

Received: 2018.03.27

Accepted: 2018.04.16

Published: 2018.05.08

Functional and *In Silico* Assessment of *GDF3* Gene Variants in a Chinese Congenital Scoliosis Population

Authors' Contribution:
Study Design A
Data Collection B
Statistical Analysis C
Data Interpretation D
Manuscript Preparation E
Literature Search F
Funds Collection G

ABCDEF 1,2,3 **Jia Chen**
BCDEF 2,3 **Xiaoxin Li**
BCDEF 2,3 **Yuchen Niu**
ABCDEG 1,2,3 **Zhihong Wu**
ABCDEFG 1,2,3 **Guixing Qiu**

1 Department of Orthopedic Surgery, Peking Union Medical College Hospital, Peking Union Medical College and Chinese Academy of Medical Sciences, Beijing, P.R. China
2 Beijing Key Laboratory for Genetic Research on Skeletal Deformity, Beijing, P.R. China
3 Medical Research Center of Orthopedics, Chinese Academy of Medical Sciences, Beijing, P.R. China

Corresponding Authors: Guixing Qiu, e-mail: qjuguixing.pumch.orth@gmail.com, Zhihong Wu, e-mail: orthoscience_pumch@163.com
Source of support: The study was funded by the National Natural Science Foundation of China (NSFC, grant numbers: 81472046, 81472045, 81772299)

Background: The present study aimed to evaluate the pathogenicity of 5 *GDF3* gene variations using functional and *in silico* assessment approaches in a Chinese congenital scoliosis population.

Material/Methods: We selected 13 patients carrying 5 variants from a congenital scoliosis cohort. The PCR products of samples were verified by Sanger sequencing. The data and sequence alignment were analyzed using Chromas and ClustalW. SIFT and PolyPhen-2 were used to predict the functional effects of each missense and amino acid substitutions. SWISS-MODEL server and Swiss-PdbViewer were used to analyze conformational changes of *GDF3* structure. DUET, UCSF Chimera, and Ligplot software were used to further explore the protein stability, side chains, and hydrophobic interaction changes, respectively. Luciferase reporter gene and Western blot assays were used to perform functional assessments for every variant from the molecular level.

Results: Of the 13 patients, the S212L variant reoccurred in 9 patients. The rest of the patients carried 1 missense mutation each. The variants of R84L and R84C were predicted as probably damaging *loci*. S212L, N215S, A251T were predicted as benign *loci*. In functional assays, R84L, S212L, and A251T display inhibitory effects on functional assays. N251S mutation showed a negative effect in protein expression assays but not in luciferase reporter gene assays. The variant of R84C displayed no negative effects on 2 functional assays.

Conclusions: Our results suggest that the 4 of the 5 variants in *GDF3* gene contribute different pathogenicity in congenital scoliosis, which may provide molecular evidence for clinical genetic testing.

MeSH Keywords: **Computer Simulation • Genetic Variation • Growth Differentiation Factor 3 • Molecular Biology • Scoliosis • Transforming Growth Factor beta**

Full-text PDF: <https://www.medscimonit.com/abstract/index/idArt/910232>

 3262  2  4  23



Background

Congenital scoliosis (CS) is a relatively rare skeletal malformation due to failure of somatogenesis in the process of embryonic development [1,2]. CS can be an independent disorder or integrate with other syndromes, such as spondylocostal dysostosis and spondylothoracic dystrophy [3,4].

GDF3, a TGF superfamily member, belongs to the subclass of bone morphogenetic protein/growth and differentiation factor (BMP/GDF) [5]. Previous studies reported that *GDF3* is a direct BMP signaling inhibitor in early embryos and pluripotent cells [6]. BMP superfamily signaling is critical to many skeletal disorders, especially in skeletal development [7]. Genetic variants in the genome that encode relevant BMP pathway molecules can cause a number of skeletal disorders in humans [8], including variants in *GDF3* gene. It had been reported that missense mutations in *GDF3* gene can result in Klippel-Feil syndrome, which involves fusion of the cervical vertebra [9].

With regard to congenital scoliosis, its spectrum of skeletal deformities can be clinically divided into 3 types: type I is failure of formation, type II is failure of segmentation, and type III is mixed anomalies [10]. Klippel-Feil syndrome is a type of CS that can involve certain genetic defects in vertebral formation [11]. Besides the phenotypes of skeletal malformation, different missense variants in *GDF3* can display ocular-associated anomalies, such as coloboma and microphthalmia, and various mutations can contribute to the molecular role of *GDF3* [9].

Currently, next-generation sequencing (NGS) technology facilitates comprehensive understanding of certain genetic disease based on the global genome setting [12]. However, interpreting and annotating some complex mutations, such as variants of uncertain significance (VUS), is very challenging. Fortunately, studies had reported that robust tools for clinical annotation of certain VUS can be used by employing functional and *in silico* predictors for those pathogenicities [13–15]. Accordingly, these approaches may facilitate the interpretation of certain VUS in clinical practice.

In our congenital scoliosis cohort, we had found 5 genetic VUS in the *GDF3* coding domain after performing whole-exome sequencing (WES). We hypothesized that these mutations play a role in dysfunction of *GDF3*. Therefore, in the present study we used functional and *in silico* assessment approaches to explore the pathogenicity of 5 *GDF3* VUS in our Chinese congenital scoliosis population.

Material and Methods

Participants

Congenital scoliosis patients were enrolled from the Department of Orthopedic Surgery in Peking Union Medical College Hospital, Beijing, China, between Jan 2007 and Feb 2017. This study was conducted in accordance with the Declaration of Helsinki and was approved by the Ethics Committee of the Peking Union Medical College Hospital, Peking Union Medical College and the Chinese Academy of Medical Sciences (ethics approval number: JS-1198). Signed informed consent was obtained from every participant or their legal guardians.

