Cellular Physiology and Biochemistry

Cell Physiol Biochem 2018;48:2134-2146
DOI: 10.1159/000492556
Published online: August 9, 2018
Accepted: July 31, 2018

Original Paper

Identification of Competing Endogenous RNA Regulatory Networks in Vitamin A Deficiency-Induced Congenital Scoliosis by Transcriptome Sequencing Analysis

Chong Chen[a] Haining Tan[a] Jiaqi Bi[a] Zheng Li[a] Tianhua Rong[a]
Youxi Lin[a] Liang Sun[b] Xingye Li[a] Jianxiong Shen[a]

[a]Department of Orthopedic Surgery, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences, Peking Union Medical College, Beijing, [b]Beijing Zhongke Jingyun Technology Company Ltd., Beijing, China

Key Words
Congenital scoliosis (CS) • Competing endogenous RNA (ceRNA) • Sequencing • Vitamin A deficiency • Bioinformatics analysis

Abstract
Background/Aims: Congenital scoliosis (CS) is a result of anomalous development of vertebrae and is frequently associated with somitogenesis malformation. Although noncoding RNAs (ncRNAs) have been recently determined to be involved in the pathogenesis of CS, the competing endogenous RNA (ceRNA) regulatory networks in CS remain largely unknown. ***Methods:*** Sequencing was conducted to explore the ncRNA expression profiles in rat embryos (gestation day 9) following vitamin A deficiency (VAD) (n = 9 for the vitamin A deficiency-induced congenital scoliosis (VAD-CS) group and n = 4 for the control group). Real-time reverse transcriptase polymerase chain reaction (RT-PCR) was conducted to verify the expression levels of selected mRNAs, long noncoding RNAs (lncRNAs), circular RNAs (circRNAs), and microRNAs (miRNAs). Bioinformatics analysis was used to discover the possible relationships and functions of the ceRNAs. ***Results:*** A total of 749 mRNAs, 56 miRNAs, 685 lncRNAs, and 70 circRNAs were identified to have significantly different expression levels in the two groups. Wnt, PI3K-ATK, FoxO, EGFR, and mTOR were found to be the most significant pathways involved in VAD-CS pathogenesis. The circRNA/miRNA/mRNA and lncRNA/miRNA/mRNA networks of CS were built, and the gene expression mechanisms regulated by ncRNAs were unveiled via the ceRNA regulatory networks. ***Conclusion:*** We comprehensively identified ceRNA regulatory networks of embryonic somite development in VAD-CS as well as revealed the contribution of different ncRNA expression profiles. Our data demonstrate the association between mRNAs and ncRNAs in the pathogenic mechanism of CS.

Jianxiong Shen

Department of Orthopedic Surgery, Peking Union Medical College Hospital
No. 1 Shuaifuyuan, Wangfujing Street, Beijing 100730 (China)
Tel. +86-10-69152813, E-Mail sjxpumch@163.com

Introduction

Congenital scoliosis (CS) is a type of segmental or mixed vertebral malformation that manifests as beyond 10 degrees of lateral curvature of the spine, which is caused by defects in vertebral formation during embryogenesis and is characterized by a longitudinal and rotational imbalance [1-5]. CS has an estimated prevalence of approximately 0.5–1‰ of live births [4, 6]. However, the etiology and molecular mechanisms of CS remain unclear and multifactorial, including both genetic and environmental factors.

Previous evidence has identified that, in mouse embryos, disturbance of the retinoic acid pathway might damage the left-right bilateral symmetry, since the retinoic acid pathway controls the segmental structure of the vertebrate body plan in embryogenesis and is important in segmentation clock development [7-9]. Recently, it has been reported that CS, which is caused by vitamin A deficiency (VAD) during pregnancy, in postnatal rats may be induced by a retinoic acid pathway abnormality in somitogenesis [10].

Noncoding RNAs (ncRNAs) represent a class of RNAs that are defined as nonprotein-encoding transcripts that functionally regulate protein expression. The ncRNAs include microRNAs (miRNAs) (shorter than 200 nucleotides), long ncRNAs (lncRNAs, longer than 200 nucleotides), and circular RNAs (circRNA, closed loop structure) [11].

The activities of intracellular lncRNAs or circRNAs plus competing endogenous RNAs (ceRNAs) might function as miRNA sponges, which will inhibit miRNAs by promoting binding with miRNA recognition elements and powerfully inhibiting miRNA activity, thus causing increased levels of miRNA target genes [12].

To the best of our knowledge, the ceRNA regulatory networks in CS and the fundamental mechanisms have not been thoroughly investigated. In this study, we predicted and investigated embryonic ceRNAs in VAD-induced CS (VAD-CS) by RNA sequencing. In addition, comprehensive prediction and investigation of the ceRNA mechanisms during the somitogenesis of CS were performed to offer novel diagnostic and therapeutic strategies against CS.

Materials and Methods

Animals

Twenty female Sprague Dawley rats (20 weeks old, body weight of 200–230 g) were obtained from the Laboratory Animal Center of Army General Hospital and SPF Biotechnology Co., Ltd. (Protocol No. SYXK 2014-0037, Beijing, China) and randomly allocated to standard cages, with five rats in each cage. The institutional animal welfare committee approved all of the animal procedures, which were in line with the national and international animal care and ethics guidelines. All rats were housed at room temperature (21–23 °C) with a 60–70% relative humidity and 12-h light/dark cycle, and had free access to standard laboratory water and diet.

CS model establishment

The rat CS model induced by VAD was established as described previously [10]. In brief, 20 female rats were allocated randomly into the VAD-CS group (n = 12) and the control group (n = 8). After 1 week of adaptation, rats in the VAD-CS group received a modified diet of AIN-93G deprived of any vitamin A source (Research Diets, USA); rats in the control group received an AIN-93G diet with adequate vitamin A (4 retinol equivalents (RE)/g diet), for more than 2 weeks. VAD (less than 2 mg of vitamin A/100 mL, 0.74 μM) was confirmed by analyzing the plasma levels using high-performance liquid chromatograpHhy in a random sample, as described previously [10, 13]. Next, the female rats receiving the designated diet described above were mated with normal male rats at 6–10 pm. During gestation, the female rats continuously received the same diet as mentioned above [14].

Tissue collection

On gestation day (GD) 9, nine VAD-CS group rats and five control group rats were detected to be pregnant. A total of 79 embryos from the VAD-CS group and 43 embryos from the control group were collected. Nine embryos from the VAD-CS group and four embryos from the control group were sequenced.

