Journal of Cardiology 66 (2015) 527–531

Contents lists available at ScienceDirect

Journal of Cardiology

journal homepage: www.elsevier.com/locate/jjcc

Original article

# The expression profile analysis of *NKX2-5* knock-out embryonic mice to explore the pathogenesis of congenital heart disease

Jian Li (MD)[a,b,1], Yinyin Cao (MD)[a,1], Yao Wu (MD)[a], Weicheng Chen (MD)[a], Yuan Yuan (MD)[a], Xiaojing Ma (MD)[a,b,**], Guoying Huang (MD)[a,b,*]

[a] *Pediatric Heart Center, Children's Hospital of Fudan University, Shanghai 201102, China*
[b] *Shanghai Key Laboratory of Birth Defect, Shanghai 201102, China*

ARTICLE INFO

*Article history:*
Received 29 August 2014
Received in revised form 9 December 2014
Accepted 18 December 2014
Available online 24 March 2015

*Keywords:*
Congenital heart disease
Embryonic mice
Differentially expressed genes
Regulation factors
Latent pathways

ABSTRACT

*Background:* Mutation of *NKX2-5* could lead to the development of congenital heart disease (CHD) which is a common inherited disease. This study aimed to investigate the pathogenesis of CHD in *NKX2-5* knock-out embryonic mice.

*Methods:* The expression profile in the *NKX2-5* knock-out embryonic mice (GSE528) was downloaded from Gene Expression Omnibus. The heart tissues from the null/heterozygous embryonic day 12.5 mice were compared with wild-type mice to identify differentially expressed genes (DEGs), and then DEGs corresponding to the transcriptional factors were filtered out based on the information in the TRANSFAC database. In addition, a transcriptional regulatory network was constructed according to transcription factor binding site information from the University of California Santa Cruz database. A pathway interaction network was constructed by latent pathways identification analysis.

*Results:* The 42 DEGs corresponding to transcriptional factors from the null and heterozygous embryos were identified. The transcriptional regulatory networks included five down-regulated DEGs (*SP1*, *SRY*, *JUND*, *STAT6*, and *GATA6*), and six up-regulated DEGs [*POU2F1*, *NFY* (*NFYA*/*NFYB*/*NFYC*), *USF2* and *MAX*]. Latent pathways analysis demonstrated that ribosome, glycolysis/gluconeogenesis, and dilated cardiomyopathy pathways significantly interacted.

*Conclusion:* The identified DEGs and latent pathways could provide new comprehensive view for understanding the pathogenesis of CHD.



## Introduction

Cardiovascular disease is the leading cause of mortality worldwide [1]. Congenital heart defect (CHD) is defined as the defective structure of heart and great vessels and the incidence of CHD could reach up to 1% in newborns [2,3]. The developmental process of the heart is complicated, requiring accurate expression of related genes over time and space, among which the transcriptional factors play key roles. Generally speaking, paired regions of anterior lateral mesoderm generate the cardiac lineages and parallelly cardiac primordia develop, which form the primitive heart tube along the ventral midline of the embryo [4,5].

Cardiac development originates from a series of rhythmical gene expressions which are concisely controlled by the transcription factors. Most of the known causes of CHD are sporadic genetic changes, for example, mutations of α-myosin heavy chain (*MYH6*) are associated with cardiac defects [6]. Many research results have indicated that transcriptional factors, such as *NKX2-5*, *GATA*, and *TBX*, play vital roles in the occurrence and development of CHD [7–10]. Among the *NK2* family, *NKX2-3*, *NKX2-5*, *NKX2-7*, *NKX2-8*, and *NKX2-10* are related to the development of the heart and *NKX2-5* has attracted more attention [11,12]. Cardiac homeobox gene *NKX2-5* (also called *Csx1*) has been confirmed to be closely related to CHD [13]. Unknown mechanisms could be responsible for why human mutations in *NKX2-5* lead to progressive cardiomyopathy and conduction defects. Previously, *NKX2-5* was identified in zebrafish and mice [14,15]. However, Gioli-Pereira et al. [16] evaluated the presence of mutations in the *NKX2-5* in 159 patients

* Corresponding author at: Pediatric Heart Center, Children's Hospital of Fudan University, 399 Wan Yuan Road, Shanghai 201102, China. Tel.: +86 021 64931990; fax: +86 021 64931990.
** Corresponding author at: Pediatric Heart Center, Children's Hospital of Fudan University, 399 Wan Yuan Road, Shanghai 201102, China. Tel.: +86 021 64931817; fax: +86 021 64931990.
*E-mail addresses:* xiaojingma021@163.com (X. Ma), gyhuang@shmu.edu.cn (G. Huang).
[1] These authors contributed equally to this work.

http://dx.doi.org/10.1016/j.jjcc.2014.12.022
0914-5087/

with a diverse range of non-syndromic congenital heart diseases, and found that there was no absolute association between *NKX2-5* and CHD.

