



Whole-exome sequencing reveals a novel *COL2A1* mutation in a patient with spondylo-epiphyseal dysplasia congenita

A. Sangsin^{1,2,3,4}, C. Srichomthong^{3,4}, M. Pongpanich^{5,6}, K. Suphapeetiporn^{3,4} and V. Shotelersuk^{3,4}

¹Department of Orthopedics, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand

²Interdepartment Program of Biomedical Sciences, Faculty of Graduate School, Chulalongkorn University, Bangkok, Thailand

³Center of Excellence for Medical Genetics, Department of Pediatrics, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand

⁴Excellence Center for Medical Genetics, King Chulalongkorn Memorial Hospital, the Thai Red Cross Society, Bangkok, Thailand

⁵Department of Mathematics and Computer Science, Faculty of Science, Chulalongkorn University, Bangkok, Thailand

⁶Program in Bioinformatics and Computational Biology, Graduate School, Chulalongkorn University, Bangkok, Thailand

Corresponding author: K. Suphapeetiporn

E-mail: kanya.su@chula.ac.th

Genet. Mol. Res. 15 (1): gmr.15017624

Received September 10, 2015

Accepted November 17, 2015

Published March 11, 2016

DOI <http://dx.doi.org/10.4238/gmr.15017624>

ABSTRACT. Skeletal dysplasia is a group of disorders with more than 450 entities, many of which cannot be differentiated, especially during infancy, but could lead to different clinical courses and prognoses. In this study, we have described a case of a Thai infant with short stature, flat face, pectus carinatum, indirect inguinal hernia, platyspondyly, and generalized delayed endochondral ossification. Using whole-exome sequencing (WES), we successfully identified a *de novo* heterozygous mutation,

c.2024G>A (p.G675D), in the *COL2A1* gene, which, to our knowledge, has not been previously reported. These molecular findings helped provide a definite diagnosis of spondyloepiphyseal dysplasia congenita, aiding in proper management of the disease and improved genetic counseling. We demonstrated that WES is an efficient and cost-effective tool for molecular diagnosis for a type II collagenopathy.

Key words: *COL2A1*; Spondyloepiphyseal dysplasia congenita; Whole-exome sequencing

INTRODUCTION

Skeletal dysplasia is a term given to a group of diseases composed of over 450 entities (Sewell et al., 2015). Clinicians generally identify patients with flat facial profiles, a short stature, platyspondyly, arthritis at a young age, and ocular abnormalities as being afflicted with type II collagenopathies, a group of skeletal dysplasia caused by heterozygous mutations in *COL2A1* (Jobling et al., 2014). However, patients with mutations in the same gene could have very different clinical courses and prognoses.

Type II collagenopathies range from *in utero* (lethal) disorders, including achondrogenesis type II or hypochondrogenesis (ACG-II/HCG; OMIM 200610) and Torrance-type platyspondylic dysplasia (OMIM 151210), to adult disorders, such as avascular necrosis of the femoral head (OMIM 608805). Neonatal or infant type II collagenopathies include spondyloepiphyseal dysplasia congenita (SEDC; OMIM 183900), spondyloepimetaphyseal dysplasia, Strudwick-type (OMIM 184250), spondyloperipheral dysplasia (OMIM 200610), Kniest dysplasia (OMIM 156550), and Stickler syndrome type I (STD-1; OMIM 108300) (Kannu et al., 2012).

All changes in the *COL2A1* gene lead to retardation of endochondral ossification. Truncating mutations cause mild skeletal alterations, but severe extraskelatal manifestations (ocular abnormalities and hearing loss), in STD-I. Splice site mutations leading to exon skipping result in an intermediate phenotype in Kniest dysplasia. Interestingly, missense mutations in the triple-helical glycine residue lead to severe skeletal, but mild extraskelatal, alterations in spondyloepiphyseal dysplasia (SED) (Nishimura et al., 2005). SED includes lethal SED (ACG-II and HCG), SEDC, and late-onset SED (Unger et al., 2001; Jung et al., 2004; Nishimura et al., 2005; Meredith et al., 2007; Xia et al., 2007; Mark et al., 2011). Identification of the underlying mutations could help in the clinical diagnosis of this disease, and in providing accurate information to the patients and their families.

The *COL2A1* gene has 54 exons. The use of conventional mutation detection methods, including polymerase chain reaction (PCR) followed by Sanger sequencing of these exons, would be time-consuming and prohibitively expensive. In this study, we show that whole-exome sequencing (WES), a relatively new and efficient genetic technique, could be used to successfully identify a novel mutation in *COL2A1* in a patient with SEDC.