RNA isolation, library construction, and sequencing

RNA was purified from the embryos collected above by TRIzol, according to the manufacturer's methods. An RNA Nano 6000 Assay Kit and Bioanalyzer 2100 system (Agilent Technologies, CA, USA) were applied to assess the RNA integrity. A Ribo-Zero Gold rRNA Removal Kit (Human/Mouse/Rat) (Epicenter Company, USA) was used to remove the ribosomal RNAs. Duplicate samples of RNA from the same embryo were subjected to RNase R (Cat. No. RNR07250, Epicentre Company, USA) or not. RNase R was used to digest the linear RNA to separate the circRNAs. Then, reverse transcriptase was further added to the two samples of RNAs with or without RNase R for cDNA construction. The RNAs treated with RNase were subjected to circRNA detection, while the RNAs without RNase R treatment were used to detect mRNAs and lncRNAs. Both of the cDNA libraries were sequenced by Illumina XTEN, paired-end 150 bp.

Meanwhile, the NEBNext Multiplex Small RNA Library Prep Set for Illumina (NEB, USA) was used to generate the sequencing libraries. The miRNA libraries were sequenced by BGISEQ-500, single-end 50 nt.

Transcriptome assembly

First, the RNA sequencing library quality was evaluated by FastQC (v0.11.5), and then the low-quality reads were removed using cutadapt. The clean RNA sequencing reads were aligned to the reference genome of rats, downloaded from UCSC, using Tophat (v2.1.0). Then, the transcriptome was re-assembled by Cufflinks (v2.2.1), the rat reference annotations from GENCODE. Next, Cuffmerge was used to merge the meta-assemblies from all samples to build the final transcriptome.

Expression analysis

Gene expression (mRNA and lncRNA) was estimated with normalized fragments per kilobase per million mapped reads by cufflinks. Differential gene expression analysis was performed using cuffdiff with default parameters. Differentially expressed (DE) genes with a fold change > 2 and a p-value < 0.05 were considered for additional investigation. For miRNA abundance analysis, miRDeep2 was used to estimate the miRNA abundance, and different levels of miRNA expression were analyzed by DESeq2 (R software package). For circRNA abundance analysis, we used the normalized reads count as the abundance. DE circRNAs were estimated by DESeq2 (R software package).

Target gene prediction and function enrichment analysis

MiRNA target genes were predicted by the overlapping results of TargetScan and miRanda, and the network was based on the target prediction.

Annotation of gene ontology (GO) and analysis of Kyoto Encyclopedia of Genes and Genomes *(KEGG) signaling pathways*

Annotation of GO and analysis of KEGG signaling pathways were performed to explore the functions of all of the DE genes. In brief, for the DE genes, GO analysis (http://www.geneontology.org) was used to clarify the interesting genetic regulatory networks by creating hierarchical categories based on the molecular function (MF), cellular component (CC), and biological process (BP) features. In addition, pathway analysis was conducted to discover the important pathways based on the KEGG (http://www.genome.jp/kegg/).

Analysis of the ceRNA regulatory network

We modified a published method to construct a co-expression network of DE genes. For each gene pair (between mRNAs and lncRNAs, between mRNAs and circRNAs), the Pearson correlation coefficient and the corresponding *P*-value by the R software package (WGCNA) were calculated. In advance, we analyzed the mRNA-miRNA-ceRNA (lncRNA, circRNA) regulatory network, which was based on the overlapping miRNA target between the mRNAs and ceRNAs.

Real-time reverse transcriptase polymerase chain reaction (RT-PCR)

The relative mRNA expression levels of retinaldehyde dehydrogenase (RALDH) 1, 2, and 3, as well as retinoic acid receptor (RAR)-α, β, and γ were evaluated by RT-PCR using an iQ5 RT-PCR detection machine (Bio-Rad, Hercules, CA, USA) for the remaining three rats of each group to validate the CS model. Following the manufacturer's instructions, the All-in-one first-strand cDNA synthesis kit (Genecopoeia) was used to synthesize the first-strand cDNA, which was further amplified in triplicate using the IQ SYBR green SuperMix reagent (Bio-Rad, Hercules, CA, USA) with an Opticon RT-PCR instrument (MJ Research, Waltham, MA, USA). RT-PCR specificity was verified according to the melting-curve assay and routine agarose gel electrophoresis. The primer sequences used are shown in Table 1. The ΔΔCt method was applied to calculate the mRNA amount, and the relative mRNA amount was quantified as $2^{-\Delta\Delta Ct}$. The expression levels of lncRNAs and circRNAs were calculated as described for mRNAs. *GAPDH* was used as a loading control. The expression of miRNA in each sample was normalized to that of U6.

Table 1. Primers used in this study

Gene	Primer Sequence (5' to 3')
RALDH1	F: ACTGTGTCATCTGCTCTG
	R: TTACTCTGCTGGCTTCTT
RALDH2	F: ACATCAACAAGGCTCTCA
	R: CCAAACTCACCCATTTCTC
RALDH3	F: AGAGGGCTGTTCATCAAA
	R: TGCTGTGAGTCCATAGTC
RAR-α	F: AACAACAGCTCAGAACAAC
	R: CGAACTCCACAGTCTTAATG
RAR-β	F: CTTGGGCCTCTGGGACAAAT
	R: TGGCGAACTCCACGATCTTAAT
RAR-γ	F: AGACTTTTCCCTCACTCTG
	R: TTCGCAAACTCCACAATC
GAPDH	F: AACCTGCCAAGTATGATGA
	R: GGAGTTGCTGTTGAAGTC
Cyp4f18	F: GGCCGTGGCACCCCGTTATC
	R: CGGCCTCAGGAAACGGTAGAAGACTC
Foxo4	F: CCGGAAATCCTGGGGGCTGTAAC
	R: CCGGGGCGCTTTCAATGGC
rno-miR-466c-3p	F: TATACATGCACACATACACAC
rno-miR-187-5p	F: AGGCTACAACACAGGACC
U6	F: CTCGCTTCGGCAGCACA
chr15_23792823_23793342_+	F: TGCCCCTGTATTCTTCCTCGTCTGTC
	R: CAGGGAAGTGGCAGGGTGGATCA
chr5_50556456_51183813_-	F: CGCGGTCAAAATGGTGGATGTAGA
	R: CCCGCTCCCCAGTACATGAACC

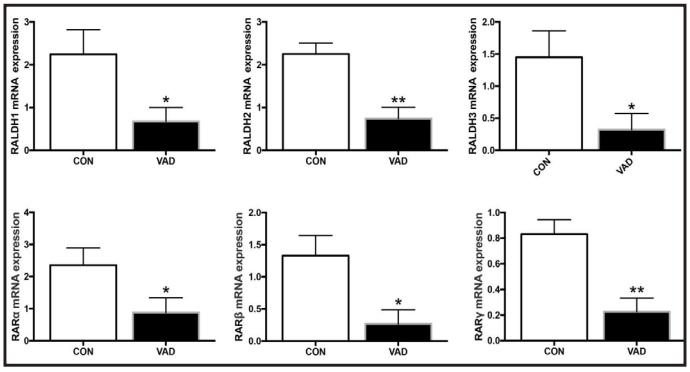

Fig. 1. Transcriptional levels of retinaldehyde dehydrogenase (RALDH1, RALDH2, and RALDH3) and retinoic acid receptors (RAR-α, RAR-β, and RAR-γ) in rat embryos of the VAD-CS (n=3) and control (n=3) groups. *P<0.05, **P<0.001, VAD-CS group vs. the control group.

