



Research Article

Open access

Implication of a *de novo* Variant in ciliary rootlet coiled-coil (CROCC) with assimilation of atlas (AOA)

Huaiyu Tong¹, Chongye Guo², Liang Liang^{2,3}, Hua Mi⁴, Meng Li², Yiheng Yin¹, Lijun Shang^{5*}, Shuangli Mi^{2,3*}, Xinguang Yu^{1*}

¹ Department of Neurosurgery, First Medical Center, General Hospital of Chinese PLA, Beijing, China

² Key Laboratory of Genomic and Precision Medicine, Beijing Institute of Genomics, Chinese Academy of Sciences, China National Center for Bioinformation, Beijing 100101, China

³ University of Chinese Academy of Sciences, Beijing, China

⁴ XuanWu TCM Hospital Beijing, Beijing, China

⁵ School of Human Sciences, London Metropolitan University, London, N7 8DB, UK

DOI: 10.31383/ga.vol6iss1pp11-26

Abstract

Assimilation of atlas is a rare skeletal malformation causing nerve compression with high risk of fatal. However, the genetic etiology of assimilation of atlas AOA is currently lacking. In this paper, the whole-exome sequencing (WES) analysis was employed to study a Chinese family having a sporadic proband son of assimilation of atlas AOA but other healthy family members. We identified a novel variant in ciliary rootlet coiled-coil gene (NM_014675.5 (CROCC): c.4702C>T (r.4702c>u, p.(Arg1568Cys)). The variant had different genotypes between the proband and healthy family members but with high conservations of “damage” to protein structure based on MutationTaster and SIFT prediction. CROCC gene can be obtained in both healthy (n=220) and non-mutated assimilation of atlas AOA patient samples (n=68) but absented in five sporadic patients with the novel variant. Furthermore, abnormal of cilia was observed after editing the target sequence on CROCC using CRISPR-Cas9. These results suggested that assimilation of atlas AOA might be caused by the mutation of CROCC: c.4702C>T (r.4702c>u, p.(Arg1568Cys)). With strong amino acid conservation and interaction regulation, the variant mutation could cause the signal disorder of skeletal development which may lead to the defective bone formation and finally cause the development of assimilation of atlas AOA.

*Correspondence

E-mail: xinguang_yu@263.net;
mishl@big.ac.cn;
l.shang@londonmet.ac.uk

Received

April, 2022

Accepted

June, 2022

Published

June, 2022

Copyright: ©2022 Genetics & Applications, The Official Publication of the Institute for Genetic Engineering and Biotechnology, University of Sarajevo

Keywords

Assimilation of atlas AOA, Whole exome sequencing, Ciliary rootlet coiled-coil (CROCC), Rootletin, Cilia, Hedgehog signaling pathway

Introduction

Assimilation of atlas (AOA, OMIM: 601076) AOA, also known as “occipitalization of the atlas” or “atlas assimilation”, is one of the junction (CVJ) disease characterized with fusion of the atlas and the base of the occipital bone (Abd et al., 2008). AOA was firstly reported in 1911 as one of the rare skeletal deformities with incidence of 0.14% - 3.36% (Al-Motabagani and Surendra, 2006; Allen, 1972; Kassim et al., 2010; Khanmanarong et al., 2013; Kim et al., 2013; Saini et al., 2009). It may show symptomatic or sometimes completely asymptomatic and can incidentally be detected at autopsies or during routine cadaveric dissections (Al-Motabagani and Surendra, 2006; Kassim et al., 2010). Normally, AOA appears solitary onset in clinic with the fusion of atlas with occipital bone, but is also conjunction with the onset of the wide variety of other hemivertebra abnormalities, such as C1-T4 spinal cord malformation (Anderson et al., 2006; Kassim et al., 2010; Nimje and Wankhede, 2014). The mild symptoms include headache, pain in the neck, and abnormal neck posture, restriction of neck movements, dizziness, syncope and weakness (Abd et al., 2008; Campos et al., 2012; Hemamalini, 2014). AOA could also cause compression of human nerve system leading to weak constitution (Hemamalini, 2014; Jayanthi et al., 2003; Kassim et al., 2010). For instance, AOA may cause vertebral artery compression and affect cerebral blood flow (Kassim et al., 2010), and it also lead to compression of the first cervical nerve (Kassim et al., 2010). This abnormality may affect the movement of extensor surface posture muscles, and can even be fatal when the illness aggravates (Hemamalini, 2014; Jayanthi et al., 2003; Kassim et al., 2010). So far, whether it is a unique bone defect or a spectrum of congenital spinal deformity remains unknown.

Due to its rare incidence, asymptomatic onset and the limited techniques available, there are very few studies on AOA. AOA patients normally show asymptomatic onset and severity with time (Al-Motabagani and Surendra, 2006; Campos et al.,

2012; Jayanthi et al., 2003; Kassim et al., 2010; Kim et al., 2013; Saini et al., 2009) and they only seek for treatment until the skeletal malformation starts to constrict the nerves (Al-Motabagani and Surendra, 2006; Campos et al., 2012; Jayanthi et al., 2003; Kassim et al., 2010; Kim et al., 2013; Saini et al., 2009). But some patients on the contrary do not show any nerve system lesions which lead to insufficiency clinical samples for further studies.