MATERIAL AND METHODS

Study subject

A Thai infant with skeletal dysplasia who was brought to the Genetics Clinic of King Chulalongkorn Memorial Hospital was studied. The medical data, pedigree, physical examina-

tion data, and laboratory results of the patient were recorded. The possible consequences of this study were thoroughly explained to the caregivers of the infant, and informed consent was obtained from them. Ethical approval was obtained from Chulalongkorn University Ethics Committee for the Faculty of Medicine.

Genomic DNA preparation and WES

Genomic DNA was isolated from peripheral blood leukocytes obtained from the infant, using Puregene Blood kit (Qiagen, Hilden, Germany). WES of the genomic DNA was performed by Macrogen Inc. (Seoul, South Korea). DNA was captured on the SureSelect Human All Exon kit (v.4; Agilent Technologies, Santa Clara, CA, USA) and sequenced on an Agilent HiSeq2000 sequencer. Base calling was performed and the quality scores determined using the Real Time Analysis software (v.1.7; Illumina, San Diego, CA, USA). Sequence reads were aligned against UCSC hg19 using the Burrows-Wheeler Alignment software (<http://bio-bwa.sourceforge.net/>). Single nucleotide variants (SNVs) and insertion/deletions (InDels) were detected using SAMTOOLS (<http://samtools.sourceforge.net/>) and annotated by dbSNP&1000G. After quality filtering, we attempted to identify variants located in the coding regions of previously reported skeletal dysplasia genes for all potential pathogenic SNVs and InDels. Variant calling exclusion criteria included a coverage <10X, quality score <20, minor allele frequency ³1% in 1000 Genomes Project, and non-coding variants and synonymous exonic variants. The remaining variants were subsequently filtered using an in-house database of 165 unrelated Thai exomes. The variants were confirmed by PCR and Sanger sequencing.

Existing SNVs or known pathogenic mutations were filtered out using the Human Gene Mutation Database (HGMD; <http://www.hgmd.cf.ac.uk/ac/index.php>) and Exome Aggregation Consortium (ExAC; <http://exac.broadinstitute.org>). Deleterious SNVs were predicted by Polyphen-2 (<http://genetics.bwh.harvard.edu/pph2/>) and SIFT (<http://sift.jcvi.org>). Evolutionarily conserved sequences among species were analyzed using Clustal X v.2 (<http://clustal.org/clustal2/>).

Case presentation

The patient is the first born son of a non-consanguineous healthy couple. At term, the infant was delivered by cesarean section because of breech presentation; the birth weight of the boy was 2640 g. At 9 months of age, the infant was referred to our genetic clinic for evaluation of his dysmorphic features. Physical examination revealed a short stature (51 cm), flattened facies, presence of pectus carinatum, rhizomelia, limited range of motion in both elbows, and reducible left indirect inguinal hernia (Figure 1A). Radiographs showed generalized delay in ossification. The anteroposterior view showed a broad and bell-shaped thorax (Figure 1B) and a lateral view of the anterior ends of the ribs showed flaring (Figure 1C). The vertebral bodies were flattened (Figure 1C). Retarded ossification of the pubic bone, with small iliac wings in the cephalocaudal dimension, was observed (Figure 1D). There was no sign of ossification in the femoral heads, necks, and around the knees. Long bones, especially the femurs and humeri, appeared short (Figure 1B and D). The infant's parents and siblings showed no dysmorphic features.



Figure 1. Clinical and radiographic features of the patient. **A.** Photograph of the patient shows flattened facies, pectus carinatum, rhizomelia, and left inguinal hernia. **B.** Chest AP radiograph shows a broad and bell-shaped thorax and short humeri. **C.** Chest lateral radiograph shows anterior end flaring of the ribs and platyspondyly. **D.** Pelvic radiograph shows retarded ossification of the pubic bone, femoral heads, and necks, and a short iliac bone.

RESULTS

WES identified a non-synonymous variant (NM_001844:c.G2024A;p.G675D) in *COL2A1*. PCR and Sanger sequencing of DNA extracted from leukocytes obtained from the infant and his parents revealed that the patient was heterozygous for the mutation, while his parents expressed only the wild-type allele (Figure 2A). Polyphen-2 predicted this variant to be damaging, with a score of 0.999; in addition, SIFT (<http://sift.jcvi.org>) predicted this mutation to be deleterious, with a score of 0.001. Analysis with Clustal X v.2 (<http://clustal.org/clustal2/>) revealed that the glycine residue at amino acid 675 was evolutionarily conserved (Figure 2B) from humans to amphibians. This mutation was absent in the HGMD (<http://www.hgmd.cf.ac.uk/ac/index.php>) and ExAC (<http://exac.broadinstitute.org>) databases, and our in-house exome database comprising exomes obtained from 165 Thai individuals.

