A novel missense COL10A1 mutation identified by next generation sequencing in a Chinese pedigree with Schmid metaphyseal chondrodysplasia

Qiong Chen1, Sheng-Nan Wu1, Yong-Xing Chen1, Lu Zhang2, Hai-Yan Wei1, Senthil Arun Kumar1.
1Department of Endocrinology, Genetics, Metabolism, Children's Hospital Affiliated to Zhengzhou University (Henan Children's Hospital, Zhengzhou Children's Hospital), Zhengzhou 450053, China
2Shanghai WeHealth Biomedical Technology Co. Ltd, Shanghai 201315, China

INTRODUCTION

Schmid metaphyseal chondrodysplasia (SMCD; OMIM #156500) is an autosomal dominant hereditary chondrodysplasia inflicted by the heterozygous mutations in the COL10A1 gene located at the chromosome 6q21-6q22.3 (1, 2). The clinical manifestations of SMCD are noticeable only after the second year of disease progression characterized by short-limbed dwarfism, bowed legs, and waddling gait (2). Clinically, SMCD diagnosis has been carried out using the radiography and genome analysis of the COL10A1 mutation (2-4). The radiography analysis includes irregular acetabular roofs, enlarged capital femoral epiphyses, coxa vara, genu varum, metaphyseal irregularities with fraying and splaying in the knee, ankle, and wrist (3, 5). COL10A1 gene is comprised of 3 exons that encodes the α1(X) chains of type X collagen, which is the short-chain collagen that regulates the hypertrophic chondrocytes of growth plate cartilage (6, 7). Type X collagen is a homotrimeric molecule made of three α1(X) chains with each chain possessed two noncollagenous globular domains at both the amino-terminal (NC2) and carboxyl-terminal (NC1) ends (8, 9). With the NC1 domain being highly prone to get infused with the deleterious mutations (1, 10), to date, there are 54 COL10A1 mutations have been recorded according to the Human Gene Mutation Database (http://www.hgmd.cf.ac.uk/ac/all.php) within which the missense mutation, nonsense, and frameshift mutations are noticeable among the SCMD affected patients (1, 3, 11). Irrespective of these earlier recorded gene mutations, our genome analysis on the SMCD affected-female child and her younger male sibling has confirmed the inheritance of a novel missense gene mutation, c.2020G>A with the heterozygous substitution of glycine replaced by arginine in the 674th codon of COL10A1 from the SMCD affected mother. Concomitantly, the modified structural characteristics of the α(X1) chain of collagen X mediated by this novel mutation, along with the other deleterious clinical symptoms have been recorded in this study to facilitate the clinicians for a better understanding of the SMCD disease progression and for its accurate prognosis.

Background of the SMCD affected child and the follow-up procedures

The child was born at 40 weeks of gestation to the nonconsanguineous parents whose mother and her younger male sibling showed similar types of phenotypic characteristics such as short-limbed dwarfism, bowed legs, waddling gait and genu varum alike the female child patient. A case report of her other close family members showed no adverse clinical symptoms of SMCD. Earlier, the patient was misdiagnosed and treated for the vitamin D deficiency while she was a year and 8 months old at the local hospital in China which may be due to the waddling gait characteristics of SMCD mimicking the clinical symptoms of a vitamin D-deficient child. Since the child didn't show any sign of improvement in her clinical symptoms even upon her treatment for the vitamin D deficiency for two years. She has been further examined by our physicians at the Genetics Out-patient Department of Children's Hospital Affiliated to Zhengzhou University within which the laboratory tests, and as well as the radiographic examination and the genome analysis have confirmed the incidence of SMCD in the female child patient. Concomitantly, these clinical tests were performed on the patient's biological mother and her younger male sibling aged 1 year and 5 months old to confirm the inheritance of SMCD in the affected child. The study was approved by the Ethics Committee of Children's Hospital Affiliated to Zhengzhou University. Informed consent was obtained from the patient's parents.