In order to explore the pathogenesis of CHD, we applied bioinformatics methods to analyze the expression profiling of GSE528. Schinke et al. created a *NKX2-5* knock-out mouse by Cre/loxP mediated excision. Homozygous mutant mice die at embryonic day 14.5 because of severe cardiac malformations, while heterozygous mutant mice demonstrate the cardiac defects associated with cardiac septation and valve morphogenesis (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE528). The expressional profile of wild embryos, *NKX2-5* knock-out null embryonic mice and heterozygous embryonic mice, were analyzed to select the differentially expressed genes (DEGs). Then relevant DEGs corresponding to transcription factors were filtered out based on the information in the TRANSFAC database [17]. Furthermore latent pathway identification analysis was applied to identify the interactions between the pathways. Finally, transcriptional regulatory network and latent pathway interaction networks were constructed.

## Materials and methods

### *Data sources*

The gene expressional profile of GSE528 from mice was downloaded from Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE528). The platforms were GPL75 and GPL76 [Mu11KsubA] Affymetrix Murine 11K SubA Arrays. Ten heart tissue samples were applied in this study, including four samples from wild-type embryonic day 12.5 mice, three from *NKX2-5* knock-out heterozygous embryonic day 12.5 mice, and three from *NKX2-5* knock-out null embryonic day 12.5 mice.

### *Preprocessing of expression data*

Preprocessing of the expression profile matrix was first performed. If there was at least one missing expression value for one probe, the probe was discarded. According to the annotation information from platform, probe name was converted to a gene name, and expression values of multiple probes corresponding to the same gene were averaged.

### *Identification of DEGs*

In order to get the DEGs, the heart tissue samples of wild embryos, *NKX2-5* knock-out null embryonic mice, and *NKX2-5* knock-out heterozygous embryonic mice were performed in a pairwise comparison with Limma package of R language [18], and screening criteria were adjusted with a $p$-value $<0.05$, and $|\log_2 \mathrm{FC}| > 0.585$.

### *Screening of the transcription factors in DEGs*

DEGs corresponding to transcription factors were obtained according to information from TRANSFAC database [17] and target genes of transcription factors according to binding sites information were provided by the University of California Santa Cruz (UCSC) database (http://genome.ucsc.edu/).

### *Latent pathways identification analysis*

Latent pathways identification analysis [19] was applied to identify the interactions between the pathways in which DEGs were involved. The stronger the significance of interaction was, the more relevant was the relationship between the pathways and diseases. Firstly, the clusterProfiler package of R language was applied to identify all gene ontology (GO) terms biological process (BP) (later referred to as $G$) and KEGG pathways (referred to as $P$) in which DEGs were involved. The expressed values of DEGs were obtained according to microarray expression profile. Secondly, the bipartite network of $G$ and $P$ was built. One node of the line was $P$ and the other side was $G$. The edges represented DEGs that were both involved in the $G$ and $P$. The weight of edge was determined by two parts: (1) the relative overlap of $P$ and $G$, presented as Jaccard similarity coefficient [20]; (2) the expression values of DEGs were averaged. The weight formula was as follows:

$$w_{GP} = \left|\frac{G \cap P}{G \cup P}\right| \times med\{DE_x : x \in G \cap P\}$$

wherein $P$ represents KEGG pathway; $G$ represents GO BP term; $|(G \cap P)/(G \cup P)|$ represents Jaccard similarity coefficient of $P$ and $G$; and $DE$ represents expression values of DEGs. $G \cup P$ represents all genes involved in pathway $P$ and $G$; $G \cap P$ represents the genes involved in $G$ and $P$. Thirdly, based on bipartite pathway network, if there is at least one common BP term between the two pathways, they will be connected. The weight of edge was calculated according to $A_{ij} = \sum_{k=1}^{G} w_{G_k P_i} \times w_{G_k P_j}$. Fourthly, the interaction of pathways in the network was calculated by random walk method. The criterion for the significant interaction pathways was $p$-value $<0.05$.