Statistical methods

All results are represented as means ± standard error of the mean. RT-PCR data were analyzed by one-way analysis of variance (ANOVA), then Tukey's multiple comparison test. Statistical significance was accepted when $P < 0.05$.

Results

CS model identification

To determine the ceRNA regulatory networks in CS rats, the VAD-induced CS model was established with Sprague Dawley rats. There were significant reductions in the mRNA expression levels of *RAR-γ* and *RALDH2* of rat embryos in the VAD-CS group versus those in the control group ($P < 0.001$). Moreover, as detected by RT-PCR, the mRNA expression levels of *RALDH1, 3* and *RAR-α, -β* in rat embryos of the VAD-CS group were notably less than those of the control group ($P < 0.05$) (Fig. 1).

DE mRNAs and ncRNAs

To investigate whether ncRNAs participate in the pathogenesis of CS, the GD 9 rat embryos were examined using the sequencing technique. The DE mRNAs and ncRNAs were analyzed by RNA sequencing significance analysis using Cuffdiff software and the criteria of $q < 0.05$. Fig. 2 shows the clustering heatmap of DE mRNAs, miRNAs, lncRNAs, and circRNAs, respectively.

There were a total of 194 upregulated and 555 downregulated mRNAs in the GD 9 rat embryos of the VAD-CS group versus the control group, respectively. In addition, there were 685 DE lncRNAs, including 185 upregulated and 500 downregulated lncRNAs; 70 DE circRNAs, including 35 upregulated and 35 downregulated circRNAs; and 56 DE miRNAs, including 53 upregulated and 3 downregulated miRNAs, in the GD 9 rat embryos of the VAD-CS group versus the control group.

The histograms summarizing the DE lncRNAs, circRNAs, miRNAs, and mRNAs are shown in Fig. 3. The expression of target genes (mRNAs) was regulated by lncRNAs *via* coregulation or colocalization. The DE mRNAs may be further indirectly or directly regulated by lncRNAs if the target genes of the lncRNAs are the same as the DE mRNAs. As shown in Fig. 3B, the Venn diagram represents the intersectional analysis between the target mRNAs of the co-expression or colocalization with lncRNAs and DE mRNAs. To acquire evidence whether DE miRNAs and DE circRNAs might reveal modified matching DE mRNAs, the intersection of DE mRNAs with DE circRNAs (Fig. 3C) or DE miRNAs (Fig. 3D) was selected.

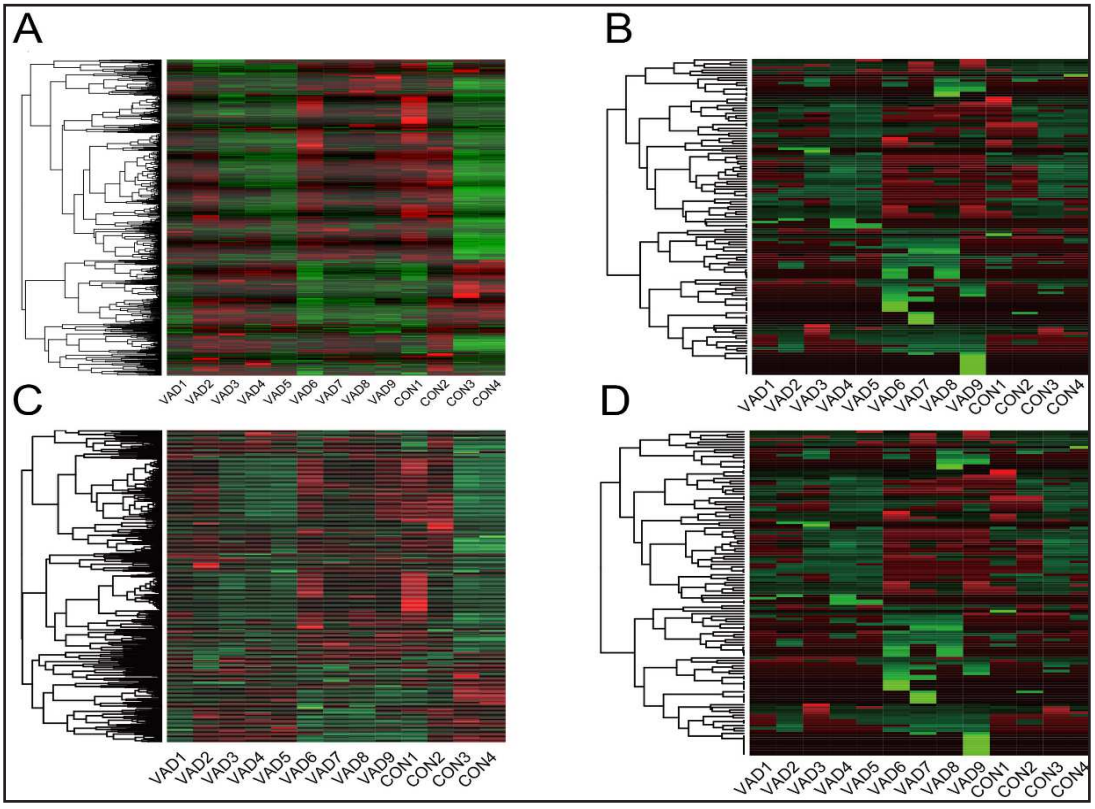

Fig. 2. Expression profile alterations of mRNA (A), miRNA (B), lncRNA (C), and circRNA (D) in rat embryos. Heat map of mRNAs and ncRNA showing hierarchical clustering of regulated ncRNAs and mRNAs in rat embryos from the VAD-CS group (n=9) compared with the control group (n=4). Red and blue represent upregulated and downregulated genes, respectively, according to clustering analysis.