Furthermore, some cilia-related proteins could cause severe bone abnormalities suggesting that the role of cilia in multiple organ development is indispensable (Zhang et al., 2017). Additionally, cilia are considered to function as chemo-sensors and mechano-sensors for bone formation and maintenance (Lee and Gleeson, 2011). Bone developmental signaling pathways, including Hedgehog signaling pathway, have been linked to primary cilia and cilia protein, such as intraflagellar transport (Ift) proteins, and their mutations could cause multiple bone disease and cranio-facial and skeletal defects (Mori et al., 2017). However, there is still no effective way to understand this complexity disease.

Although sequencing technique has been successfully applied in studying small sample size of many CVJ diseases, for instance, a homozygous mutation on gene mesenchyme homeobox 1 (MEOX1) was detected and proved to be one of the pathological mutations for type II Klippel-Feil syndrome using the whole-exome sequencing (WES) technique (Bayrakli et al., 2013), there is no such an application of studying the genetic etiology of AOA cases so far. In this study, the WES analysis was employed to study a Chinese family having a sporadic proband son of AOA but other healthy family members. The genetic etiology of AOA cases was consequently explored.

Material and methods

Ethical Compliance and Study participants

This study has the ethics approval (No.2014S010) from the China PLA General hospital. The clinical

investigations were informed to all participants and received written informed consent to allow the publication of this study. A total of 293 participants were selected for this study, including 220 healthy volunteers and 73 AOA patients who have been diagnosed in the outpatient department of the China PLA General hospital. Among the above participants, the AOA patient (son) and his four healthy family members including his parents and two female siblings were particularly studied in this project. The peripheral blood samples of 5 ml were collected from patients' arms using standard hospital operational protocol after they were diagnosed at the China PLA General hospital. All medical reports were produced by professionally the hospital medical staffs. Furthermore, the exome sequencing data were stored at the internal database of Beijing Institute of Genomics and it is available upon requests for research purposes only.

DNA preparation

The peripheral bloods from all participants were collected for DNA isolation. Genomic DNA was isolated using QIAamp DNA Blood Mini Kit (QIAGEN, Cat No.51104) following the manufacturer's instructions. The quantity of DNA was detected by NanoDrop ND-1000 Spectrophotometer (ThermoScientific).

Whole exome sequencing (WES)

Exome enrichment was conducted following the manufacturer's protocol with Illumina HiSeq2500 sequencing technique. The exomes were captured using the Agilent SureSelected (50Mb) QXT Library Prop Kit (Cat No. G9682B). Whole exome capture was carried out on an Illumina HiSeq2500 platform and the indexed DNA libraries were subjected to paired-end 125bp using V4 reagent. DNA of 5µg per sample were submitted. Sequencing reads aligned mapping to the human reference genome assembly (hg19/GRCh37) using

Burrows-Wheeler Alignment tool (BWA) (Abuín et al., 2015), and the quality of raw data was checked using FastQC (Babraham Bioinformatics). Variants were called, filtered, and prioritized according to their pathogenicity scores (>0.95) obtained from the Polyphen-2 web interface and GATK ver3.3 (Figure 1) (Bahe et al., 2005; Doerks et al., 2002; I.V. et al., 2016; Schoichet et al., 1995). Annovar was used for SNV annotation (Yang et al., 2002). Furthermore, variants were detected in the Human Gene Mutation Database (HGMD, <http://data.mch.mcgill.ca/phexdb>), and genes were intensively examined against the genes with implications of skeletal abnormal diseases to detect the candidate variants. Several SNVs project databases have been used during the analysis procedure, including dbSNP build 138 (Sherry et al., 2001), the 1000 Genomes Project (Consortium, 2003), the International HapMap Project (Consortium, 2003) and the Exome Sequencing Project (ESP) (<https://esp.gs.washington.edu/drupal/>).

Sanger sequencing

Primers for PCR and qPCR amplifications of all target variants were designed using the web-based Primer 3 software (<http://biotools.umassmed.edu/bioapps/prim3>). The primer sequences were listed in Supplementary Table 1. A standard PCR protocol was applied. In brief, initial denaturation at 94°C for 5 min, followed by 35 cycles starting at 94°C for 30 sec; denaturation at 58°C for 30 sec, and extension at 72°C for 2 min. The final extension was carried out at 72°C for 5 min. PCR products were diluted 10-fold in double distilled water and submitted for Sanger sequencing on an Applied Biosystems capillary DNA sequencer (Model 3730 XL) without further purification.