We solicited detailed case histories and performed thorough physical examinations. The inclusion criteria were: (1) congenital scoliosis and (2) complete medical records. The exclusion criteria were: (1) idiopathic scoliosis, neuromuscular scoliosis, and adult degenerative scoliosis; (2) other associated syndromes, such as Marfan's syndrome, VACTERL syndrome, Alagille syndrome, and Jarcho-Levin syndrome; and (3) scoliosis caused by other diseases, and without genetic factors, such as fractures or laminectomy, thoracoplasty, scoliosis caused by radiation therapy, infections, and spinal tumors. Because some *GDF3* gene-associated patients had been reported to suffer from ocular dysplasia, we had 2 experienced ophthalmologists check the ocular-relevant symptoms, including unilateral or bilateral microphthalmia, coloboma, and nystagmus. Other associated abnormalities, such as those of the heart, kidneys, and skin, were also examined. All enrolled individuals were then screened by standing-position, full-length spinal X-ray and magnetic resonance imaging (MRI), and the imaging results were interpreted by at least 2 experienced radiologists.

A 5-mL peripheral blood sample was obtained from every participant. The blood was collected in 2% ethylenediaminetetraacetic acid anticoagulant tubes and fully mixed to prevent clotting. DNA was extracted using the QIAGEN Whole Blood Genomic DNA Mini Kit (QIAGEN Inc., Valencia, CA) according to its standard protocol. Then, the DNA was preserved in 0.5-mL Eppendorf tubes at -80°C for whole-exome sequencing.

PCR Sanger sequencing and *in silico* assessment

After screening WES data according to ACMG guidelines [16], the putative pathogenic mutation needed to be confirmed by PCR Sanger verification in original samples. PCR amplification and program setting were performed according to a previous study [9], followed by bi-directional amplification. Then, PCR products were sent to the Beijing Genomics Institute (BGI) for Sanger sequencing. The data and sequence alignment were analyzed using Chromas (Version 2.6.4, Technelysium Pty Ltd., Australia) and ClustalW (<http://www.genome.jp/tools-bin/clustalw>) [17].

We used SIFT, PolyPhen-2, and MutationTaster to predict the functional effects of each missense and amino acid substitution. We used MutationMapper (http://www.cbioportal.org/mutation_mapper.jsp) to simulate the localization of mutation sites in protein structure domains. In order to predict the 3-dimensional protein structure and the internal molecular segments, we construct mimetic human GDF3 protein by using the SWISS-MODEL (<http://swissmodel.expasy.org>) server [18]. We then analyzed the structure using Swiss-PdbViewer version 3.7b1 (<http://www.expasy.org/spdbv/>) [19]. To further clarify the stability of protein structure, we used DUET, a web server for studying the effect of certain variants on protein stability [20]. In addition, we used the UCSF Chimera to predict the atomic contacts (i.e., steric clashes) of mutant proteins among their side chains. Chimera is an extensible program for interactive visualization and analyzing molecular structures and related data [21]. Moreover, we used Ligplot (www.ebi.ac.uk/thornton-srv/software/LIGPLOT/) to analyze and present the hydrophobic interaction changes during the mutation [22].

Gene cloning and mutagenesis

Full-length coding wild-type *GDF3* was generated by reverse transcription of whole-blood RNA (Takara Shuzo, Kyoto, Japan) and then was cloned into the expression vector pcDNA3.1-Myc-His A (Clontech, Palo Alto, CA, USA). Point mutations were introduced by site-directed mutagenesis (Takara Shuzo, Kyoto, Japan). The SOX9 luciferase reporter vector plasmid was constructed using pGL3-basic and introduced the promoter upstream 1000bp sequence of *SOX9*.

Cell culture and luciferase reporter assay

ATDC5 and COS-7 cells were cultured in high-glucose Dulbecco's modified Eagles medium (DMEM; Gibco) and DMEM/F12 (Gibco), respectively. The 2 cell lines were supplemented with 10% (v/v) fetal bovine serum (FBS, Sigma-Aldrich) and 1% (v/v) penicillin/streptomycin (Gibco). Cells were purchased from the BeNa Cell Collection (Beijing, China) and were routinely tested for mycoplasma. Cell transfection processing was performed using the Lipofectamine 3000 kit (Invitrogen, Carlsbad, CA, USA), following the manufacturer's instructions.

ATDC5 cells were seeded at 10 000 cells per well on 24-well plates for 48 h before transfection. Each well of a 24-well plate was transfected with 50 ng of pcDNA3.1-GDF3, 400 ng of pGL3-SOX9, and 5 ng of pRL-SV40-Renilla reporter. After 48 h, ATDC5 cells were lysed with 100 μ L of 1 \times Passive Lysis Buffer (Promega, Madison, WI). Next, we performed the luminescence using the Dual Luciferase assay system (Promega, Madison, WI) according to the manufacturer's instructions. The data were read using a microplate luminometer (Thermo Scientific, Inc.). The results were reduplicated in 3 independent assays.

Western blot assay

To assess the effect of mutations on protein expression, COS-7 cells were used for GDF3 protein expression. COS-7 cells were seeded at 20 000 cells per well on 6-well plates for 48 h before transfection. Each well of a 6-well plate was transfected with 2 μ g of pcDNA3.1 vector plasmid for expressing GDF3 proteins. Cell lysate and culture medium were collected and extracted proteins at 48 h post-transfection.

Proteins were extracted from cell lysate and supernatant, and then were separated by 15% SDS-PAGE. After transfer to 0.45-mm nitrocellulose membranes (BioRad, CA, USA), proteins were incubated with mouse anti-HIS antibody (1: 1000, cat. no. ab18184, Abcam, Cambridge, MA, USA) and then with HRP-conjugated anti-mouse IgG (1: 10 000, cat. no. ab6789, Abcam, Cambridge, MA, USA). The Western blot bands were developed by ECL reactions (Clarity™ and Clarity Max™ Western ECL Blotting Substrates, BioRad, CA, USA) in a chemiluminescence imager (Amersham Imager 600, Healthcare Life Sciences, GE, Pittsburgh, PA, USA).