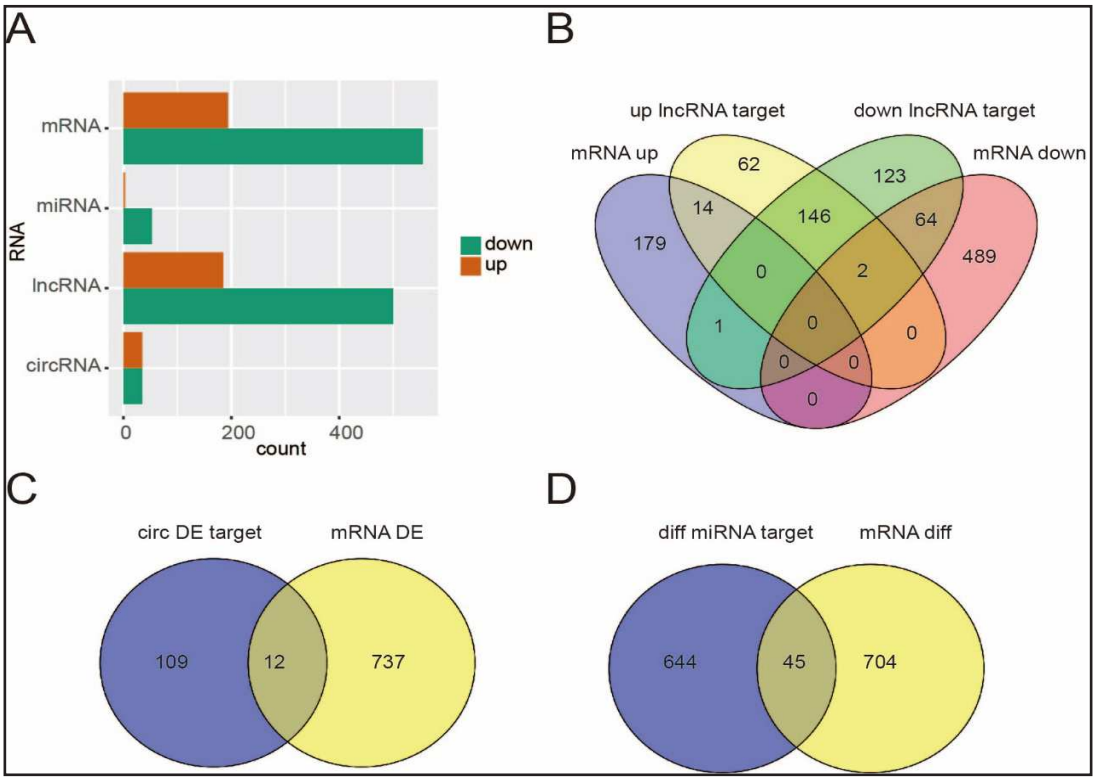

Fig. 3. Relative expression of mRNAs and ncRNAs in rat embryos. (A) Histogram representing the numbers of upregulated and downregulated mRNAs and ncRNAs in rat embryos from VAD-CS group (n=9) vs control group (n=4); (B) Venn diagram representing the overlapping numbers of upregulated lncRNA-targeted mRNAs, downregulated lncRNA-targeted mRNAs, upregulated mRNAs, and downregulated mRNAs; (C) Venn diagram representing the overlapping numbers of targeted mRNAs of DE circRNAs and DE mRNAs; (D) Venn diagram representing the overlapping numbers of targeted mRNAs of DE miRNAs and DE mRNAs.

Confirmation of ncRNA and mRNA expression by real-time PCR

To confirm the dependability of the outcomes by sequencing and to offer a foundation for advanced investigation, the expression variations of selected mRNAs and ncRNAs in GD 9 embryos were studied. mRNAs (Cyp4f18 and Foxo4), miRNAs (rno-miR-466c-3p and rno-miR-187-5p), lncRNAs (NONRATG024332.1 and NONRATG027649.1), and circRNAs (chr15_23792823_23793342_+ and chr5_50556456_51183813_-) were analyzed by real-time PCR (Fig. 4A). Fig. 4B represents the ncRNA and mRNA expression levels detected by sequencing. All the confirmed mRNAs and ncRNAs were in line with the data acquired by the second generation sequencing.

DE ncRNA functional prediction in VAD-CS rats

To further study the gene functions of ncRNAs in VAD-CS rats, genes with an absolute correlation value >0.99 were selected and used to predict the function of ncRNAs by GO and KEGG pathway analyses. The nearby protein-coding genes might be regulated by lncRNAs.

The colocalization threshold was set as 100 kb downstream and upstream of the lncRNAs, and the key functions of DE lncRNAs were forecasted by enrichment analysis.

GO is the classic cataloging system of gene function [15]. Directed acyclic graphs (DAGs) graphically exhibit GO enrichment analysis outcomes for the DE genes. The branch symbolizes the inclusion association, which describes the range from progressively small from the top to the bottom. The top 10 outcomes of GO enrichment analysis were chosen to be the DAG master node and were shown with the GO term, including the association. The color depth denotes the degree of enrichment. The DAGs of CC, BP, and MF as well as the GO

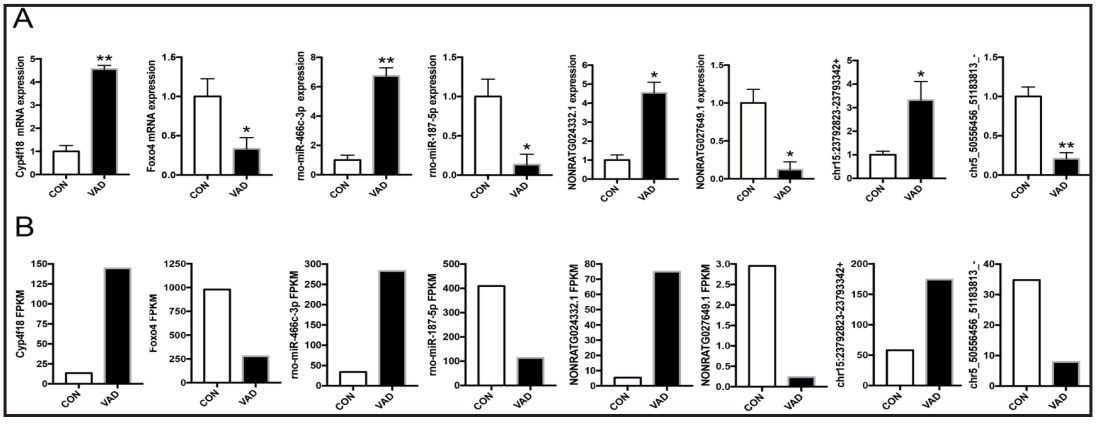

Fig. 4. Real-time PCR verifications of the eight regulated RNAs in rat embryos from the VAD-CS group (n=9) and control groups (n=4). (A) The expression levels of mRNAs, miRNAs, lncRNAs, and circRNAs were notably regulated at GD 9 after VAD. One-way ANOVA was used, followed by Tukey's multiple comparison test. *P<0.05, **P<0.001. (B) Sequencing results of mRNAs, miRNAs, lncRNAs, and circRNAs with real-time PCR validation.

