre-Roig C, Ortega-Gomez A, Lemnitzer P, Poelman H, Schumski A, Winter J, Drechsler M, de Jong R, Immler R et al: Chrono-pharmacological Targeting of the CCL2-CCR2 Axis Ameliorates Atherosclerosis. *Cell Metab* 2018, 28(1):175-182.e175.
34. Miyabe C, Miyabe Y, Bricio-Moreno L, Lian J, Rahimi RA, Miura NN, Ohno N, Iwakura Y, Kawakami T, Luster AD: Dectin-2-induced CCL2 production in tissue-resident macrophages ignites cardiac arteritis. *J Clin Invest* 2019, 129(9):3610-3624.
35. Adlam D, Olson TM, Combaret N, Kovacic JC, Iismaa SE, Al-Hussaini A, O'Byrne MM, Bouajila S, Georges A, Mishra K et al: Association of the PHACTR1/EDN1 Genetic Locus With Spontaneous Coronary Artery Dissection. *J Am Coll Cardiol* 2019, 73(1):58-66.
36. Liu S, Jiang Z, Qiao L, Guo B, Xiao W, Zhang X, Chang L, Li Y: Integrin  $\beta$ -3 is required for the attachment, retention and therapeutic benefits of human cardiospheres in myocardial infarction. *J Cell Mol Med* 2018, 22(1):382-389.
37. Pierpont EI, Semrud-Clikeman M, Pierpont ME: Variability in clinical and neuropsychological features of individuals with MAP2K1 mutations. *Am J Med Genet A* 2017, 173(2):452-459.
38. Davies B, d'Udekem Y, Ukoumunne OC, Algar EM, Newgreen DF, Brizard CP: Differences in extracellular matrix and myocyte homeostasis between the neonatal right ventricle in hypoplastic left heart syndrome and truncus arteriosus. *Eur J Cardiothorac Surg* 2008, 34(4):738-744.
39. Lowery JW, Rosen V: The BMP Pathway and Its Inhibitors in the Skeleton. *Physiol Rev* 2018, 98(4):2431-2452.
40. Jiang Y, Habibollah S, Tilgner K, Collin J, Barta T, Al-Aama JY, Tesarov L, Hussain R, Trafford AW, Kirkwood G et al: An induced pluripotent stem cell model of hypoplastic left heart syndrome (HLHS) reveals multiple expression and functional differences in HLHS-derived cardiac myocytes. *Stem Cells Transl Med* 2014, 3(4):416-423.
41. Gambetta K, Al-Ahdab MK, Ilbawi MN, Hassaniya N, Gupta M: Transcription repression and blocks in cell cycle progression in hypoplastic left heart syndrome. *Am J Physiol Heart Circ Physiol* 2008,

## Figures



**Figure 1**

The workflow of this study: data acquisition, preprocessing, analysis and verification.



**Figure 2**

Identification of DEGs and pathways for LV and RV HLHS.