### *Transcriptional regulatory networks and potential pathways interaction networks*

The transcriptional regulatory networks were constructed and the latent pathway interaction networks were built according to the weight of pathway interactions using Cytoscape software (Cytoscape 3.X; http://www.cytoscape.org) [21].

## Results

### *Expression matrix normalization and screening of DEGs and transcription factors*

The expression profile matrixs before and after normalization were shown in Fig. 1. The normalized median line after normalization was almost in a straight line, indicating that normalization was successful.

The genes with $|\log_2 \mathrm{FC}| > 0.585$ and $p$-value $<0.05$ were selected as significant DEGs. Information of transcription factors was obtained from TRANSFAC database. Table 1 shows the number of up-regulated DEGs, down-regulated DEGs, and their corresponding transcription factors in heterozygous and null embryonic mice compared with wild embryonic mice.

### *Latent pathways analysis*

To explore the interactions between the pathways affected by DEGs in *NKX2-5* knock-out embryonic mice, the associated network between the pathways was constructed and a total of 8361 interactions involving 130 pathways in *NKX2-5* knock-out heterozygous embryos were selected. A total of 8385 interactions involving 130 pathways in *NKX2-5* knock-out null embryos were screened.

Figs. 2 and 3 showed the interaction network of latent pathways, demonstrating that there were 13 significant pathways in heart tissue from heterozygous embryos and null embryos. The significance of pathways in the interaction network was evaluated. Thirteen pathways with $p$-value $<0.05$ were identified in heterozygous embryos (Table 2) and thirteen pathways with $p$-value $<0.05$ were identified in null embryonic heart tissue

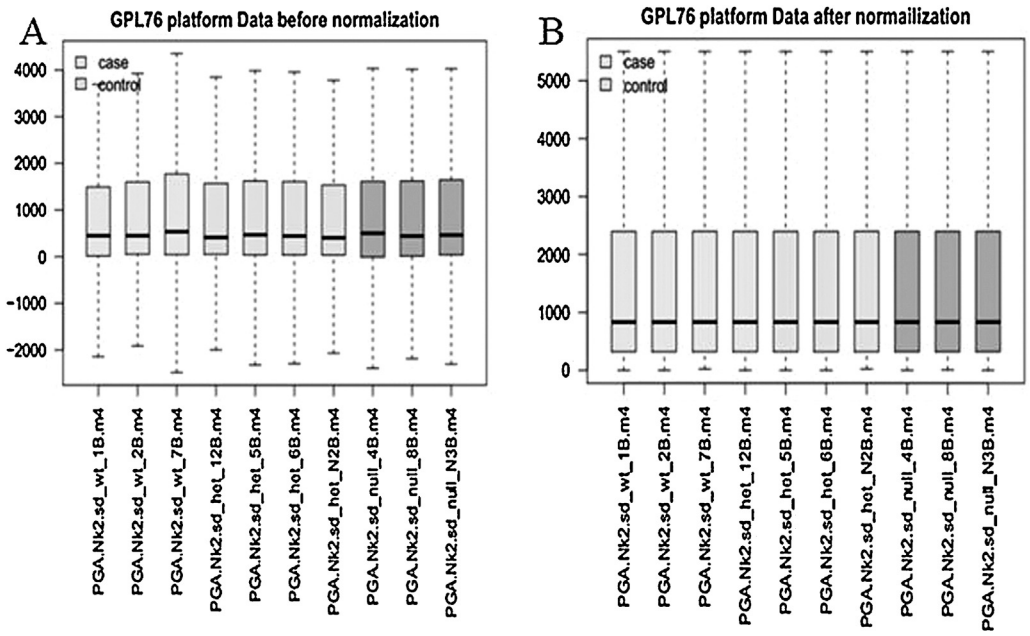


**Fig. 1.** The expression cassette of samples before normalization (A) and after normalization (B). The horizontal axis indicates the name of the samples, and the ordinate axis represents the expression values. The black lines in each cassette are the median of data.

**Table 1**
The number of differentially expressed genes (DEGs) and their corresponding transcription factors (TF).

| ID | Up-DEG/TF | Down-DEG/TF |
|---|---|---|
| Heterozygous vs. wild embryo | 970/46 | 1063/46 |
| Null vs. wild embryo | 983/45 | 1052/45 |
| Overlap | 391/21 | 403/21 |

(Table 3). The pathways included ribosome, glycolysis or gluconeogenesis, pyruvate metabolism, cardiac muscle contraction, porphyrin and chlorophyll metabolism, fructose and mannose metabolism, hypertrophic cardiomyopathy, and dilated cardiomyopathy pathways.