Gene polymorphism analysis by the next generation gene sequencing

DNA was extracted from the peripheral blood collected from the patient as well as from the patient's family using the QIAamp DNA Blood Midi kit (Qiagen, Hilden, Germany). Exome enrichment was performed using the Seq Cap EZ Exome Probes v3.0 (Roche, Switzerland). Polymorphic variants that showed more than 0.5% allele frequency in general populations were filtered out (based on the database of 1000 genomes sequencing project: http://www.1000genomes.org/). The SIFT, PolyPhen, MutationTaster and M-CAP were used to assist in predicting the functional impact of identified missense variants.

Oligonucleotides flanking the genomic locations of identified variants were designed using the Primer3Plus browser. Polymerase chain reactions (PCR) was performed with the following primers: 5'-TTGTTAGTGCCAACCAGGGG-3' and 5'- CTGCTCACTTTTCAGGGGGA-3'. PCR products were sequenced bidirectionally using Big Dye Terminator chemistry v3.1 and sequenced using an ABI 3730XL sequencer (Applied Biosystems/Life Technologies, Carlsbad, CA, USA). Sequences were reviewed manually and using Mutation Surveyor (SoftGenetics, State College, PA) and compared to the wild type reference sequences of COL10A1, NM_000493.

Evaluation of the effect of mutant-collagen X protein structure compared with its wild type in an in-silico setup

Collagen X protein sequence was downloaded from Uniprot Database (12) id (Q03692) which basically exist in trimeric form. Structure based search from the Protein Data Bank (13) listed only the monomeric structure instead of the trimer. Henceforth, to generate the trimer, we have opted the SWISSMODEL server (14) with which the wild type collagen X protein was considered as the query sequence. Of the listed templates, the 3D structure with better query coverage and amino acid identity was considered as the template and the 3D model was generated. Furthermore, mutant collagen protein was generated wherein the Gly674 was replaced with positively charged Arg674 in the sequence and was further considered for 3D modelling. Both the wild and the mutant structures were validated using PROSA (15) and Ramachandran Plot (16). To understand the effect of mutation on protein flexibility, we have chosen the CABS-flex 2.0 online server (17) that has been widely acknowledged for its efficacy as a tool to study the fast simulations of protein structure flexibility. The monomeric subunit of the modelled wild and the mutant collagen proteins were submitted while keeping the parameters default. The Root Mean Square Plot (RMSF) from the result files were considered for further analysis.

The partial picture of the article

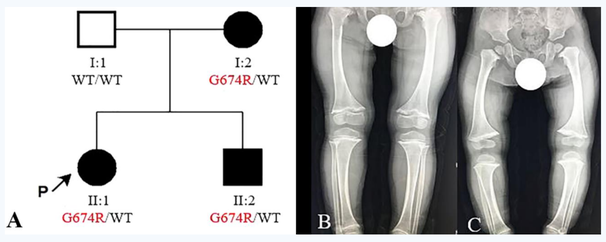

Figure 1: Familial tree and clinical characteristics of the Schmid Metaphyseal Chondrodysplasia (SMCD)-affected family: A, Pedigree tree of the SMCD-affected family : Symbols colored black denotes the SMCD-affected individuals; P, denotes the SMCD-affected female child patient; I:1 and I:2 implies the biological father and mother of the children; II.1 and II.2 implies the daughter and the son of the parents. B and C, Radiography picture of the SMCD-affected female child patient (B) and the younger male sibling (C).