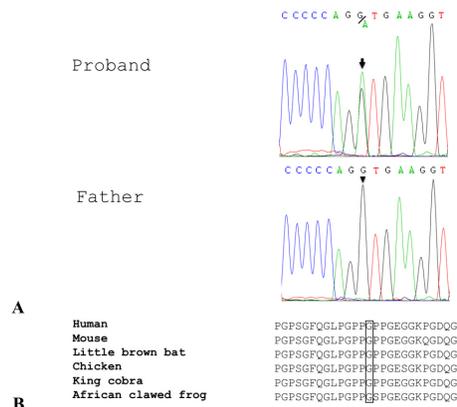


Figure 2. Mutation analysis. **A.** Direct sequencing shows that the proband is heterozygous for a *de novo* c.G2024A (p.G675D) mutation (upper panel), while his father expresses only the wild-type G allele (lower panel). **B.** The mutation observed in this patient is in a region that is highly conserved among many species.

DISCUSSION

Type II collagenopathies caused by mutations in the *COL2A1* gene have a wide range of severity. Identification of the underlying mutation in an infant with a type II collagenopathy would confirm a clinical diagnosis, in addition to providing valuable information for effective management of the disease. Here, we have shown that WES, a relatively new molecular technique, could be used to identify the causative mutation in the *COL2A1* gene in a 9-month-old infant with skeletal dysplasia.

WES, a next-generation sequencing (NGS) method, starts with fragmenting DNA into millions of short fragments, followed by massively parallel sequencing of all exons in all genes in the genome; the output obtained after one run (5 days) is in the range of megabases. It is suitable for the detection of single nucleotide substitutions, insertions, or deletions smaller than 8 to 10 nucleotides (Biesecker and Biesecker, 2014). Compared to the conventional mutation detection technique involving PCR-Sanger sequencing, WES is less time-consuming and less expensive. The cost of whole human genome sequencing by NGS dramatically decreased from hundreds of millions of dollars prior to 2001 to less than ten thousand dollars now (Collins and Hamburg, 2013). The current cost of WES is around \$900. Additionally, WES and whole-genome sequencing methods can be used to uncover new disease genes. Since 2010, 28 novel causative genes of skeletal dysplasia have been identified using NGS (Lazarus et al., 2014). WES also provides researchers with the opportunity to identify incidental actionable pathogenic variants (Collins and Hamburg, 2013). Mutations in *COL2A1*, a gene containing 54 coding exons, cause the spread of SEDC throughout the entire gene without hotspots. Molecular diagnosis using PCR-Sanger sequencing of these 54 exons would be prohibitively expensive and time-consuming. WES, therefore, is the more appropriate choice in this case.

In this study, we identified a previously undescribed heterozygous mutation, c.G2024A in exon 31 of *COL2A1*, theorized to induce an amino acid substitution (p.G675D) in the triple-helical glycine residue. In addition to altering the glycine in the triple-helical glycine residue, evidence has suggested that this may be a causative mutation: 1) analyses of the parental samples revealed that neither carried the mutation, indicating that c.G2024A was a *de novo* mutation; 2) this mutation was also absent in 100 unrelated Thai controls; 3) Clustal X analysis revealed that this mutation is evolutionarily conserved among species (Figure 2B); and 4) various prediction softwares, such as PolyPhen-2 and SIFT, predicted that this mutation could be damaging or deleterious.

COL2A1 regulates the production of components required to construct the α -chain of pro-collagen type II, an important structural protein present in the bones and cartilages. A previous study of the phenotypic spectrums of *COL2A1* mutations in a triple-helical glycine residue showed that glycine to serine substitutions resulted in variable phenotypes, while glycine to non-serine substitutions was responsible for the more severe phenotypes, including HCG and SEDC. Missense mutations in HCG cluster near the N-terminal or C-terminal amino acid boundaries of the triple-helical glycine residue, while mutations in SEDC are spread throughout the exons (Nishimura et al., 2005). So far, 34 *COL2A1* mutations have been reported in SEDC. The mutation observed in our patient occurred at amino acid residue 675, in the middle region of the triple-helical glycine residue, which is consistent with the genotype of SEDC. Moreover, all patients with glycine to non-serine substitutions in the triple-helical glycine residue have been reported to present ocular abnormalities (Nishimura et al., 2005). Therefore, the molecular pathology seen in our patient, p.G675D, a glycine-to-non-serine substitution in the triple-helical glycine residue, led us to highly recommend a thorough eye examination for signs of, and to regularly monitor, myopia, vitreo-retinal degeneration, and retinal detachment, despite ocular abnormalities not being observed in our patient so far.