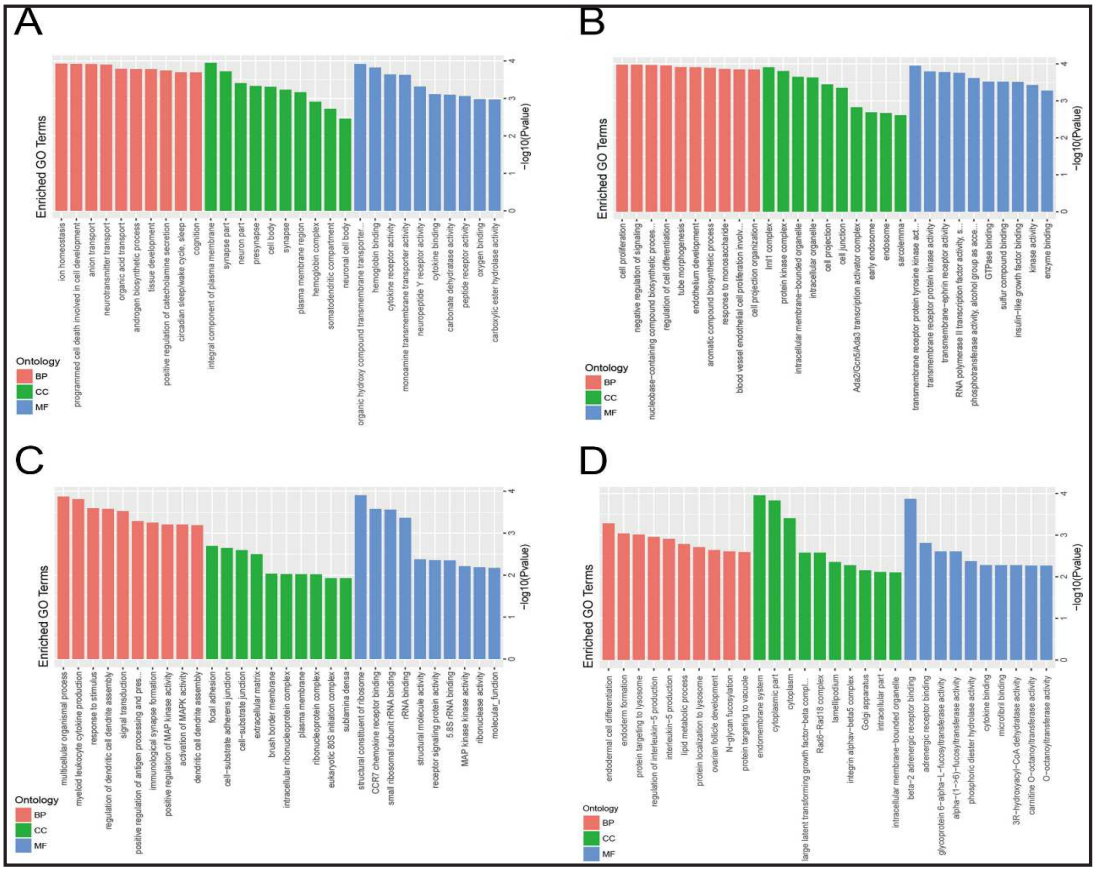

Fig. 5. GO analysis of mRNAs (A), miRNAs (B), lncRNAs (C), and circRNAs (D) in rat embryos of VAD-CS group (n=9) and control group (n=4). Gene numbers in the GO term were presented in histographs.

enrichment analysis by co-expression and co-localization of genes of DE mRNAs, miRNAs, lncRNAs, and circRNA are shown in Fig. 5.

The gene functions of DE ncRNAs were elucidated based on the predicted target gene dissemination in GO. The gene numbers were statistically analyzed by each GO term enrichment and are shown in the form of histograms.

The GO analysis of DE mRNAs showed that the most meaningfully enriched BPs were ion homeostasis and programmed cell death involved in cell development; the most remarkably enriched CCs were the integral component of the plasma membrane and synapse; and the most considerably enriched MFs were allowing the transfer of organic hydroxy compounds from one side of a membrane to the other side and hemoglobin binding (Fig. 5A). The targeted gene GO analysis for DE miRNAs showed that the most meaningfully enriched BPs were cell proliferation, negative regulation of signaling, and nucleobase-containing compound