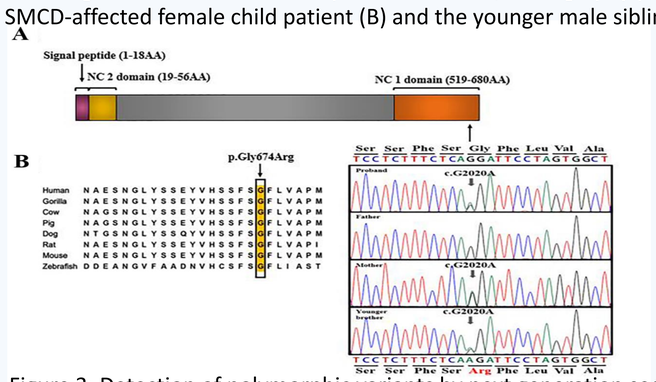

Figure 3: Detection of polymorphic variants by next generation sequencing: A, a novel missense mutation, c.2020G>A; p.Gly674Arg detected at the NC1 domain loci of the COL10A1 gene on the Schmid Metaphyseal Chondrodysplasia (SMCD)-affected female child (proband) followed by the healthy father, and the SMCD-affected mother and the younger brother. B, A diagrammatic representation of the highly conserved glycine residue in the 674th codon of the COL10A1 gene of higher mammals shown in the following chronological order of hierarchy (from top to bottom): Humans; Gorilla; Cow; Pig; Dog; Rat; Mouse, and Zebrafish.

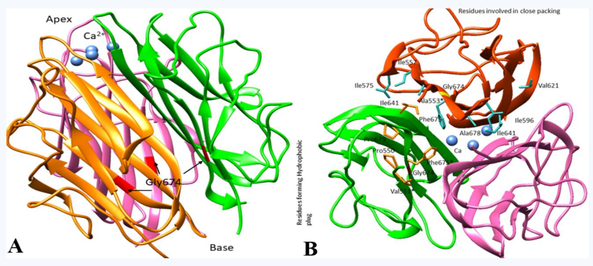

Figure 5: Structural characteristics of the NC1 trimer and the orientation of the amino acid residues engaged in the hydrophobic plug formation and the intact trimer packing: A, a modelled NC1 trimer with highlighted Apex and Base region having intact Ca2+ cluster at the apex. B, chain (marked in green) with the hydrophobic plug forming residues at the base of the monomeric subunit. Phe675 being an active member of the hydrophobic plug is proximal to Gly674 (marked by Asterix * symbol). Residues Val551 and Ala553 are highlighted by Asterix * symbol which flanks Ser552 and Phe554 and forms hydrogen bonds in the wild and mutant protein.

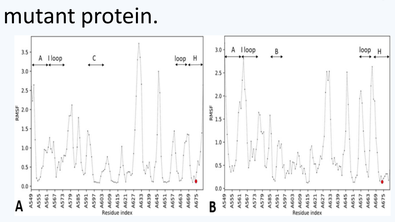

Figure 7: Determination of structural rigidity of the mutant NC1 domain of the α1(X) chain monomer compared with its wild type by the Root Mean Square Fluctuation (RMSF)

Clinical background and the follow-up examination of the patient and other family inmates

During the early stages of examination, the patient was diagnosed with short-limbed dwarfism, bowed legs, waddling gait, and genu varum. The patient was recorded with an average height measured about 85.2 cm (-3.7 SD) at the age of 3 years and 8 months old. Blood-serum level measurements showed higher concentrations of 25-hydroxyvitamin D (>70ng/ml) which may be due to vitamin D treatment upon the misdiagnosis of rickets over SCMD at the local hospital. Further, serum level measurements of calcium, phosphorous, alkaline phosphatase, calcitonin, and parathyroid hormone didn't show any sign of abnormalities on the patient (data not shown). The radiographic picture of the lower limbs confirming the bowed legs in the patient has been shown in Figure 1B. A follow-up examination of these clinical symptoms associated with the phenotypic characteristics of SCMD was carried out on the patient's younger sibling (aged 1 year and 5 months old) as well as her mother within which both showed the marked symptoms of SCMD associated with the short-limbed dwarfism, bowed legs, waddling gait, and genu varum alike the patient. The radiographic picture of the patient's younger sibling has been shown in Figure 1C. However, the younger male sibling of the patient showed less severity in the clinical symptoms of SCMD with the average height measured about 72.8 cm (-3.1 SD), and with the normal 25-hydroxyvitamin D levels. The prepared growth chart of the patient and her younger brother monitored over a period of 10 months has been shown in Figure 2A and Figure 2B. With the patient's mother who showed abnormal phenotypic characteristics alike her daughter was recorded with an average height ranging about 145 cm (-2.9 SD) compared to her normal healthy partner (the patient's father) with an average height measured about 170 cm (-0.44 SD). The family tree has been shown in Figure 1A. Furthermore, genome analysis of the inflicted mutations associated with the SCMD on the patient and the affected immediate family members has confirmed the incidence of SCMD. Alternatively, compared with the other earlier recorded clinical evidences of SCMD, we found both the patient and her younger sibling showed obvious skeletal lesions during the early stages of the development of SCMD (1, 5, 18).