Despite the identification of some genotype-phenotype correlations, there remain several ambiguities in type II collagenopathies. Significant differences in adult heights were observed in the same family carrying the same mutation, indicating the multifactorial effects of *COL2A1* dysplasia (Terhal et al., 2012). In addition, while the index case in a family with a truncating *COL2A1* mutation presented typical Stickler syndrome, two other family members with the same mutation were adjudged healthy (Faber et al., 2000). Therefore, the identification of genotype-phenotype correlations in type II collagenopathies requires further research with comprehensive clinical and genomic data.

In conclusion, WES was used to identify a *de novo* novel heterozygous missense *COL2A1* mutation (c.G2024A, p.G675D) in an infant with skeletal dysplasia, leading to a diagnosis of SEDC. This emphasizes the usefulness of WES in identifying causative mutations, guiding proper management, and extending the mutational spectrum of type II collagenopathies.

Conflicts of Interest

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

Research supported by the Academic Advancement into Its 2nd Century Project (#RES560530177-HR), and the Thailand Research Fund.

REFERENCES

- Biesecker LG and Biesecker BB (2014). An approach to pediatric exome and genome sequencing. *Curr. Opin. Pediatr.* 26: 639-645. <http://dx.doi.org/10.1097/MOP.0000000000000150>
- Collins FS and Hamburg MA (2013). First FDA authorization for next-generation sequencer. *N. Engl. J. Med.* 369: 2369-2371. <http://dx.doi.org/10.1056/NEJMp1314561>
- Faber J, Winterpacht A, Zabel B, Gnoinski W, et al. (2000). Clinical variability of Stickler syndrome with a *COL2A1* haploinsufficiency mutation: implications for genetic counselling. *J. Med. Genet.* 37: 318-320. <http://dx.doi.org/10.1136/jmg.37.4.318>
- Jobling R, D'Souza R, Baker N, Lara-Corrales I, et al. (2014). The collagenopathies: review of clinical phenotypes and molecular correlations. *Curr. Rheumatol. Rep.* 16: 394. <http://dx.doi.org/10.1007/s11926-013-0394-3>
- Jung SC, Mathew S, Li QW, Lee YJ, et al. (2004). Spondyloepiphyseal dysplasia congenita with absent femoral head. *J. Pediatr. Orthop. B* 13: 63-69.
- Kannu P, Bateman J and Savarirayan R (2012). Clinical phenotypes associated with type II collagen mutations. *J. Paediatr. Child Health* 48: E38-E43. <http://dx.doi.org/10.1111/j.1440-1754.2010.01979.x>
- Lazarus S, Zankl A and Duncan EL (2014). Next-generation sequencing: a frameshift in skeletal dysplasia gene discovery. *Osteoporos. Int.* 25: 407-422. <http://dx.doi.org/10.1007/s00198-013-2443-1>
- Mark PR, Torres-Martinez W, Lachman RS and Weaver DD (2011). Association of a p.Pro786Leu variant in *COL2A1* with mild spondyloepiphyseal dysplasia congenita in a three-generation family. *Am. J. Med. Genet. A.* 155A: 174-179. <http://dx.doi.org/10.1002/ajmg.a.33762>
- Meredith SP, Richards AJ, Bearcroft P, Pouson AV, et al. (2007). Significant ocular findings are a feature of heritable bone dysplasias resulting from defects in type II collagen. *Br. J. Ophthalmol.* 91: 1148-1151. <http://dx.doi.org/10.1136/bjo.2006.112482>
- Nishimura G, Haga N, Kitoh H, Tanaka Y, et al. (2005). The phenotypic spectrum of *COL2A1* mutations. *Hum. Mutat.* 26: 36-43. <http://dx.doi.org/10.1002/humu.20179>
- Sewell MD, Chahal A, Al-Hadiithy N, Blunn GW, et al. (2015). Genetic skeletal dysplasias: a guide to diagnosis and management. *J. Back Musculoskeletal Rehabil.* 28: 575-590. <http://dx.doi.org/10.3233/BMR-140558>
- Terhal PA, van Dommelen P, Le Merrer M, Zankl A, et al. (2012). Mutation-based growth charts for SEDC and other *COL2A1* related dysplasias. *Am. J. Med. Genet. C. Semin. Med. Genet.* 160C: 205-216. <http://dx.doi.org/10.1002/ajmg.c.31332>
- Unger S, Korkko J, Krakow D, Lachman RS, et al. (2001). Double heterozygosity for pseudoachondroplasia and spondyloepiphyseal dysplasia congenita. *Am. J. Med. Genet.* 104: 140-146. <http://dx.doi.org/10.1002/ajmg.10062>
- Xia X, Cui Y, Huang Y, Pan L, et al. (2007). A first familial G504S mutation of *COL2A1* gene results in distinctive spondyloepiphyseal dysplasia congenita. *Clin. Chim. Acta* 382: 148-150. <http://dx.doi.org/10.1016/j.cca.2007.04.005>