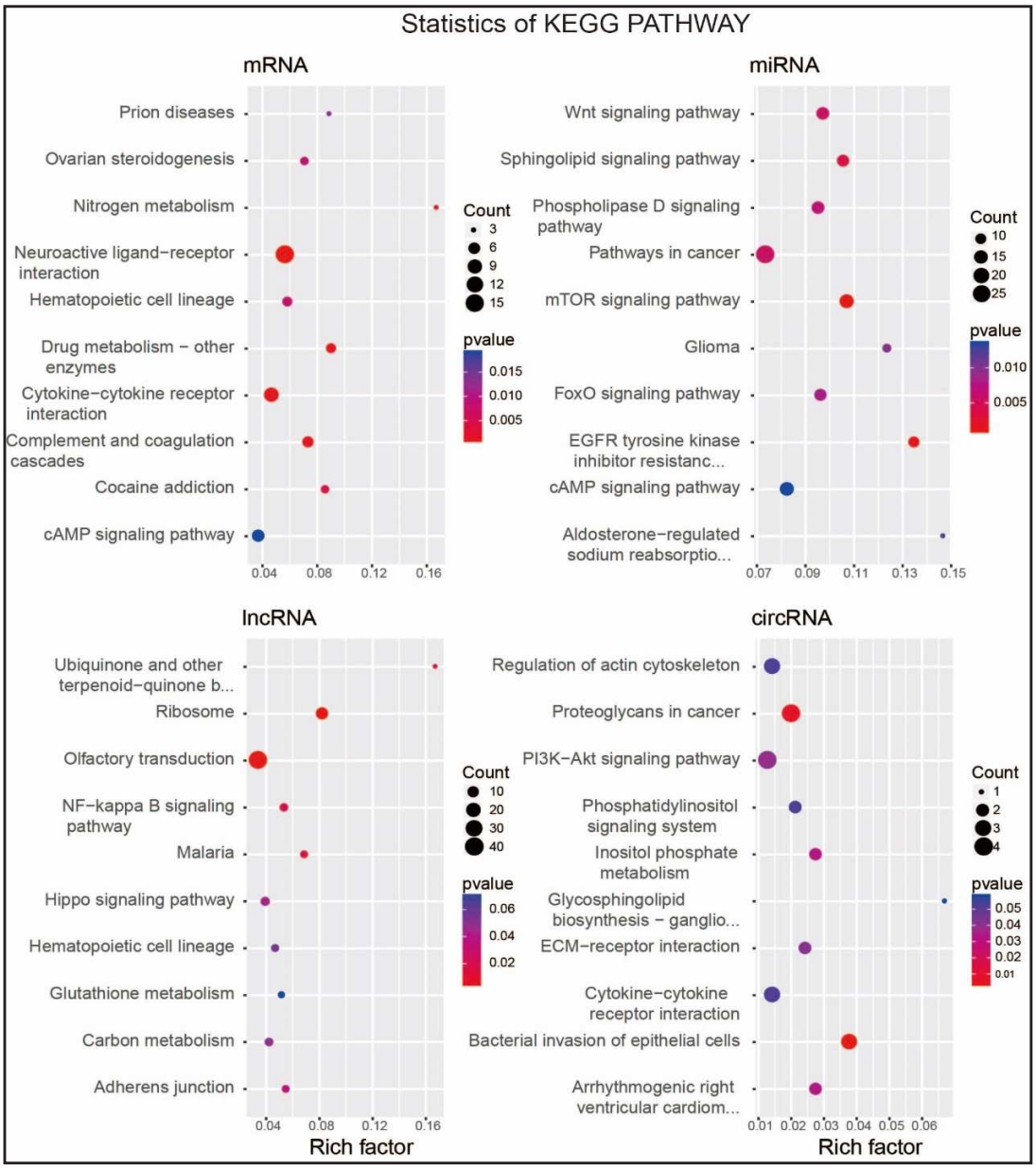

Fig. 6. mRNAs (A), miRNAs (B), lncRNAs (C), and circRNAs (D) enriched in the KEGG pathway scatterplot in rat embryos from VAD-CS group (n=9) and control group (n=4).

biosynthetic processes; the most notably enriched CCs were Iml1 complex and protein kinase complex; and the most meaningfully enriched MF was transmembrane receptor protein tyrosine kinase activation (Fig. 5B). The GO analysis for the co-expressed and colocalized genes of DE lncRNAs showed that the most significantly enriched BPs were the evolution process, localization, and development of multicellular organisms; the most significantly enriched CCs were focal adhesion, cell substracte adherens junction, and extracellular matrix; and the most significantly enriched MF were structural constituent of ribosome and CCR7 chemokine receptor binding (Fig. 5C). According to the GO analysis for the intersection of DE circRNA- and DE miRNA-targeted genes with predicted mRNAs, the most noteworthy enriched BP was endodermal cell differentiation; the most significantly enriched CCs were the endomembrane system and the cytoplasmic part; and the most significantly enriched MF was beta-2 adrenergic receptor binding (Fig. 5D). The most noticeable gene function classification will be the focus of an upcoming investigation.

Diverse genes work together to achieve their biological functions *in vivo*. The significant enrichment pathway could be used to determine the key signaling and biochemical pathways associated with the selected target genes.

KEGG is a primary public pathway database. Based on the entire genome sequence information, KEGG can determine the most noteworthy enrichment pathway in the selected target genes by adopting the KEGG pathway as a unit and applying the hypergeometric test [16, 17]. The graphic exhibition of KEGG enrichment analysis represents the augmented scatter diagram of the selected target genes. In this graph, the enriched degree of KEGG is evaluated using the Rich factor, gene numbers, and *Q*-value. The enrichment will be more significant with a greater Rich factor, a *Q*-value closer to zero, and a greater gene number. The top 10 pathways are shown in Fig. 6. The most notably associated pathways in CS pathogenesis were Wnt, PI3K-ATK, FoxO, EGFR, and mTOR when the target mRNAs were predicted using KEGG analysis based on the DE miRNAs. These results were the same as the enriched pathways examined with expected genes during CS pathogenesis (Fig. 6). The key signaling and biochemical pathways explored by KEGG may afford a better assessment of upcoming investigation guidelines of ncRNAs.

ceRNA regulatory network

To discover the molecular mechanism of ncRNAs associated with CS pathogenesis, analysis of a regulatory network for ncRNAs and mRNAs was conducted. Our results showed that lncRNAs have wide regulatory functions and directly control the DNA structure, transcription, and RNA translation. Furthermore, lncRNAs and circRNAs can constrain target

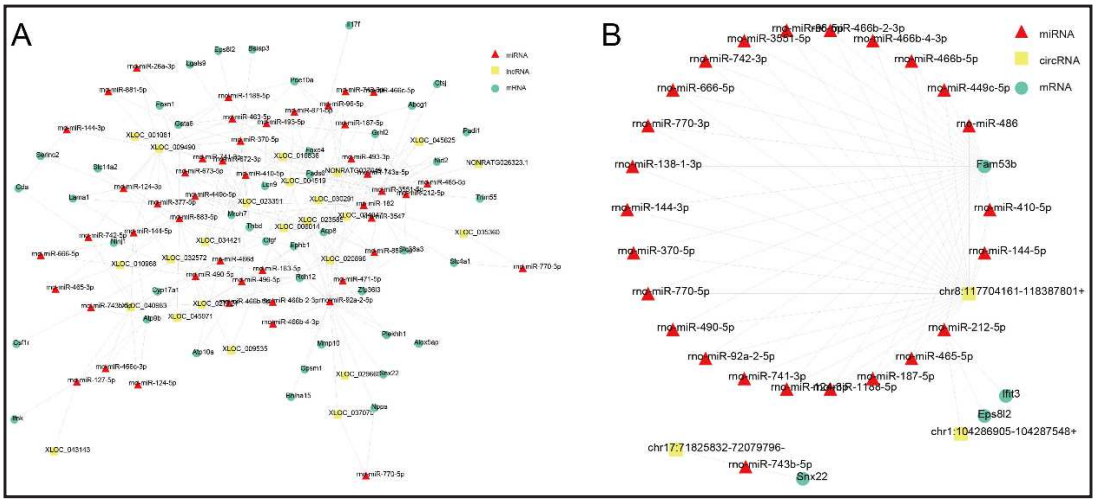

Fig. 7. Regulatory network analysis of lncRNAs-miRNAs-mRNAs (A) and circRNAs-miRNAs-mRNAs (B) in rat embryos from VAD-CS group (n=9) and control group (n=4).

gene regulation of miRNAs, thus indirectly adjusting gene expression and serving as a miRNA sponge. According to the theory of ceRNA, pairs of lncRNA/circRNA with identical miRNA binding sites were recognized. By establishing pairs of lncRNA/circRNA-miRNA-gene and lncRNA/circRNA (as a decoy), mRNA (as the target), and miRNA (as the center), the ceRNA regulation networks were constructed (Fig. 7).