Identification of candidate variant

All variants were selected among family subjects, including the proband, his parents, his two sisters

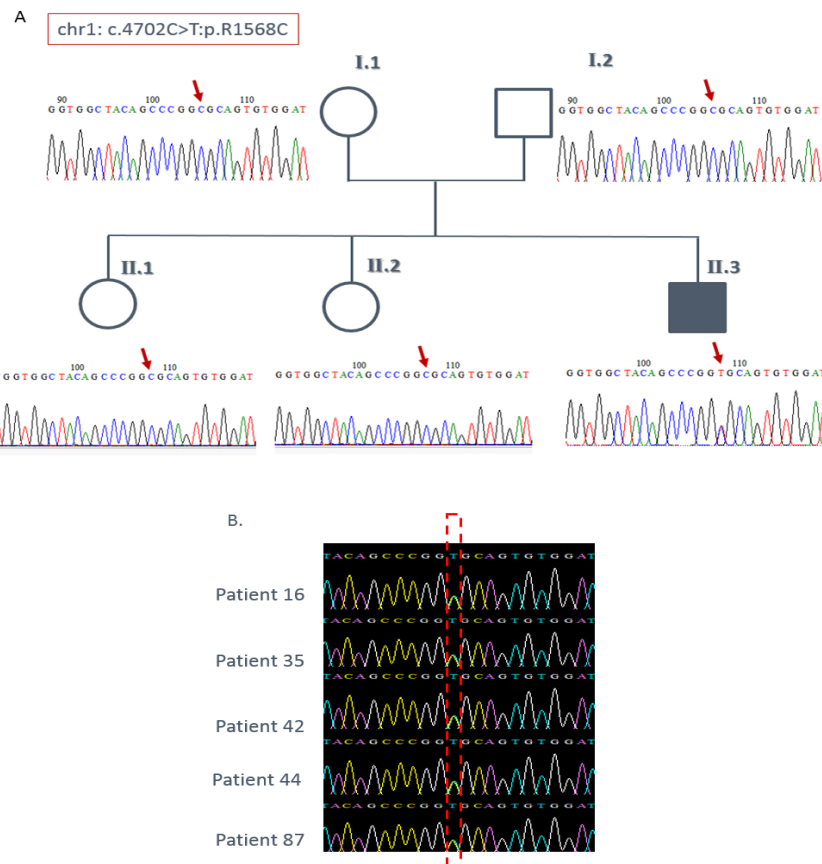


Figure 1. (A) The pedigree chart for the genotype of the proband family with mutation on gene (NM_014675.5 (CROCC): c.4702C>T (r.4702c>u, p.(Arg1568Cys)). The Sanger sequencing results showed the proband with a mutated “T” which was different from the rest of the family. **(B)** Sanger sequencing results identified for the candidate variant in sporadic patients number 16, 35, 42, 44 and 87.

and the health controls. Finally, all the overlapped variants were then mapped to two main databases to check their genotype frequency occur in population. We set the cut-off value for the frequency as $MAF < 0.1\%$. The databases included our internal WES database, 1000 Genomes and Genome Aggregation Database (GnomAD; <http://gnomad.broadinstitute.org/>), which contained aggregated WES data of more than 123,136 individuals.

Genetic testing was then performed to inspect the candidate variant. Sanger sequencing was used to verify all the target variants identified by WES. The variants and the nearby sequences were downloaded from Genome Bioinformatics Database at University of California Santa Cruz (UCSC, <http://genome.ucsc.edu/>), and then Sanger

sequencing was applied based on the above description. The results were analyzed by BioEdit version 7.0.9.0. Ultimately, we used Sanger sequencing to confirm which gene that the candidate variant belonged to.

Verification of candidate variant

The candidate variant was verified by conservation of species and Gene Ontology Consortium (GO) cluster (Stevens, 2004). Then, National Center for Biotechnology information (NCBI) database was mainly used for literature reviews on the functions of candidate variant and gene. In addition, GeneCards (<https://www.genecards.org/>) was adopted among others for the basic information of candidate variant and gene. Also, Ingenuity

Pathway Analysis (IPA) were performed to seek for the interaction features of the candidate gene. We input the candidate gene into the IPA online tool and detected the biological pathway related to the gene. Finally, real-time PCR was carried out to ensure the expression of the candidate gene in sporadic patients and healthy individuals. Fermentas Maxima SYBR Green qPCR Master Mix (2×) was used in this experiment followed the provided protocol. The RT-PCR results was analyzed using Bio-Rad CFX Manager 2.1.

CRISPR-Cas9 gene editing system was then applied to edit CROCC for its verification. The standard procedure was followed according to the description of Ran et al (Ran et al., 2013). The oligo DNA sequence were designed using online tool “CRISPR Design” (<http://crispr.mit.edu/>). The plasmid was PX330 and the plasmid extraction using TIANprep Mini Plasmid Kit (TIANGEN). For the cell transformation, 293T cells was chosen. Immunofluorescence was used to observe the change of cilia after knocked down CROCC. The antibody was produced by Abcam, including primary antibody (Anti-Rootletin antibody, Ab121653) and secondary antibody (Goat Anti-Rabbit IgG H&L, Ab150077). Before immunofluorescence, the cells were cultured on cover glass for 3 days. Then the experiment was preformed following with the standard procedure of Abcam.

Results and Discussion

Clinical diagnosis

The epidemiological characteristics of both groups are shown in Table 1. No difference was shown regarding age, sex, smoking habits, alcohol consumption, overweight, sedentary lifestyle and inflammatory comorbidities. Concerning exposed individuals, 57.1% did not refer to any use of PPE and 21.4% reported a history of acute pesticide poisoning. The mean time of exposure was 6.0 (\pm 3.8) hours per day for 20.1 (\pm 8.5) years.

It is important to emphasize that the absence of a significant association between the epidemiological parameters in the evaluated groups is essential to minimize confounding factors in the context of immunological studies since several clinical and epidemiological factors can interfere with the immune function (Dowd and Aiello 2009; Brodin and Davis 2017).