*Construction of transcriptional regulatory networks*

The 42 common up-regulated and down-regulated transcription factors in heterozygous and null embryonic mice were identified, and transcriptional regulatory networks were constructed according to the transcription factor binding sites information provided by the UCSC database. As can be seen from Fig. 4, 11 DEGs corresponding to transcription factors and regulatory target genes formed the umbrella regulatory networks. *SP1* (specificity protein 1), *SRY* (sex-determining region Y, testis

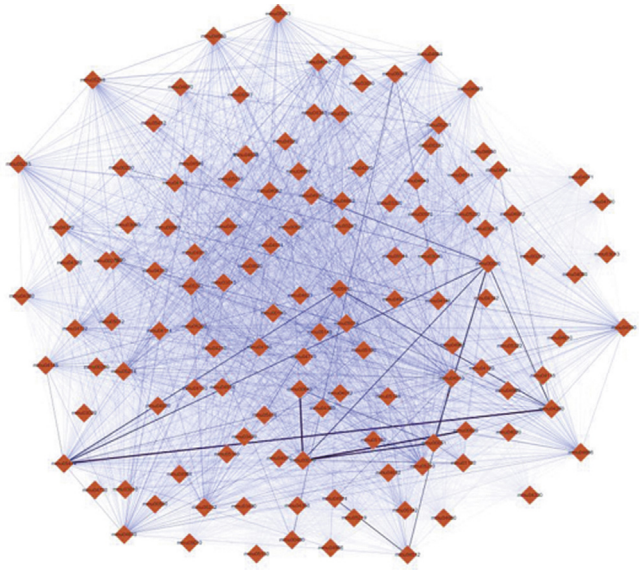

**Fig. 2.** The interaction network of pathways that were involved in the differentially expressed genes (DEGs) in *NKX2-5* knock-out heterozygous embryonic day 12.5 mice. Red diamonds represent pathways and the edges represent interactions between pathways. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

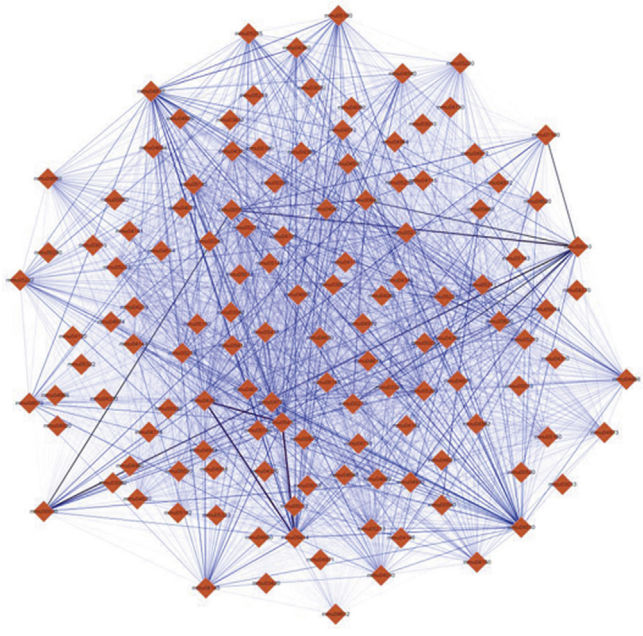

**Fig. 3.** The interaction network of pathways that were involved in the differentially expressed genes (DEGs) in *NKX2-5*-knocked null embryonic day 12.5 mice. Red diamonds represent pathways and the edges represent interactions between pathways. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

**Table 2**
The significant pathways that differentially expressed genes in *NKX2-5* knock-out heterozygous embryonic day 12.5 mice involved by latent pathways analysis.