Finally, we used stripping buffer (cat. no. 46430; Thermo Scientific, Inc.) to wash off the primary and secondary antibodies. Membranes were subsequently incubated with control primary antibodies. The anti-secreted alkaline phosphatase (1: 5000, cat. no. ab54778, Abcam, Cambridge, MA, USA) and anti-alpha-tubulin (1: 5000, cat. no. ab7291, Abcam, Cambridge, MA, USA) were control antibodies, which were used for detecting cytosol and secreted proteins, respectively. In our study, all the experiments were repeated 3 times independently, the minimum sample size was 3 in luciferase report and Western blot assays.

Statistical analyses

Measurement data of luciferase reporter gene assays were analyzed and presented using GraphPad Prism version 7.03 (GraphPad Software, Inc., CA, USA). The intensity of the Western blots was analyzed by ImageJ software version 1.51 (National Institutes of Health, Bethesda, MD, USA). The data were analyzed by unpaired *t* test. P-values less than 0.05 were considered as statistically significant differences.

Results

Clinical characteristics of enrolled patients

We enrolled 13 of the patients who carried 5 *GDF3* rare variants in our cohort. The clinical imaging features of these patients were: 6 patients had type I vertebral formation, 4 patients had type II vertebral formation, and 3 patients had hybrid vertebral formation type III (Table 1).

Table 1. Clinical features of congenital scoliosis patients with *GDF3* gene mutations.

Patients' No.	Gender	Age	CS type	Spine malformation	Rib malformation	Mutation Type	HGMD transcript	Nucleotide changes
CS01	F	15	I	T12, L1 wedge vertebrae	Missing 12 th rib on the left side	Missense	NM_020634.1	250C>T
CS02	F	4	II	T1~2, L3~4 right fused vertebrae	No	Missense	NM_020634.1	251G>T
CS03	M	4	I	T10 right hemivertebrae	Missing 10 th rib on the left side	Missense	NM_020634.1	635C>T
CS04	F	16	III	Butterfly vertebrae, T3~T4 failure of segmentation	No	Missense	NM_020634.1	635C>T
CS05	F	15	I	L4~L5 semisegmented hemivertebrae	No	Missense	NM_020634.1	635C>T
CS06	F	10	III	T10 left hemivertebrae, T11 wedge vertebrae	Missing 10 th rib on the right side	Missense	NM_020634.1	635C>T
CS07	M	14	I	L1 right hemivertebrae	No	Missense	NM_020634.1	635C>T
CS08	F	5	I	T9 right hemivertebrae	No	Missense	NM_020634.1	635C>T
CS09	F	13	I	L2~L3 semisegmented hemivertebrae	No	Missense	NM_020634.1	635C>T
CS10	M	5	III	T10~L1 failure of segmentation, T10, T11 butterfly vertebrae, T12 left semisegmented hemivertebrae	Missing 12 th rib on the right side	Missense	NM_020634.1	635C>T
CS11	M	14	II	L4~L5 semisegmented hemivertebrae	No	Missense	NM_020634.1	635C>T
CS12	M	10	II	T11~L1 failure of segmentation	No	Missense	NM_020634.1	644A>G
CS13	F	15	II	L4~5, L6~7 failure of segmentation	7 th and 8 th fused ribs	Missense	NM_020634.1	751G>A

CS – congenital scoliosis; T – thoracic vertebra; L – lumbar vertebra; NM – Nucleotide Sequence Number; HGMD – Human Gene Mutation Database.

The Cobb's angle of the main curve in these patients was up to 83 degrees (patient CS10) and the minimum was 36 degrees (patient CS09). After thorough physical examinations, 13 patients in our cohort were determined to have no other abnormal characteristics, such as microphthalmia or coloboma. The main characteristic phenotype of these patients was thoracolumbar scoliosis with sporadic rib deformities. There were 3 segments cuneate vertebral, 9 hemivertebra, 4 butterfly vertebrae, and 18 fusion vertebrae (Figure 1).

Sequencing verification and structural mapping

The DNA samples of all 13 enrolled patients were verified through PCR Sanger sequencing. The sequencing peaks showed that each patient had 1 single missense mutation in the *GDF3* coding domain (Figure 1). R84C, R84L, S212L, and N215S occurred in the propeptide domain. A251T was located in the TGF- β domain. S212L variant recurred in 9 patients. The other 4 variants occurred in single individuals (Figure 2A).