Whole exome sequencing results

In order to identify the genetic cause underlying the AOA disease of the proband, DNA from all 5 family members were submitted for WES. The GATK analysis process was employed for the WES data which were from the sequence for all the proband family members (Supplementary Figure 2).

The samples were sequenced and yielded sufficient high-quality sequencing data with an average 11.36GB raw data after quality control. The exome sequencing data was used to map with the human genome version hg19 with the read depth of 100×. The exome capture efficiency and the mapping data coverage were evaluated among the whole samples set. On average, 99.86% of target regions were sequenced in these samples and a 138.30 folds mapping-reads-coverage were achieved in the sample set. Comparison analysis between family members was performed using mapped data from all five samples. The potential contribution of variants and total 53,550 coding variants were located and their distribution situation on proband's chromosome was drawn (Supplementary Figure 3). Furthermore, those variants were filtered out if they showed high frequency and polymorphic in the results. All the variants with minor allele frequency (MAF) of 0.1% or below were selected for this study.

In order to test whether the genetic cause of AOA was caused by craniovertebral junction and cervical vertebra regulatory genes, including HOX genes (HOXA1, HOXB1, HOXA2, HOXB2,

HOXA3, HOXB3, HOXD3 and HOXD4), MEOX1, OCN, OPG, COLA1 and RUNX2. The WES data of all patients showed that none of the craniovertebral junction and cervical spine related genes had rare mutations.

The genotype of the proband and 4 healthy family members were compared, and a total 432 variants was obtained. Then, the respective allele frequencies ($\leq 0.1\%$) in 65,000 population were screened by GnomAD. The total number of variants was reduced to seven in non-synonymous single nucleotide variants (SNV) and non-frameshift deletion mutation groups. Sanger sequencing showed similar results in the proband and 4 healthy family members. Then, SIFT/Polyphen2 analysis was done based on these seven variants and only one un-annotated variant of these SNVs (NM_014675.5 (CROCC): c.4702C>T p.(Arg1568Cys)) was assessed as “Damage” with SIFT score of 0.03, polyphen2 score of 0.964 and MutationTaster score of 1 (Table 1). While SNVs were not reported in the databases including the GnomAD database, the 1000 Genome Project database and the Human Gene Mutation Database, gnomAD, ExAC etc.

proband (Figure 1A). Subsequently, the DNA of 220 healthy samples and 73 sporadic patients were also tested for candidate variants by Sanger sequencing analysis. All of the genotype from 220 healthy volunteers and 68 sporadic patient samples appeared to be the same as the reference sequence genotype “C”, but there were 5 AOA sporadic patients containing the same SNV as the proband genotype “T” (Figure 1B). Furthermore, this mutation is absent from both the GnomAD exome database, as well as from the 1000 Genomes Projects database. Sanger sequencing revealed that this novel mutation was absent from both parents’ and the two healthy sisters’ blood DNA samples and therefore it was considered as a de novo mutation in the family.

The Candidate variant conservation analysis

The location of candidate variants on CROCC gene was charted by pathway builder Tool 2.0 (The ProteinLounge, USA) (Figure 2). According to UCSC database, NM_014675.5 (CROCC): c.4702C>T (r.4702c>u, p.(Arg1568Cys) showed high conservation of arginine in 72 out of 100 species with “C” which was the same as the

Table 1. Seven variants that were collected after SIFT/Polyphen2 analysis.

Gene	Variants	SIFT	ACMG criteria
NM_014675.5 (CROCC)	c.1756G>C (r.1756g>c, p.(Asp568His)	T	BP1
	c.4702C>T (r.4702c>u, p.(Arg1568Cys)	D	PP3
	c.1734C>G (r.1734c>g, p.(Asp578Glu)	T	BS1
	c.116G>A (r.116g>a, p.(Arg39Gln)	T	Uncertain Significance
NM_021012.5 (KCNJ12)	c.1214G>T (r.1214g>u, p.(Ser405Ile)	T	Uncertain Significance
	c.167A>C (r.167a>c, p.(Glu56Ala)	T	Uncertain Significance
	c.631C>T (r.631c>u, p.(Leu211Phe)	T	PP3

*Bold font was the final target variant. In ‘SIFT pred’ column, ‘D’ means ‘damaging’; ‘T’ means ‘tolerated’.

Genotype accuracy of candidate SNV

We sequenced samples from the proband and his family to determine the accuracy of genotype results. Firstly, for the candidate variant (NM_014675.5 (CROCC): c.4702C>T (r.4702c>u, p.(Arg1568Cys)), the reference genotype was “C”, the same as the genotype of healthy parents and sisters, but the genotype was “T” in the offspring

calculated by different methods (Figure 2A). Furthermore, the high conservation of the amino acid on the same chromosome locations were also revealed by UCSC (Figure 2B). In order to check the effects on candidate SNV, SWISS-MODE simulation was performed to compare the protein structure between wild type and the variant. The variant replaced arginine at position 772 with cysteine, thereby inducing intermolecular hydrogen bond.