| KEGG pathway ID | Description | *p*-Value | Degree |
|---|---|---|---|
| mmu03010 | Ribosome | 0.000 | 129 |
| mmu00010 | Glycolysis/gluconeogenesis | 0.000 | 129 |
| mmu00620 | Pyruvate metabolism | 0.004 | 128 |
| mmu04260 | Cardiac muscle contraction | 0.006 | 128 |
| mmu00860 | Porphyrin and chlorophyll metabolism | 0.010 | 129 |
| mmu00051 | Fructose and mannose metabolism | 0.020 | 129 |
| mmu05414 | Dilated cardiomyopathy | 0.024 | 129 |
| mmu05410 | Hypertrophic cardiomyopathy (HCM) | 0.026 | 128 |
| mmu00270 | Cysteine and methionine metabolism | 0.026 | 129 |
| mmu04146 | Peroxisome | 0.028 | 129 |
| mmu00640 | Propanoate metabolism | 0.040 | 128 |
| mmu00480 | Glutathione metabolism | 0.042 | 129 |
| mmu03040 | Spliceosome | 0.042 | 129 |

**Table 3**
The significant pathways that differentially expressed genes in *NKX2-5* knock-out null embryonic day 12.5 mice involved by latent pathways analysis.

| KEGG pathway ID | Description | *p*-Value | Degree |
|---|---|---|---|
| mmu03010 | Ribosome | 0.000 | 129 |
| mmu00010 | Glycolysis/gluconeogenesis | 0.000 | 129 |
| mmu04260 | Cardiac muscle contraction | 0.002 | 129 |
| mmu00620 | Pyruvate metabolism | 0.002 | 129 |
| mmu00051 | Fructose and mannose metabolism | 0.014 | 129 |
| mmu00860 | Porphyrin and chlorophyll metabolism | 0.022 | 129 |
| mmu05410 | Hypertrophic cardiomyopathy (HCM) | 0.024 | 129 |
| mmu00270 | Cysteine and methionine metabolism | 0.024 | 129 |
| mmu00640 | Propanoate metabolism | 0.028 | 129 |
| mmu05414 | Dilated cardiomyopathy | 0.032 | 129 |
| mmu03040 | Spliceosome | 0.032 | 129 |
| mmu00030 | Pentose phosphate pathway | 0.040 | 129 |
| mmu00480 | Glutathione metabolism | 0.049 | 129 |

determining factor), *JUND*, *STAT6*, and *GATA6* were down-regulated and *POU2F1* (POU class 2 homeobox 1 gene), *NFY* (nuclear factor Y) (*NFYA*/*NFYB*/*NFYC*), *USF2* (upstream stimulatory factor 2), and *MAX* were up-regulated.

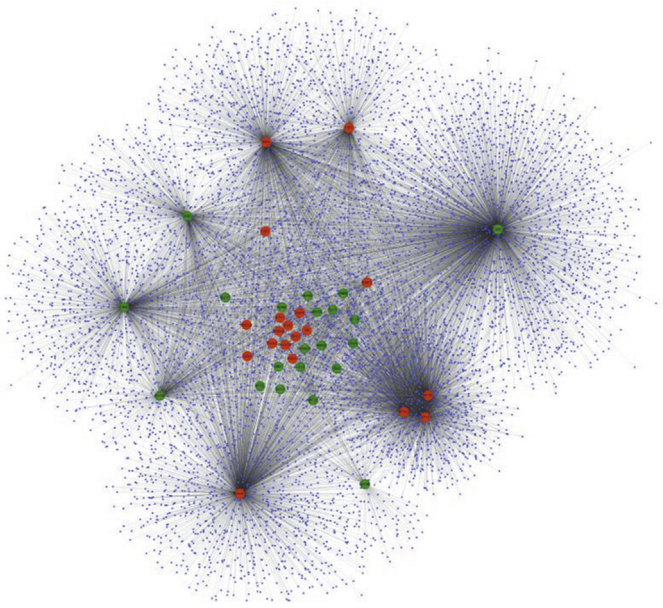

**Fig. 4.** Regulatory network of transcription factors with the same trend and their target genes in *NKX2-5* knock-out heterozygous and null embryonic mice. Red nodes represent up-regulated genes corresponding to transcription factors; green nodes represent down-regulated genes corresponding to transcription factors; purple nodes represent regulatory target genes. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

## Discussion

CHD includes congenital malformations caused by fetal abnormalities in heart and great vessels during childhood [22]. In our study, a total of 42 common up-regulated and down-regulated DEGs corresponding to transcription factors were identified in *NKX2-5* knock-out heterozygous and null embryonic mice compared with wild type.