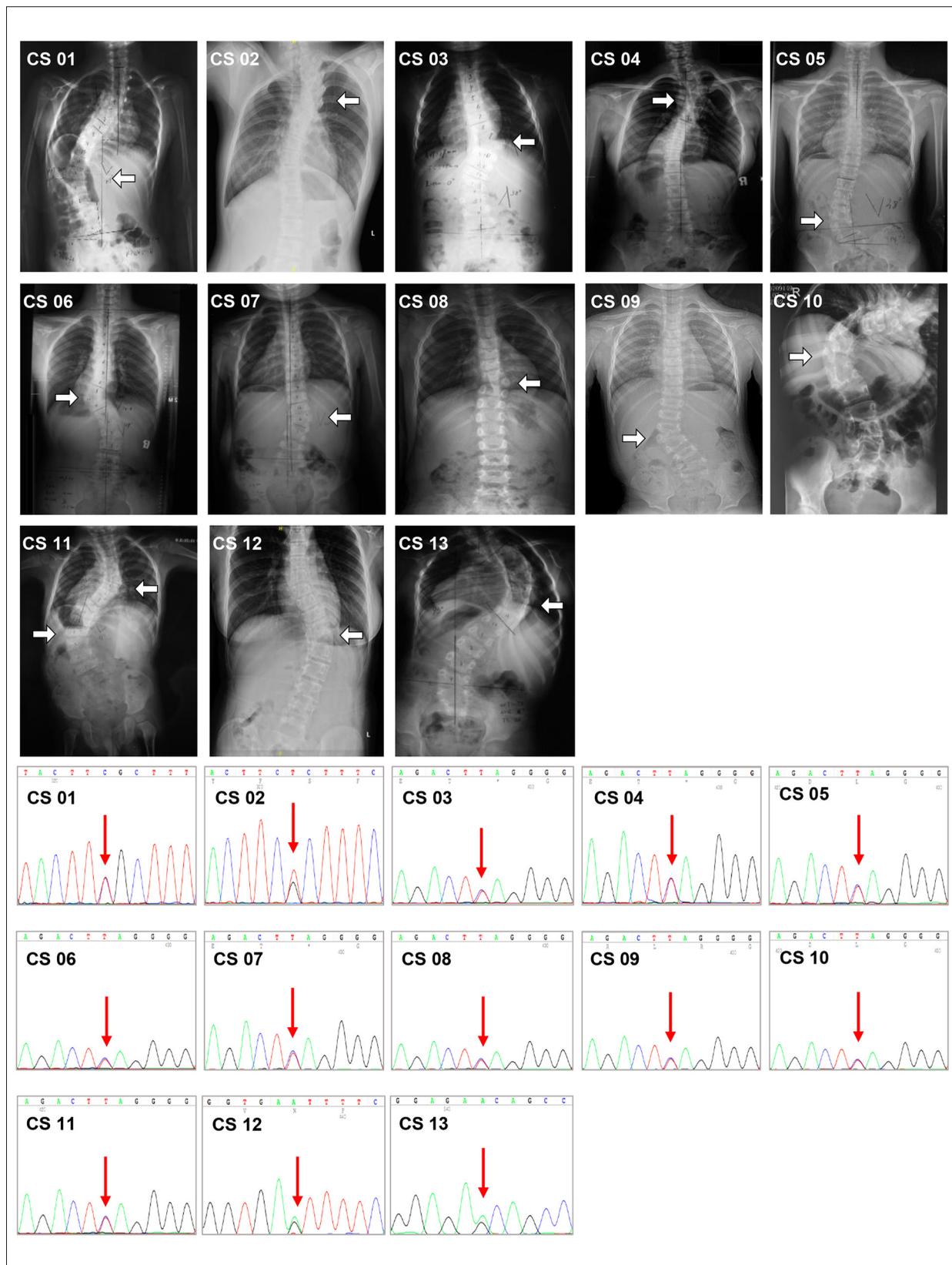


Figure 1. Clinical X-ray examination and corresponding Sanger sequencing for each patient. The white arrows show the apex of scoliosis. The red arrows indicate the location of missense variants in each peak profile.

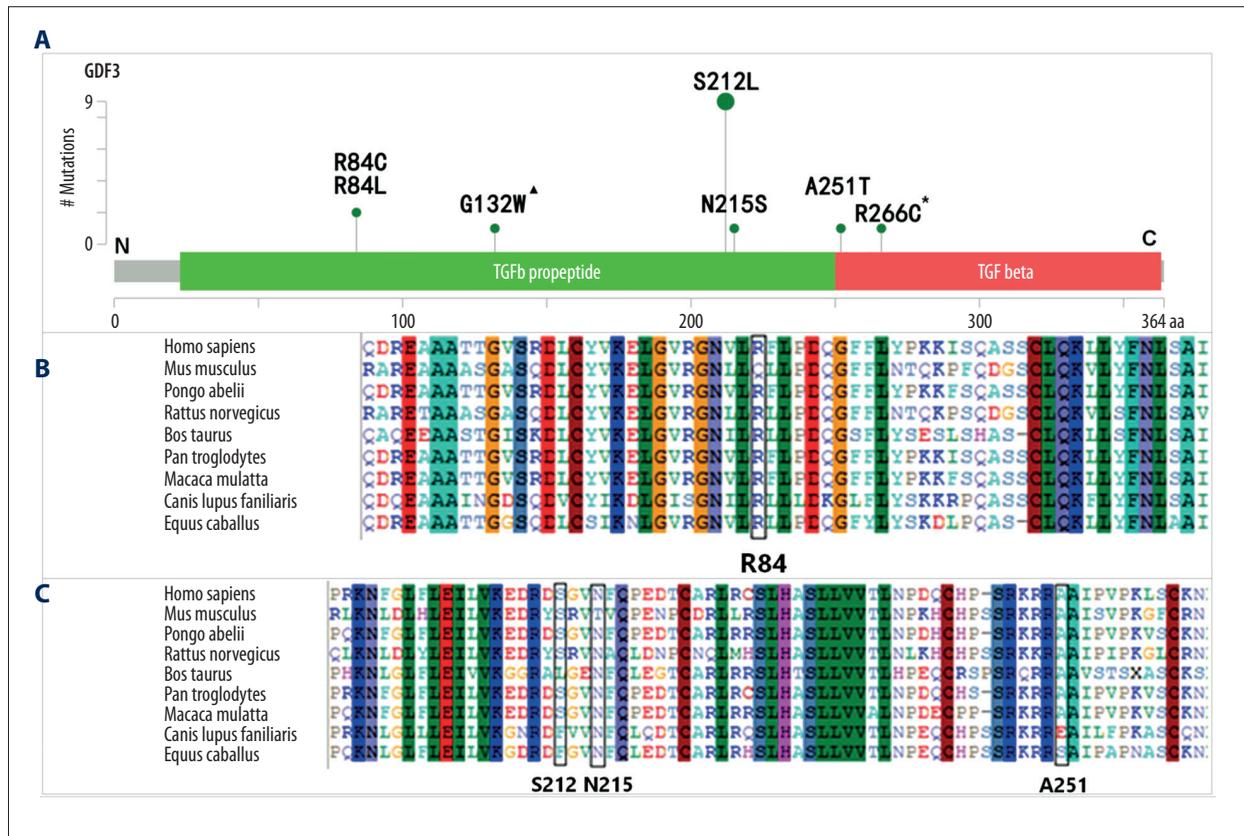


Figure 2. Distribution and multiple sequence alignment of variants. (A) Green box indicates TGF-β propeptide domain, red box indicates TGF-β domain. The green dot indicates the location of each mutation site. The horizontal ordinate indicates the length of the amino acid of GDF3. The vertical ordinate indicates the absolute frequency of mutations in our cohort. The triangle sign indicates the location of G132W, which correspond to negative reference control, and the asterisk sign indicates the location of R266C, which correspond to positive reference control. (B, C) Conservative comparison between *Homo sapiens* and 8 vertebral species for each variant.