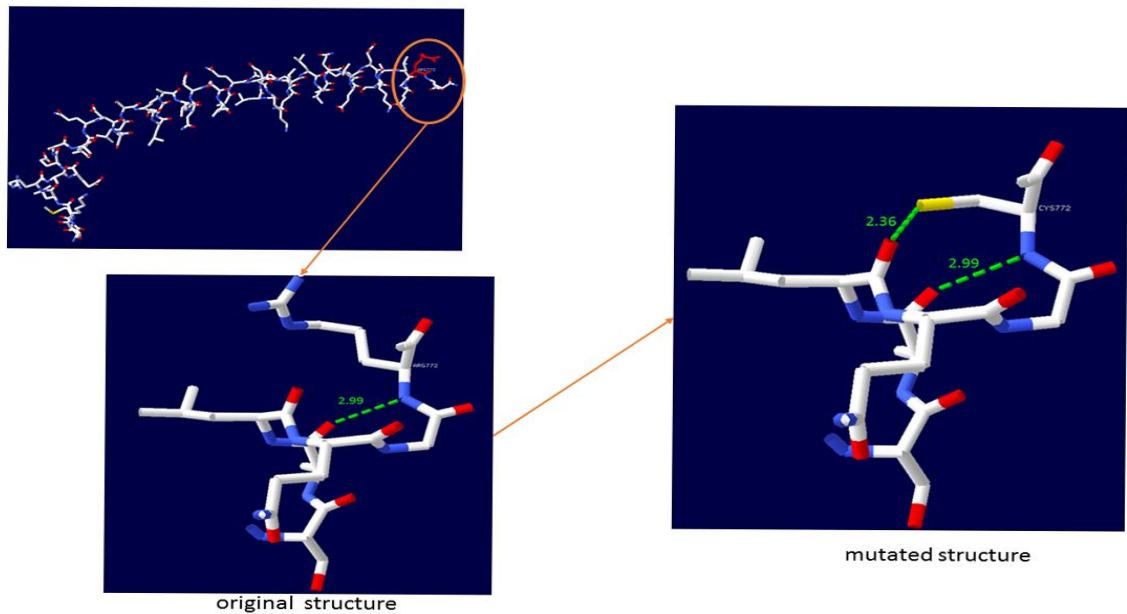
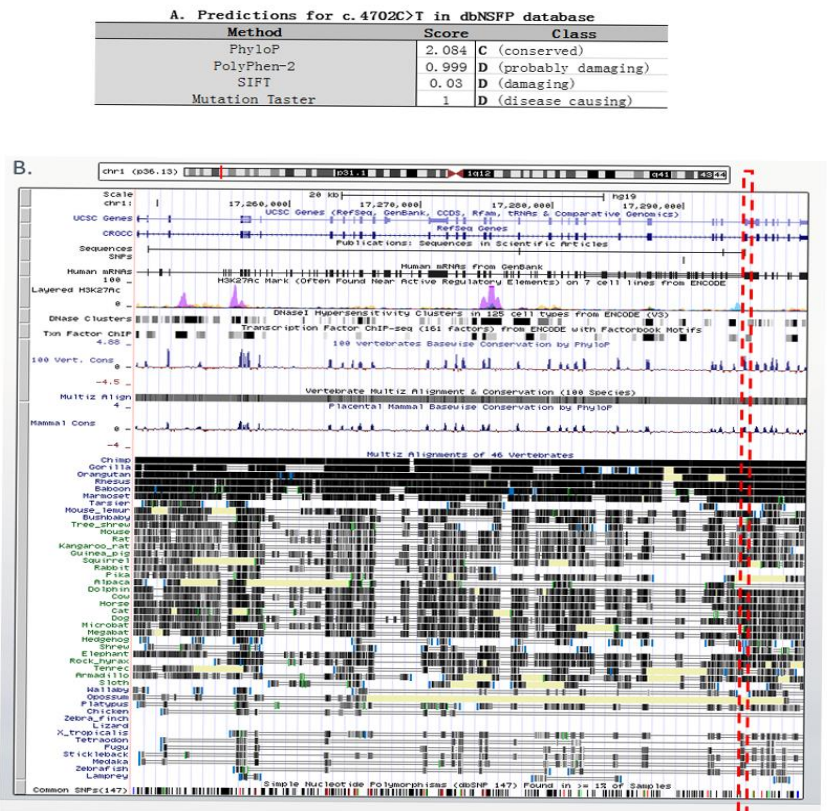


Figure 2. The allelic conservation (NM_014675.5 (CROCC): c.4702C>T (r.4702c>u, p.(Arg1568Cys)) and the structure simulations between wild type and mutated CROCC. (A) Pathogenicity prediction for (NM_014675.5 (CROCC): c.4702C>T (r.4702c>u, p.(Arg1568Cys)) in dbNSFP database. (B) The amino acid conservation results on CROCC and the amino acid containing the candidate SNV showed highly conservation. (C) Simulations on the protein structure after the candidate SNV changing from "C" to "T" showed the amino acid changed from arginine to cysteine.

This caused a shift in the electrostatic potential (Figure 2C) and finally changed the protein structure during translation which might cause functional changes of the protein.