Discussion

In the current work, we found that certain mRNAs and ncRNAs were significantly regulated in embryos from rats with VAD-CS. The potential functions of DE ncRNAs were also predicted in the VAD-CS model using GO and KEGG pathway analyses. Moreover, the ceRNA regulatory networks were built. Based on these findings, we proposed that ceRNAs play a critical role in CS pathogenesis. Furthermore, sequencing analysis disclosed prospective pathogenic targets for CS.

CS is caused by irregular vertebrae growth, including unsuccessful formation and segmentation through embryogenesis, and is often related to other organ defects [18, 19]. The etiology and mechanism of CS remain unknown.

Retinoic acid, the active vitamin A, plays vital roles in numerous physiological processes, such as chondrogenesis, osteochondral development, and cell differentiation [20-22]. Including retinoic acid signaling pathways, at least four signaling pathways have been reported to regulate the segmentation clock, such as the Notch, Fgf, and Wnt pathways [7, 8, 23].

The expression levels of RARα2 and RARβ2 are severely reduced in VAD embryos. Supplementation of retinoic acid or retinol to VAD embryos at or before the 4/5 somite stage restores the expression levels of RARα2 and RARβ2 within about 45 min and rescues normal development. RARβ2 expression requires RARα2 expression [24]. Embryogenesis needs the active form of vitamin A and retinoic acid to active the retinoid receptor during neurulation. RARγ is the most powerful mediator of retinoid signaling at this specific time of development. In addition, it has been reported that a distinct retinoid receptor pathway regulates the critical RA-required developmental events in early avian embryos [25]. In our current study, VAD suppressed the expression levels of RARs and RALDHs, which are key elements of retinoic acid signaling in VAD-CS rat embryos.

ncRNA dysregulation is related to many illnesses, such as CS. To detect the potential ncRNAs associated with the pathogenesis of CS, in this research, we selected the embryos on GD 9 to perform transcriptome sequencing during the stage of somitogenesis [26, 27]. Meanwhile, the DE ncRNAs were further investigated, and the regulatory interaction and function between ncRNAs and mRNAs in VAD-CS were predicted.

Though we identified DE ncRNAs in rat embryos of VAD-CS and confirmed some ncRNAs, the fundamental mechanisms of ncNRAs in CS are rarely known. The GO is a key bioinformatics method to unify the illustration and product attributes of genes across all species. Terms and annotations of GO have demonstrated the designated worthy predictors of gene function and trends. KEGG pathway databases are more extensively used in present enrichment analysis, which includes the advanced order functional evidence for gene function systematic analysis.

Somitogenesis is an accurately controlled multistep process that is controlled by a molecular oscillator of the segmentation clock. This temporal periodicity is then translated into a tightly regulated spatial one, which is revealed by the expression of genes such as *EGF*, *DLL3*, and notch ligand [28]. Meanwhile, PI3K, as a PDGFR alpha downstream effector, is crucial for cell migration of the somite-derived dorsal mesenchyme, and interruption of receptor signaling in these cells results in spina bifida [29].

Therefore, ncRNA-associated functions and related pathways in the embryos of VAD-CS rats using GO and KEGG term enrichment have been analyzed in our current study, and Wnt, PI3K-ATK, FoxO, EGFR, and mTOR were shown to be the most comprehensively involved

pathways in CS pathogenesis. However, the ncRNA functions forecasted by KEGG and GO analysis in CS must be extensively investigated in upcoming research.

RNA with miRNA response elements may serve as ceRNAs, including pseudogenes, circRNAs, lncRNAs, and mRNAs, which function as natural miRNA sponges to inhibit miRNA functions by sharing miRNA response elements to regulate each other [30].

The underlying mechanisms of the interaction between ncRNAs and mRNAs in CS still largely remain unknown. Therefore, ncRNA/circRNA-miRNA-mRNA of CS was constructed for the first time according to our data from second generation sequencing. These groundbreaking findings will improve our knowledge on the function of ncRNAs in the pathogenesis of CS. The ceRNA network prediction and bioinformatics analysis also provide comprehensive understanding of ncRNA functions, which may be involved in the initiation and progression of disease [31].

It is worth stating that embryos contain diverse cell populations, such as numerous types of embryonic stem cells, somite cells, and notochord cells. The GO, KEGG, and ceRNA regulatory network analyses conducted in this work only considered the embryo as a whole. As a result, further *in vitro* and *in vivo* experiments must be conducted. The single-cell sequencing analysis with somite or notochord cells of the embryo is currently being conducted by our team. The data will complement our current findings, thus illuminating functions of ncRNAs in CS embryos more accurately.

In summary, for the first time, our present study identified the comprehensive expression profile of ncRNAs in CS and offered a groundbreaking data incorporation analysis of ceRNAs in CS. A catalog of predicted ncRNAs in CS was also provided. These abundant data prompted us to propose that ceRNAs may adjust the expression of their associated protein-coding genes and play a crucial role in CS.

Future studies should further explore these predicted ceRNAs regarding the comprehensive proteomic and relevant signaling pathways, which may eventually achieve the complete discovery of the underlying mechanisms of CS.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (81330044 and 81772424 to Jianxiong Shen).

Disclosure Statement

The authors declare no conflicts of interest.

References

1 Burnei G, Gavriliu S, Vlad C, Georgescu I, Ghita RA, Dughila C, Japie EM, Onila A: Congenital scoliosis: An up-to-date. J Med Life 2015;8:388-397.
2 Wynne-Davies R: Congenital vertebral anomalies: Aetiology and relationship to spina bifida cystica. J Med Gen 1975;12:280-288.
3 Giampietro PF, Dunwoodie SL, Kusumi K, Pourquie O, Tassy O, Offiah AC, Cornier AS, Alman BA, Blank RD, Raggio CL, Glurich I, Turnpenny PD: Progress in the understanding of the genetic etiology of vertebral segmentation disorders in humans. Ann N Y Acad Sci 2009;1151:38-67.
4 Pourquie O: Vertebrate segmentation: From cyclic gene networks to scoliosis. Cell 2011;145:650-663.
5 Wu N, Ming X, Xiao J, Wu Z, Chen X, Shinawi M, Shen Y, Yu G, Liu J, Xie H, Gucev ZS, Liu S, Yang N, Al-Kateb H, Chen J, Zhang J, Hauser N, Zhang T, Tasic V, Liu P et al.: Tbx6 null variants and a common hypomorphic allele in congenital scoliosis. New Engl J Med 2015;372:341-350.