Luciferase reporter gene and protein expression assays

We used SOX9-reporter gene according to the method described in a previous study [9,23]. Because *GDF3* is involved in osteogenic signaling pathways, functional changes may affect downstream signaling transduction. The luciferase reporter gene assays in our study showed that some variants had negative effects on SOX9 activation. Compared to the negative control (c.394G>T) and wild-type vector plasmid, the 3 variants (c.251 G>T, c.635 C>T, and c.751 G>A) can interfere with SOX9 activation, while c.250 C>T and c.644 A>G have no effects on downstream signaling transduction (Figure 3A, * P<0.05).

Due to the property of secretory protein, GDF3 can mature by cutting the propeptide domains in cytosol and supernatant. Therefore, protein expression assays were used to evaluate whether the existence of certain unknown variants interfered in the process of GDF3 protein maturation. The results of protein expression demonstrated that 4 variants (R84L, S212L, N215S, and A251T) interfered in maturation of GDF3 protein

in the cytosol (Figure 3B, 3D) and 3 variants (S212L, N215S, and A251T) interfered in maturation of GDF3 protein in the supernatant (Figure 3C, 3E). Compared to control strips, R84C variant had little effect on maturation of GDF3 protein. Beside the wild-type group, we also used G132W (c.394 G>T) as a negative reference control group, which was screened from the control group. In addition, we used R266C (c.796 C>T) as positive reference control, which has been reported to be a pathogenic mutation [9].

In silico assessment of GDF3 variants

By using SIFT and Polyphen-2 for software prediction, the 2 variants (c.250 C>T and c.251 G>T) are predicted as probably damaging mutations. The other 3 mutations (c.635 C>T, c.644 A>G, c.751 G>A) are predicted as benign. The gene frequency of c.635 C>T and c.751 G>A variants are 0.0002883 and 0.000008237, respectively, according to the ExAC database. The 5 variants have not been reported in the HGMD database (Table 2).