Dysfunction of the protein containing candidate variant affected the hedgehog pathway

Five pathway essential proteins were tested by real-time PCR, including Indian Hedgehog Signaling Molecule (IHH), Patched 1 (PTCH), GLI Family Zinc Finger 1 (GLI1), Glypican 3 (GPC3) and Aliases for PTHLH Gene (PTHrp). The positions of these proteins in hedgehog pathway were shown in Figure 3. The results of real-time PCR showed that CROCC in samples with the candidate SNV was barely expressed ($p < 0.01$). But it was highly expressed in healthy and the non-mutated patient samples ($p > 0.5$). The expression of above five proteins were consistent with that of CROCC (Figure 4). These results indicated that the decreased expression of CROCC blocked the signal transmission of hedgehog signaling pathway and led to the dysfunction of ciliary. This might be one of the pathogenic causes of AOA.

Candidate variant affected cell cilia

In order to confirm if the candidate SNV on CROCC might indeed affect the cilia, we employed CRISPR-Cas9 gene editing system to knock out CROCC using the oligo DNA for the candidate SNV. Two pairs of oligo DNA were designed (Supplementary Table 1). The rate of cell transformation was 96% (Figure 5A). qPCR was performed to check if CROCC has been knocked out in monoclonal 293T cell. The CROCC knocked out monoclonal cells did not show any expression of the target sequence (Figure 5B), and the result was validated by target sequencing (Figure 5C). The expression of CROCC in 293T cell was detected by immunofluorescence and the RPE cell was used as the positive control. Fluorescence can still be seen in CROCC knockout cells, which indicated that CROCC may be related to incomplete knockout or the reparation (Figure 5D). Furthermore, immunofluorescence was used to detect the cilia on the surface of cells, and the results showed that there was almost no fluorescence in cells, indicating that the growth of cilia was affected (Figure 5E).

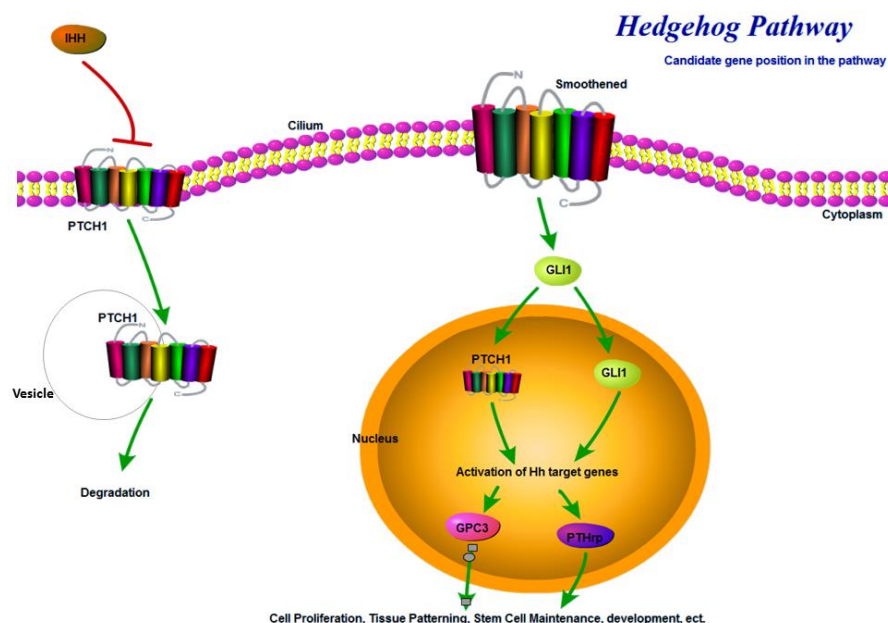


Figure 3. Location of the genes (IHH, Smoothened, PTCH1, GLI1, GPC3, PTHrp) in the hedgehog signaling pathway. This pathway can be affected by the candidate SNV mutation of CROCC and the genes activation.