Polymorphic variants detected on the SMCD affected family by the next-generation gene sequencing

Next generation sequencing was performed on the peripheral blood DNA samples isolated from the SCMD affected female child patient, and as well as from the patient's younger male sibling and their biological mother. The detected polymorphic gene variants were prioritized and selected according to ACMG (American College of Medical Genetics and Genomics) guidelines. The diagrammatic representation of the next-generation sequencing results of the patient has been shown in Figure 3A[I1] . With the careful evaluation of the coding exons with its flanked introns by the PCR-direct sequencing, we spotted a novel missense mutation, c.2020G>A accompanied by the heterozygous amino acid substitution, p.Gly674Arg located in the NC1 domain of α1(X) chain (Figure 3A). This mutation hasn't been reported in any of the recommended genomic databases generated by the dbSNP such as the 1000 Genome Project, the Exome Variant Server, and the NHLBI (National Heart, Lung and Blood Institute) Exome Sequencing Project. Based on these evaluation-based scores created by Polyphen-2 (score at 1.00), SIFT (score at 0.006), Mutation Taster (score at 1), and M-CAP (0.317), we speculate that this novel missense mutation c.2020G>A; p.Glyc674Arg at the NC1 domain could be deleterious as the glycine amino acid residue at 674 positions is the highly conserved residue since the evolution of organisms begins from Zebrafish to humans (Figure 3B) (8-10). With the familial genome analysis performed using the same methodology and genomic tools discussed above, our study has confirmed inheritance of the novel mutation, c.2020G>A; p.Gly674Arg in the patient's younger male sibling from the SMCD affected mother. Furthermore, based on the comparison of the SMCD affected female child, including the patient's male sibling and the mother with the healthy-control father, we propose the novel missense mutation, c.2020G>A; p.Glyc674Arg in the COL10A1 gene could be detrimental to the patients.

Structural alterations of the α1(X) chain of collagen X

During the template search in SWISMODEL server, wild type scored a query coverage of 24 percent (residues 521-680) with a 100 percent amino acid identity with the human collagen X NC1 trimer (PDB ID: 1GR3) (19). This template was considered for the generation of collagen X NC1 trimer (Figure 4). Regarding mutant, the query coverage remained the same, nevertheless, the identity was 99.38 percentage after the substitution of glycine with arginine. Structure validation through PROSA online server confirms a better model quality with the value below zero for the window size of 40 (Figure 5A and 5B). As per the Ramachandran plot analysis, 94.6% residues are in most favoured regions; 4.5% in additional allowed regions and 0.9% in generously allowed regions (Figure 6A and 6B). There are no residues within the disallowed region, thus, confirming it to be a best modelled structure. Similar report was observed for the mutant.

The modelled trimeric wild and the mutant collagen proteins were considered for hydrogen bond analysis using the Discovery Studio 2019 client (9). In wild type protein, Gly674 from β strand H interacts with Phe554 (β strand A) and Tyr598 (β strand C). While in the mutant protein, the Arg674 interacts with Ser552, Phe554 (β strand A) and Phe589 (β strand B). Following which an overall shift in hydrogen bonding pattern was observed between the wild and the mutant (Figure 6A and 6B). Concomitantly, the RMSF plot of the wild and the mutant proteins were analysed. It was observed that in wild type, Gly674 brings rigidity to the loop I which is following strand A and also the loop preceding the strand H. Furthermore, the strand H is rigid in wild type that contradicts its flexible stature in the mutant type (Figure 7A and 7B).