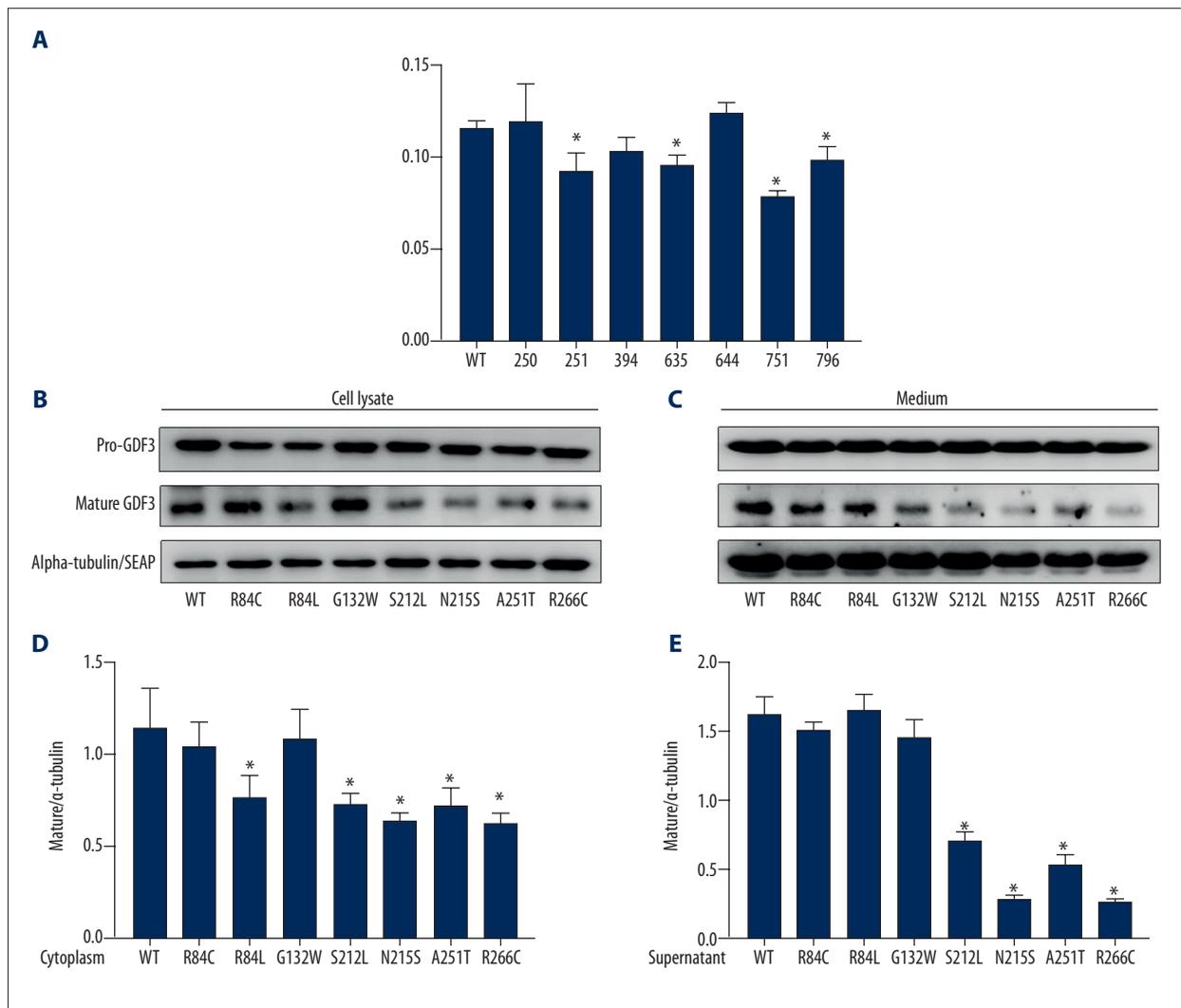


Figure 3. Luciferase reporter gene and protein expression assays. **(A)** Influence of each variant on activity of SOX9 reporter gene. WT group was transfected with normal *GDF3* gene plasmid. The horizontal ordinate indicates each mutation group. The vertical ordinate represents relative luciferase activities. The ratio of firefly luciferase activity to Renilla luciferase activity is presented as the relative luciferase activities. * $P < 0.05$ shows the variant has a significantly negative impact on downstream activation compared with the WT group. **(B, C)** Western blot analyses display the effects of each variant on pro-GDF3 (40 kDa) and mature GDF3 (22 kDa) proteins, respectively. G132W and R266C are negative and positive reference controls, respectively. Alpha-tubulin and secreted alkaline phosphatase represent loading controls for cytosolic and secreted proteins, respectively. SEAP: secreted alkaline phosphatase. **(D)** Columns indicates the relative expression level of mature GDF3 protein in cytoplasm. **(E)** Columns indicates the relative expression level of mature GDF3 protein in supernatant. * $P < 0.05$ shows the variant has a significantly negative effect on the maturation of GDF3 protein compared with the WT group.

Aligning multiple sequences between different species, R84 is evolutionarily conserved (Figure 2B), which indicates that this site has the most significant function. The other 3 mutations have at least 5 conserved vertebrate species in sequence alignment (Figure 2C). We used DUET, a web server, for an integrated computational approach for studying missense mutations, to predict the effects of mutations on protein stability. PolyPhen-2 was also used for predicting the effect of an amino acid substitution on the structure and function of a human protein. The results are shown in Table 2.

The 3D homology structure model of GDF3 was predicted by Swiss-model server according to the template of 4yci. The order of structure sequence is from the N to the C terminus (Figure 4A). The model is highly similar to 4yci in overall folding and secondary structures. We used UCSF Chimera to predict the atomic contacts of mutated proteins among their side chains. If clash occurred, the protein structure will become unstable. According to the analysis results, only p.N215S has 1 steric clash with Leu105 (Figure 4B1, 4B2), and the other 4 mutations all have no clash. In order to assess the conformational change of S212L,

Table 2. Characteristic description and pathogenic prediction of each *GDF3* variant by bioinformatics analyses.

Missense mutation	Amino acid variation	SIFT	PolyPhen-2 score	Pathogenicity prediction	DUET	ExAC allele frequency
250C>T	R84C Arg84Cys	0	1.000	Probably damaging	-0.819 Kcal/mol (Destabilizing)	NA
251G>T	R84L Arg84Leu	0	1.000	Probably damaging	0.202 Kcal/mol (Stabilizing)	NA
635C>T	S212L Ser212Leu	0.37	0.000	Benign	-0.051 Kcal/mol (Destabilizing)	0.0002883
644A>G	N215S Asn215Ser	0.61	0.013	Benign	-1.431 Kcal/mol (Destabilizing)	NA
751G>A	A251T Ala251Thr	0.31	0.124	Benign	-1.739 Kcal/mol (Destabilizing)	0.000008237

NA – not available.