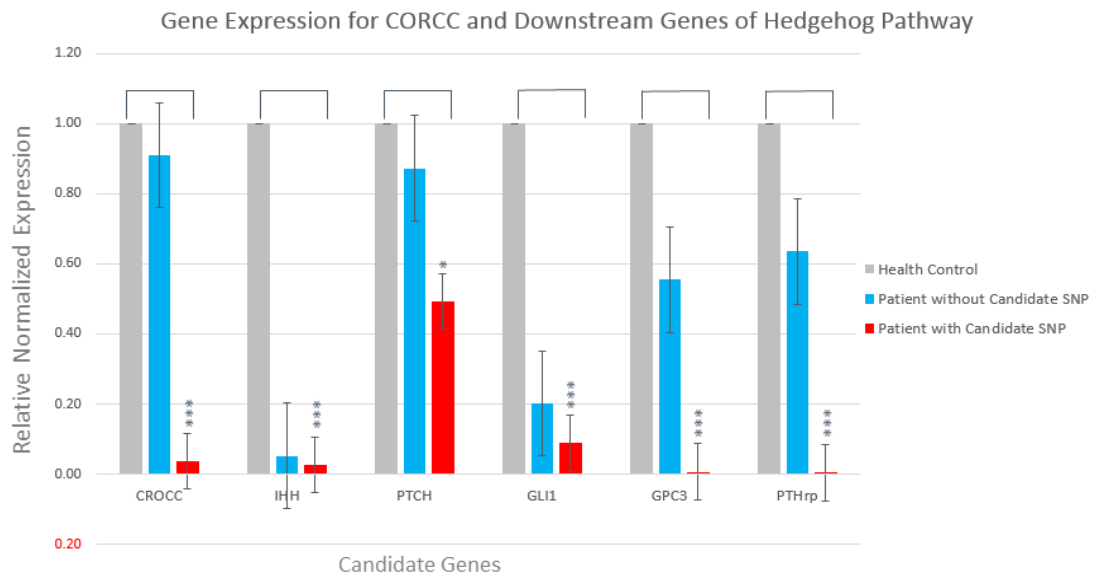


Figure 4. The real-time PCR results for CROCC, IHH, PTCH1, GLI1, GPC3, PTHrp. The red column was the gene expression for the patients with candidate SNV, the gray column for the health donors' gene expression, and the blue column for the patients without candidate SNV. Each bar in this chart is the average cq values of real-time PCR. The gene expression level (CROCC and downstream genes including GLI1, GPC3, PTHrp) of AOA patients with candidate SNV (red bar) were reduced compared with those of healthy samples (blue bar).

In this study, a missense novel mutation (NM_014675.5(CROCC): c.4702C>T (r.4702c>u, p.(Arg1568Cys)) on gene CROCC was identified from a Chinese family with a patient of AOA as our proband. The SNV with low incidence rate (< 0.01) could cause structure changes of CROCC protein, and it led to the decrease of the key proteins expression in hedgehog pathway. After CRISPR-Cas9 editing, knockdown of the CROCC in 293T cells resulted in dysfunction of cilia. Meanwhile, hedgehog proteins IHH, PTCH, GLI1, GPC3 and PTHrp have been decreasingly expressed in patients with the SNV, which indicated a similar conclusion. Therefore, target SNV caused dysfunction of skeletal development and finally led to the morbidity of AOA. Although it is very likely that the target SNV (NM_014675.5 (CROCC): c.4702C>T (r.4702c>u, p.(Arg1568Cys)) could be one of the AOA causing mutations, future experiment of single nucleotide knockout in vivo can finally confirm this result. In addition, we also did the autosomal recessive inheritance analysis. However, there was no variant identified or

observed from the analysis (data not shown). Stability is essential to maintain the normal function of proteins, and structural changes may lead to their dysfunction. According to Genecards, NCBI, IPA and KEGG database, the function of CROCC was ensured as the major structural component of the ciliary rootlet which was the supporter for long-time maintenance of cilia (Bahe et al., 2005; Yang et al., 2002). In 2012, Conroy et al. reported that CROCC (rootletin) could restrain DNA damage-induced centriole splitting and facilitate ciliogenesis (Conroy et al., 2012). In the study, authors found that the primary cilium formation would be reduced if CROCC was knocked down. Primary cilium plays sensory roles in normal formation of the appendicular skeleton through disruption of multiple signaling pathways, including hedgehog signaling pathway (Kinnebrew et al., 2019). According to previous studies, the dysfunction of CROCC may cause the lack of ciliary rootlet which leads to unstable primary cilia and dysfunction of the cilia as time goes by (Bahe et al., 2005).