In the latent pathway interactions network, significant pathways included ribosome, glycolysis or gluconeogenesis, pyruvate metabolism, cardiac muscle contraction, porphyrin and chlorophyll metabolism, fructose and mannose metabolism, hypertrophic cardiomyopathy, and dilated cardiomyopathy pathways. It has been reported that the synthesis of ribosome was accelerated in cardiac hypertrophy in the rat heart [23]. DEGs were involved in glycolysis or gluconeogenesis, pyruvate metabolism, and fructose and mannose metabolism, indicating that the energy metabolism was changed in mice after the NKX2-5 knock-out. Dilated cardiomyopathy and hypertrophic cardiomyopathy were also the significant pathways in the pathways interaction network. Thus, the expression genes involved in the dilated and hypertrophic cardiomyopathy may be related to CHD.

In the regulatory network of transcription factors and target genes, *SP1*, *SRY*, *JUND*, *STAT6*, and *GATA6* were down-regulated, and *POU2F1*, *NFY* (*NFYA*/*NFYB*/*NFYC*), *USF2*, and *MAX* were up-regulated. Lack of *JUND* could promote dilated cardiomyopathy in fra-1 transgenic mice [24]. *GATA6*, encoding a cardiac transcription factor, is extensively expressed in developing heart and non-synonymous *GATA6* variants (A178V and L198V) were not found in control population but identified in two individuals with CHD [25]. But there were no reports about the association of other transcription factors with CHD. Cyclooxygenase-2 (COX-2) is expressed in the heart failure and overexpression of *SP1* that is a transactivation factor could enhance the cox-2 promoter activity [26]. *STAT6* is activated when ischemic injury of heart is induced [27]. Although there was no direct evidence, their association with heart disease or heart development may imply their potential roles in CHD. No reports were found about the relation between *SRY*, *POU2F1*, *NFY* (*NFYA*/*NFYB*/*NFYC*), *USF2*, MAX, and heart disease or heart development. However, we could speculate that the transcription factors identified in our study may be related to CHD, providing potential view for the pathogenesis of *NKX2-5* knock-out CHD.

In conclusion, latent pathways analysis indicated that knock-out of *NKX2-5* in embryonic mice could impact ribosome pathway, energy metabolism process, myocardium development, and providing a new and comprehensive view for pathogenesis of CHD. The transcription factors *SP1*, *JUND*, *STAT6*, and *GATA6* may be related to CHD. However, the function of genes and pathways identified in our study needs large experiments to investigate their roles in the *NKX2-5* knock-out CHD.

## Authors' contributions

Guoying Huang and Xiaojing Ma participated in the design of this study, and they both performed the statistical analysis. Jian Li and Yinyin Cao carried out the study, together with Yuan Yuan, and collected important background information, and also drafted the manuscript. Yao Wu and Weicheng Chen conceived this study, and participated in the design and helped to draft the manuscript. All authors read and approved the final manuscript.

## Funding

This study was supported by the Basic Research Projects of China (2010CB529504), the Natural Science Foundation of China

(81370198), Youth Programs of National Natural Science Fund (81400240), Shanghai Municipal Commission of Science and Technology Research Project (11JC1401400), and Shanghai Health and Family Planning Commission Scientific Research Project (201440632).

## Conflict of interest

The authors declare that there is no conflict of interest.