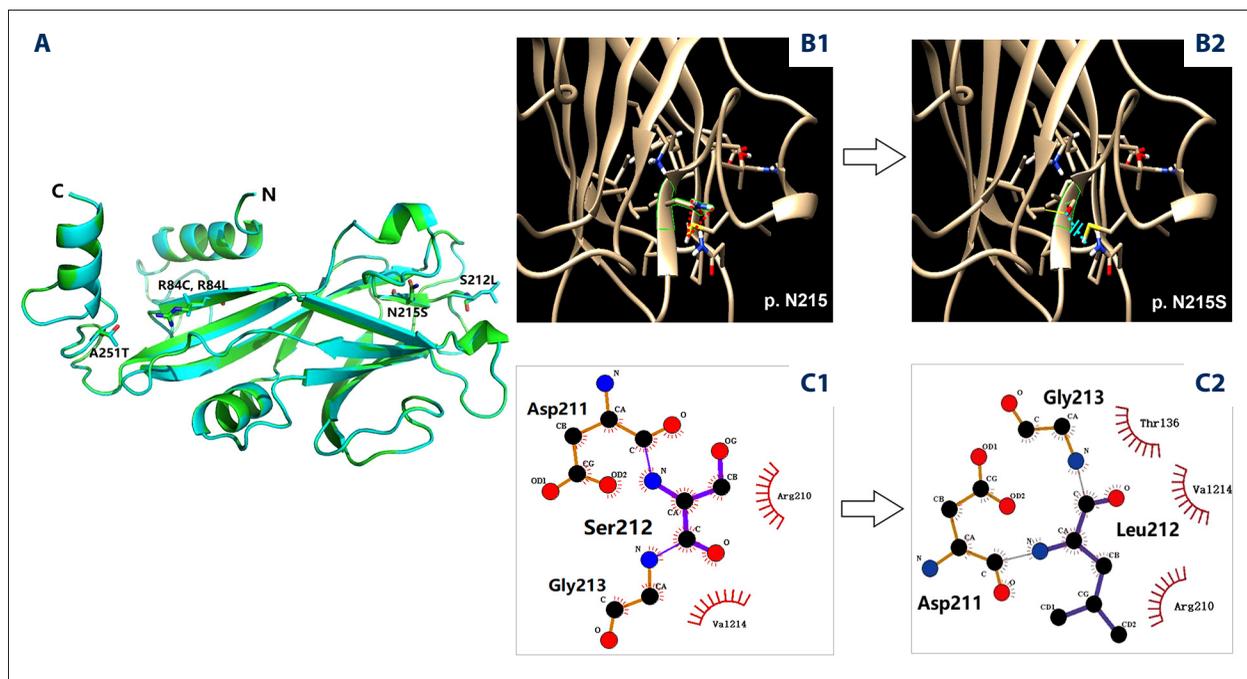


Figure 4. Homology modeling and mutation analysis. **(A)** The 3D homology structure modeling of *GDF3* is predicted by Swiss-model server according to the template of 4yci. The order of structure sequence is from the N to the C terminus. Mutant *loci* are showed in stick style. **(B)** Steric clash analysis of N215S in human *GDF3* protein structure. **(B1)** Before mutation, there are 2 hydrogen bonds (show in the red dotted lines) between the 2 amino acids. **(B2)** After mutation, 1 steric clash occurs between S215 and L105. The blue dotted lines show this relationship. **(C)** **(C1)** Ser212 has 2 hydrophobic bonds with Arg210 and Val1214. **(C2)** After mutation, another hydrophobic bond of Thr 136 occurs in relation with Leu212.

we used Ligplot to analyze and present the hydrophobic interaction changes during the mutation. For the hydrophobic property, Ser and Leu have huge differences, with scores of -0.8 and 3.8, respectively. Ser212 has 2 hydrophobic bonds with Arg210 and Val1214. After mutation, there is another hydrophobic bond of Thr 136 in relation with Leu212 (Figure 4C1, 4C2).

Discussion

To date, studies had found some *GDF3* variants appeared in patients with congenital abnormality. A previous study reported that patients who carry *GDF3* missense variants display a spectrum of oculo-skeletal disorders [9]. One *GDF3* variant, R266C, had been demonstrated to have pathogenic functions at the molecular level. Another study found that *GDF3* S212L mutation may be related to non-syndromic congenital heart diseases (CHDs). However, they did not perform further functional or *in silico* assessment of this novel mutation.

Interestingly, we found 5 novel mutations by use of WES data in our congenital scoliosis cohort. However, the pathogenicity of these 5 *GDF3* variants have not been previously elucidated. Therefore, the aim of the present study was to evaluate the pathogenicity of 5 different *GDF3* gene variants at the molecular level and through *in silico* assessment.

R84C and R84L variants locate on the first exon of *GDF3* gene. Although the 2 mutations occur at the same site, they can result in different amino acid changes. Through SIFT and PolyPhen-2 prediction, the 2 variants were evaluated as pathogenic mutations. Functional assays showed that R84L mutation can cause inhibitory effects on downstream SOX9 reporter gene, while R84C has no obvious influence on the activation of SOX9. These results are also consistent with protein expression assays, in which R84L was shown to interfere with the maturation of GDF3 in cytoplasm, while R84C has little effect on GDF3 protein. This may contribute to the amino acid change of arginine to cysteine. The cysteine may take advantage of the property of hydrophobic to have potential effects on the formation of surrounding amino acid folding, further interfering with the shearing of propeptide. Accordingly, our study suggests that R84L has a functionally pathogenic effect, while R84C may be but was predicted as probably damaging.

In our cohort, GDF3 S212L mutations were sequenced in 9 patients whose phenotypes were congenital scoliosis with lateral curvature on thoracic lumbar segment (T3–L5). However, through *in silico* assessment, the mutation was predicted as benign. Due to its recurrence in our specific cohort, we needed to perform functional assays to determine its true pathogenicity. The results of both luciferase reporter gene and protein expression assays demonstrated that S212L has functional pathogenicity either in downstream activation or in protein maturation. We used Ligplot software to simulate the hydrogen bonds near S212L *loci*. The conformational transformation demonstrates that S212L may change the number of local hydrogen bonds. However, an additional hydrophobic bond is introduced between S212 and Thr 136, which may enhance the local van der Waals force to interfere with the maturation of the protein. This result agrees with the presentation of protein expression. S212L results in amino acid change from serine to leucine, which may interfere with the shear of propeptide in the cytoplasm and the supernatant, thus further affecting the maturity of GDF3 protein. Accordingly, a preliminary judgment can be made that S212L change is a functionally pathogenic locus.

The fourth N215S mutation locates near the previous S212L variant. Interestingly, N215S does not significantly interfere with SOX9 downstream activation, in contrast with the negative effects on the maturation of secreted protein. N215S variant is able to interfere with the GDF3 protein maturation

in both cytoplasm and supernatant. Then, we used the UCSF Chimera software to predict the change in space conformation, showing that GDF3 215th serine may conflict with the 105th isoleucine. The conformation conflicts could cause the instability of the structure of the local protein. The fifth mutation, A251T, locates in the TGF- β domain. Molecular biological experiments show that GDF3 251th may cause functional changes in the maturation of GDF3-secreted protein as well as the activation in downstream elements. Considering that the 3 variants (S212L, N215S, and A251T) were predicted as benign *loci* by SIFT and Polyphen-2, while molecular assays demonstrate that these 3 sites have functional effects on downstream activation or protein maturation, these variants can be classified as functional polymorphisms or potential pathogenicity mutations.

To ensure the objectivity of our assays, we introduced 2 variants as negative reference control and positive reference control. G132W, which had been screened in our control group, was introduced as a negative reference control in the present study. No spinal deformities or other malformations have been shown in the control subjects who carry this mutation. In line with results of our simultaneous assays, G132W mutation had no changes in transcriptional activation or protein expression. We also introduced R266C variant as a positive reference control. R266C has been reported to be a functional pathogenic variant [9]. Patients who carry R266C variant display diverse phenotypes of ocular and skeletal anomalies. Consistent with a previous study, our assays also confirm the pathogenicity of R266C variant in both luciferase reporter gene and protein expression assays.