6 Giampietro PF, Blank RD, Raggio CL, Merchant S, Jacobsen FS, Faciszewski T, Shukla SK, Greenlee AR, Reynolds C, Schowalter DB: Congenital and idiopathic scoliosis: Clinical and genetic aspects. Clin Med Res 2003;1:125-136.
7 Vermot J, Gallego Llamas J, Fraulob V, Niederreither K, Chambon P, Dolle P: Retinoic acid controls the bilateral symmetry of somite formation in the mouse embryo. Science 2005;308:563-566.
8 Vermot J, Pourquie O: Retinoic acid coordinates somitogenesis and left-right patterning in vertebrate embryos. Nature 2005;435:215-220.
9 Vilhais-Neto GC, Maruhashi M, Smith KT, Vasseur-Cognet M, Peterson AS, Workman JL, Pourquie O: Rere controls retinoic acid signalling and somite bilateral symmetry. Nature 2010;463:953-957.
10 Li Z, Shen J, Wu WK, Wang X, Liang J, Qiu G, Liu J: Vitamin a deficiency induces congenital spinal deformities in rats. PLoS One 2012;7:e46565.
11 Thum T: Noncoding rnas and myocardial fibrosis. Nat Rev Cardiol 2014;11:655-663.
12 Hansen TB, Jensen TI, Clausen BH, Bramsen JB, Finsen B, Damgaard CK, Kjems J: Natural rna circles function as efficient microrna sponges. Nature 2013;495:384-388.
13 White JC, Shankar VN, Highland M, Epstein ML, DeLuca HF, Clagett-Dame M: Defects in embryonic hindbrain development and fetal resorption resulting from vitamin a deficiency in the rat are prevented by feeding pharmacological levels of all-trans-retinoic acid. Proc Nat Acad Sci U S A 1998;95:13459-13464.
14 Azevedo PS, Minicucci MF, Chiuso-Minicucci F, Justulin LA, Jr., Matsubara LS, Matsubara BB, Novelli E, Seiva F, Ebaid G, Campana AO, Zornoff LA, Paiva SA: Ventricular remodeling induced by tissue vitamin a deficiency in rats. Cell Physiol Biochem 2010;26:395-402.
15 Young MD, Wakefield MJ, Smyth GK, Oshlack A: Gene ontology analysis for rna-seq: Accounting for selection bias. Genome Biol 2010;11:R14.
16 Mao X, Cai T, Olyarchuk JG, Wei L: Automated genome annotation and pathway identification using the kegg orthology (ko) as a controlled vocabulary. Bioinformatics 2005;21:3787-3793.
17 Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, Katayama T, Kawashima S, Okuda S, Tokimatsu T, Yamanishi Y: Kegg for linking genomes to life and the environment. Nucleic Acids Res 2008;36:D480-484.
18 Shen J, Zhang J, Feng F, Wang Y, Qiu G, Li Z: Corrective surgery for congenital scoliosis associated with split cord malformation: It may be safe to leave diastematomyelia untreated in patients with intact or stable neurological status. J Bone Joint Surg Am 2016;98:926-936.
19 Shen J, Wang Z, Liu J, Xue X, Qiu G: Abnormalities associated with congenital scoliosis: A retrospective study of 226 chinese surgical cases. Spine (Phila Pa 1976) 2013;38:814-818.
20 Zhang H, Li N, Tang Y, Wu W, Zhang Q, Yu Z: Negative functional interaction of retinoic acid and tgf-beta signaling mediated by tg-interacting factor during chondrogenesis. Cell Physiol Biochem 2009;23:157-164.
21 Diederichs S, Zachert K, Raiss P, Richter W: Regulating chondrogenesis of human mesenchymal stromal cells with a retinoic acid receptor-beta inhibitor: Differential sensitivity of chondral versus osteochondral development. Cell Physiol Biochem 2014;33:1607-1619.
22 Mendoza-Parra MA, Walia M, Sankar M, Gronemeyer H: Dissecting the retinoid-induced differentiation of f9 embryonal stem cells by integrative genomics. Mol Syst Biol 2011;7:538.
23 Wahl MB, Deng C, Lewandoski M, Pourquie O: Fgf signaling acts upstream of the notch and wnt signaling pathways to control segmentation clock oscillations in mouse somitogenesis. Development 2007;134:4033-4041.
24 Cui J, Michaille JJ, Jiang W, Zile MH: Retinoid receptors and vitamin a deficiency: Differential patterns of transcription during early avian development and the rapid induction of rars by retinoic acid. Dev Biol 2003;260:496-511.
25 Romeih M, Cui J, Michaille JJ, Jiang W, Zile MH: Function of rargamma and raralpha2 at the initiation of retinoid signaling is essential for avian embryo survival and for distinct events in cardiac morphogenesis. Dev Dyn 2003;228:697-708.
26 Narotsky MG, Schmid JE, Andrews JE, Kavlock RJ: Effects of boric acid on axial skeletal development in rats. Biol Trace Elem Res 1998;66:373-394.
27 Kimmel CA, Cuff JM, Kimmel GL, Heredia DJ, Tudor N, Silverman PM, Chen J: Skeletal development following heat exposure in the rat. Teratology 1993;47:229-242.
28 Eckalbar WL, Lasku E, Infante CR, Elsey RM, Markov GJ, Allen AN, Corneveaux JJ, Losos JB, DeNardo DF, Huentelman MJ, Wilson-Rawls J, Rawls A, Kusumi K: Somitogenesis in the anole lizard and alligator reveals evolutionary convergence and divergence in the amniote segmentation clock. Dev Biol 2012;363:308-319.

29 Pickett EA, Olsen GS, Tallquist MD: Disruption of pdgfralpha-initiated pi3k activation and migration of somite derivatives leads to spina bifida. Development 2008;135:589-598.
30 Ashwal-Fluss R, Meyer M, Pamudurti NR, Ivanov A, Bartok O, Hanan M, Evantal N, Memczak S, Rajewsky N, Kadener S: Circrna biogenesis competes with pre-mrna splicing. Mol Cell 2014;56:55-66.
31 Liu W, Zhang J, Zou C, Xie X, Wang Y, Wang B, Zhao Z, Tu J, Wang X, Li H, Shen J, Yin J: Microarray expression profile and functional analysis of circular rnas in osteosarcoma. Cell Physiol Biochem 2017;43:969-985.