## References

[1] Grimaldi V, Mancini FP, Casamassimi A, Al-Omran M, Zullo A, Infante T, Napoli C. Potential benefits of cell therapy in coronary heart disease. J Cardiol 2013;62:267–76.
[2] Kohut R, Rusen ID. Congenital anomalies in Canada: a perinatal health report, 2002. Ottawa: Public Health Agency of Canada; 2002.
[3] Hoffman JI, Kaplan S. The incidence of congenital heart disease. J Am Coll Cardiol 2002;39:1890–900.
[4] Olson EN, Srivastava D. Molecular pathways controlling heart development. Science 1996;272:671–6.
[5] Fishman MC, Chien KR. Fashioning the vertebrate heart: earliest embryonic decisions. Development 1997;124:2099–117.
[6] Carniel E, Taylor MR, Sinagra G, Di Lenarda A, Ku L, Fain PR, Boucek MM, Cavanaugh J, Miocic S, Slavov D. α-Myosin heavy chain a sarcomeric gene associated with dilated and hypertrophic phenotypes of cardiomyopathy. Circulation 2005;112:54–9.
[7] Kasahara H, Lee B, Schott J-J, Benson DW, Seidman J, Seidman CE, Izumo S. Loss of function and inhibitory effects of human CSX/NKX2. 5 homeoprotein mutations associated with congenital heart disease. J Clin Invest 2000;106: 299–308.
[8] Zhang W, Li X, Shen A, Jiao W, Guan X, Li Z. GATA 4 mutations in 486 Chinese patients with congenital heart disease. Eur J Med Genet 2008;51:527–35.
[9] Gelb BD. Genetic basis of congenital heart disease. Curr Opin Cardiol 2004;19: 110–5.
[10] Schott J-J, Benson DW, Basson CT, Pease W, Silberbach GM, Moak JP, Maron BJ, Seidman CE, Seidman JG. Congenital heart disease caused by mutations in the transcription factor NKX2-5. Science 1998;281:108–11.
[11] Holland ND, Venkatesh TV, Holland LZ, Jacobs DK, Bodmer R. Amphink2-tin, an amphioxus homeobox gene expressed in myocardial progenitors: insights into evolution of the vertebrate heart. Dev Biol 2003;255:128–37.
[12] Newman CS, Reecy J, Grow MW, Ni K, Boettger T, Kessel M, Schwartz RJ, Krieg PA. Transient cardiac expression of the tinman-family homeobox gene, XNkx2-10. Mech Dev 2000;91:369–73.
[13] Moskowitz IP, Kim JB, Moore ML, Wolf CM, Peterson MA, Shendure J, Nobrega MA, Yokota Y, Berul C, Izumo S. A molecular pathway including Id2, Tbx5, and Nkx2-5 required for cardiac conduction system development. Cell 2007;129: 1365–76.
[14] Chen J-N, Fishman MC. Zebrafish tinman homolog demarcates the heart field and initiates myocardial differentiation. Development 1996;122:3809–16.
[15] Lints TJ, Parsons LM, Hartley L, Lyons I, Harvey RP. Nkx-2.5: a novel murine homeobox gene expressed in early heart progenitor cells and their myogenic descendants. Development 1993;119:419–31.
[16] Gioli-Pereira L, Pereira AC, Mesquita SM, Xavier-Neto J, Lopes AA, Krieger JE. NKX2. 5 mutations in patients with non-syndromic congenital heart disease. Int J Cardiol 2010;138:261–5.
[17] Matys V, Fricke E, Geffers R, Gößling E, Haubrock M, Hehl R, Hornischer K, Karas D, Kel AE, Kel-Margoulis OV. TRANSFAC®: transcriptional regulation, from patterns to profiles. Nucleic Acids Res 2003;31:374–8.
[18] Sanges R, Cordero F, Calogero RA. oneChannelGUI: a graphical interface to Bioconductor tools, designed for life scientists who are not familiar with R language. Bioinformatics 2007;23:3406–8.
[19] Pham L, Christadore L, Schaus S, Kolaczyk ED. Network-based prediction for sources of transcriptional dysregulation using latent pathway identification analysis. Proc Natl Acad Sci U S A 2011;108:13347–52.
[20] Real R, Vargas JM. The probabilistic basis of Jaccard's index of similarity. Syst Biol 1996;45:380–5.
[21] Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003;13: 2498–504.
[22] Allen HD, Driscoll DJ, Shaddy RE, Feltes TF. Moss & Adams' heart disease in infants, children, and adolescents: including the fetus and young adult. Wolters Kluwer Health; 2013.
[23] Siehl D, Chua BH, Lautensack-Belser N, Morgan HE. Faster protein and ribosome synthesis in thyroxine-induced hypertrophy of rat heart. Am J Physiol 1985;248:C309–19.
[24] Ricci R, Eriksson U, Oudit GY, Eferl R, Akhmedov A, Sumara I, Sumara G, Kassiri Z, David J-P, Bakiri L. Distinct functions of junD in cardiac hypertrophy and heart failure. Genes Dev 2005;19:208–13.
[25] Maitra M, Koenig SN, Srivastava D, Garg V. Identification of GATA6 sequence variants in patients with congenital heart defects. Pediatr Res 2010;68:281–5.
[26] Xu Q, Ji Y-S, Schmedtje JF. Sp1 increases expression of cyclooxygenase-2 in hypoxic vascular endothelium. Implications for the mechanisms of aortic aneurysm and heart failure. J Biol Chem 2000;275:24583–89.
[27] Mascareno E, Siddiqui M. The role of Jak/STAT signaling in heart tissue renin–angiotensin system. Mol Cell Biochem 2000;212:171–5.