Although we tested the pathogenicity of these variants by functional assays and *in silico* assessment, the present study has some limitations. First, due to lack of full pedigree for each proband, the inheritance of every patient was not clear, although the variants enrolled in our study all belong to the heterozygote model. Whether those variants are *de novo* mutations is not clear based on the present study. Secondly, 3 variants were predicted as benign, while functional assays showed them as being pathogenic. Further evidence is needed to clarify these conflicting observations. For example, further work is needed to test and verify these variants in a larger sample to find whether the 3 variants are polymorphisms. Also, comprehensive studies are needed to verify the pathogenicity in the signal transduction setting. Some polymorphisms may be defined as functional through further mechanism research. Thirdly, we expect that the actual structure of GDF3 protein will be defined, which could facilitate understanding of the conformational change for each variant. Moreover, *in vivo* assays, such as site mutagenesis by *crispr/cas9* in a mouse model, are needed.

Conclusions

The present study evaluated the pathogenicity of 5 variants through functional and *in silico* assessment. One variant, R84L, was proved to be pathogenic mutation from reporter gene and protein expression assays. R84C may be functionally benign, although it was predicted as probably damaging. The other 3 variants, S212L, N215S, A251T, may be functional

polymorphisms and need further studies to expand this definition. The determination of these variants' pathogenicity in the present study will provide molecular evidence to assist interpretation of high-throughput data for clinical genetic testing.

Conflicts of interests

None.

References

1. Hedequist D, Emans J: Congenital scoliosis: A review and update. *J Pediatr Orthop*, 2007; 27(1): 106–16
2. Kose N, Campbell RM: Congenital scoliosis. *Med Sci Monit*, 2004; 10(5): RA104–10
3. Bulman MP, Kusumi K, Frayling TM et al: Mutations in the human delta homologue, *DLL3*, cause axial skeletal defects in spondylocostal dysostosis. *Nat Genet*, 2000; 24(4): 438–41
4. Giampietro PF, Raggio CL, Blank RD et al: Clinical, genetic and environmental factors associated with congenital vertebral malformations. *Mol Syndromol*, 2013; 4(1–2): 94–105
5. Jones CM, Simon-Chazottes D, Guenet JL, Hogan BL: Isolation of *Vgr-2*, a novel member of the transforming growth factor-beta-related gene family. *Mol Endocrinol*, 1992; 6(11): 1961–68
6. Levine AJ, Brivanlou AH: *GDF3*, a BMP inhibitor, regulates cell fate in stem cells and early embryos. *Development*, 2006; 133(2): 209–16
7. Wu X, Shi W, Cao X: Multiplicity of BMP signaling in skeletal development. *Ann NY Acad Sci*, 2007; 1116: 29–49
8. Salazar VS, Gamer LW, Rosen V: BMP signalling in skeletal development, disease and repair. *Nat Rev Endocrinol*, 2016; 12(4): 203–21
9. Ye M, Berry-Wynne KM, Asai-Coakwell M et al: Mutation of the bone morphogenetic protein *GDF3* causes ocular and skeletal anomalies. *Hum Mol Genet*, 2010; 19(2): 287–98
10. Hensinger RN: Congenital scoliosis: etiology and associations. *Spine (Phila Pa 1976)*, 2009; 34(17): 1745–50
11. Xue X, Shen J, Zhang J et al: Klippel-Feil syndrome in congenital scoliosis. *Spine (Phila Pa 1976)*, 2014; 39(23): E1353–58
12. Bahassi EM, Stambrook PJ: Next-generation sequencing technologies: Breaking the sound barrier of human genetics. *Mutagenesis*, 2014; 29(5): 303–10
13. Guidugli L, Shimelis H, Masica DL et al: Assessment of the clinical relevance of *BRCA2* missense variants by functional and computational approaches. *Am J Hum Genet*, 2018 [Epub ahead of print]
14. Bravo-Alonso I, Navarrete R, Arribas-Carreira L et al: Nonketotic hyperglycinemia: Functional assessment of missense variants in *GLDC* to understand phenotypes of the disease. *Hum Mutat*, 2017; 38(6): 678–91
15. Park HK, Kim MC, Kim SM, Jo DJ: Assessment of two missense polymorphisms (rs4762 and rs699) of the angiotensinogen gene and stroke. *Exp Ther Med*, 2013; 5(1): 343–49
16. Richards S, Aziz N, Bale S et al: Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*, 2015; 17(5): 405–24
17. Thompson JD, Gibson TJ, Higgins DG: Multiple sequence alignment using ClustalW and ClustalX. *Curr Protoc Bioinformatics*, 2002; Chapter 2: Unit 2.3
18. Schwede T, Kopp J, Guex N, Peitsch MC: SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Res*, 2003; 31(13): 3381–85
19. Guex N, Peitsch MC: SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis*, 1997; 18(15): 2714–23
20. Pires DE, Ascher DB, Blundell TL: DUET: A server for predicting effects of mutations on protein stability using an integrated computational approach. *Nucleic Acids Res*, 2014; 42(Web Server issue): W314–19
21. Pettersen EF, Goddard TD, Huang CC et al: UCSF Chimera – a visualization system for exploratory research and analysis. *J Comput Chem*, 2004; 25(13): 1605–12
22. Wallace AC, Laskowski RA, Thornton JM: LIGPLOT: A program to generate schematic diagrams of protein-ligand interactions. *Protein Eng*, 1995; 8(2): 127–34
23. Asai-Coakwell M, French CR, Ye M et al: Incomplete penetrance and phenotypic variability characterize *Gdf6*-attributable oculo-skeletal phenotypes. *Hum Mol Genet*, 2009; 18(6): 1110–21