Disscusion

The study has reported a novel missense mutation c.2020G>A with the heterozygous substitution of glycine with arginine, p.Gly674Arg in the NC1 domain of the α1(X) chain of collagen X (Figure 2B)[I1] ; that triggered the incidence of SMCD in the female child patient. Also, the inheritance of this novel COL10A1 gene mutation from the SMCD affected mother (as a carrier) has been confirmed in the patient's younger male sibling (Figure1A), who were diagnosed with clinical symptoms like short-limbed dwarfism, bowed legs, waddling gait, and genu varum. The missense mutation inflicted by the glycine replacement with arginine in the 674th codon of COL10A1 gene encoding the NC1 domain of the α1(X) chain could intervene the collagen X structure and function by attenuating its trimer formation initiated by the three α1(X) chains (8-10). Earlier evidence has shown the residues Ala553, Ile557, Leu575, val621, Ile641, Phe675 and Ala678 to be directly associated with the intact packing of the α1 (X) chains-trimer supported by the Pro550, Val551, Phe675 and Pro679 residues that forms the hydrophobic plug at the base of the NC1 trimer generating a firm collagen X protein structure (20). Having confirmed the stable interaction of both the Gly674 (wild type) and Arg674 (mutant type) amino acid residues with the Phe554 residue, which is proximal to the Ala553 residue (Figure 6A and 6B); we explored an extended interaction between the Arg674 with the Ser552 residue (Figure 6B). Also, the residue Phe675 which is crucial for the compact trimer packing as well as for the hydrophobic plug formation turns out to be a flexible one in the mutant α1(X) chains of the collagen X protein compared with its wild type (Figure 5A, 5B and 7A, 7B). These structural changes inflicted by the missense p.Gly674Arg mutation has indeed affected the overall rigidity of the α1(X) chains interrupting its assembly as a trimer to generate a compact collagen X structure. There could be a direct correlation between the heterozygous p.Gly674Arg mutation and the adverse clinical symptoms like the bowed legs in the SCMD affected patients preceded by the effaced collagen X structure and function (9, 10, 21).

Since the SMCD clinical symptoms overlap with the other similar clinical symptoms of achondroplasia and rickets (18, 22, 23), we recommend the differential diagnosis of SMCD to all the clinicians to prevent any misdiagnosis. The differential diagnosis of SMCD can be achieved by performing the diagnosis at various levels on the patients. At the first level, a thorough examination of the patient's clinical symptoms, including their family members, must be carried out diligently. The second level of radiography examination must be performed cautiously to differentiate the patients with rickets showing similar phenotypic characteristics like the SMCD affected child patients.[I2] (Figure 1B and 1C). At the third-final level of examination, the polymorphism analysis must be performed diligently to categorize the SMCD affected patients over the other patients with achondroplasia and rickets-based on the inheritance of the polymorphic gene variants. Despite the noticeable improvements eventually seen in the abnormal body skeleton structure corrected by the surgical interventions in the SMCD patients, a majority of the patients still retain their short stature with coxa vara (21, 24). Surgical intervention is employed to rectify the symptomatic deformity detected in the SMCD affected patients (21).

In conclusion, our study is the first to report the inheritance of SMCD in the Chinese family inflicted by the novel missense mutation, c.2020G>A; p.Gly674Arg, affecting the structural rigidity of the α1(X) chains and its trimer formation disrupting the collagen X protein structure and function. We have also found a direct correlation between the disturbed genotypic and phenotypic characteristics in the SMCD affected-Chinese family followed with the proposed diagnostic module for early diagnosis and treatment of SMCD symptoms.