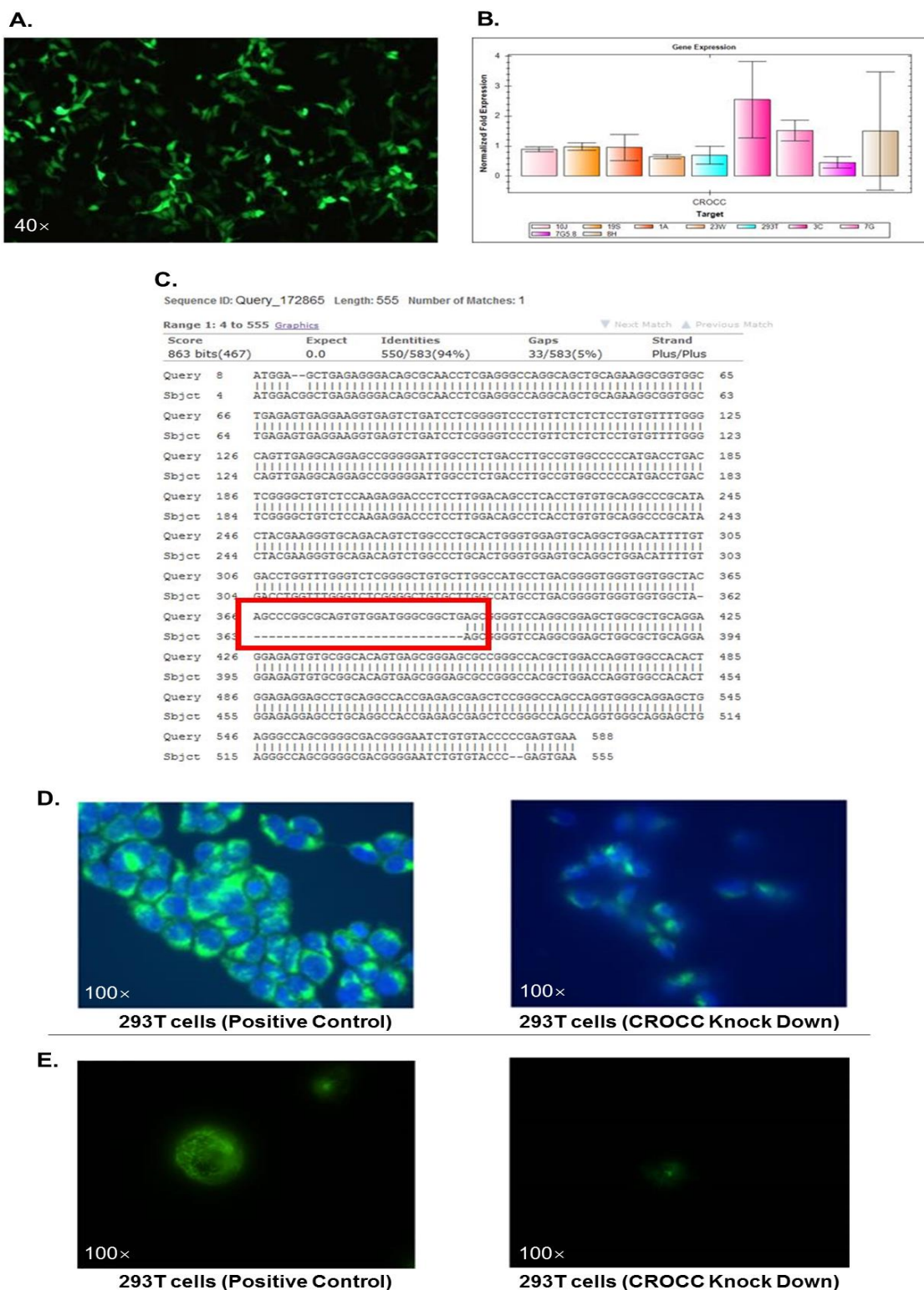


Figure 5. CRISPR-Cas9 and immunofluorescence results for *CROCC* knock-down cells. (A) Cell transformation. (B) qPCR for transformed monoclonal cells. (C) Sanger sequencing results showed that target sequence was knocked down in the testing cell line. (D) Immunofluorescence for the control cell 293T and RPE, and the testing cell line of 293T knocked down cells. (E) Immunofluorescence for cilia in the control and the testing cell line.

Hedgehog signaling pathway plays a critical role in development of bone formation. The hedgehog receptor complex is composed of two trans-membrane proteins, smoothened (Sonic Hedgehog Signaling Molecule, SHH) and patched (PATH1) (Hu and Song, 2019). Previous studies showed that the abnormal expression of Indian hedgehog signaling molecule (IHH) led to smaller calvaria with reduced expanse, thickness and mineralization and widened suture in mice (Zhou et al., 2014). It promotes bone formation by regulating osteogenic differentiation (Zhou et al., 2014). Furthermore, the primary cilia are indispensable for regulating skeletal malformations, developments and maintenance (Moore and Jacobs, 2018), reinforcing the idea that the dysfunction of CROCC may cause the lack of ciliary rootlet which leads to unstable primary cilia and dysfunction of the cilia as time goes by (Yang et al., 2002). Since cilia is the occurrence places for hedgehog signaling pathway, comparing the main component genes' expression of hedgehog signaling pathway between patients and healthy donors could be an effective way to ensure if the hedgehog pathway were dysfunction. WES technology makes it easier to find the relationship between mutations on coding exon, gene functions and development of diseases (Mori et al., 2017). Familial WES analysis was performed in this study and a novel variant (NM_014675.5 (CROCC): c.4702C>T (r.4702c>u, p.(Arg1568Cys)) was discovered and collected as the only candidate variant. The mutation of the candidate variant could lead to abnormal express of CROCC protein. It might affect the formation of cilia and block the signal of hedgehog signaling pathway, which might finally lead to the development of AOA. All these suggested that (NM_014675.5 (CROCC): c.4702C>T (r.4702c>u, p.(Arg1568Cys)) led to cilia disorder, blocked the hedgehog signaling pathway, and potentially caused skeletal malformation.

Due to the low incidence rate of AOA and the absence of symptoms, the study of AOA is rare, and genome research is more difficult. In this study, we used WES as the main method to study AOA from the perspective of genomics. We observed a candidate mutation which led to abnormal protein expression and then diagnosed as AOA. Our study demonstrated that WES has great potential in detecting target gene and establishing systematical database, which can be used in clinical for individual diagnosis (Okou et al., 2014; Prada et al., 2014). In addition, it is cost-efficient and can save money for a single clinical trial. As the results of WES may be specific, it could also be used to develop effective treatments for individual patient.

To the best of our knowledge, our study is the first study of AOA using WES. Our results may have positive implications for the clinical research of CVJ. The method used in this study certainly provides a new direction for understanding the pathogenesis of the disease from the gene level and potentially can be used to develop effective treatments.

Conclusion

Our results suggested that assimilation of atlas AOA might be caused by the mutation of CROCC: c.4702C>T (r.4702c>u, p.(Arg1568Cys)). With strong amino acid conservation and interaction regulation, the variant mutation could cause the signal disorder of skeletal development which may lead to the defective bone formation and finally cause the development of assimilation of atlas AOA.

Acknowledgements

We would like to thank all donors for the sample collections, and we are also grateful to all people

who participated in this study for their time, effort and collaboration.

Conflict of interest

The authors declare no conflict of interest.

References

- Abd OHE, Rosenberg D, Gomba L, Isaac Z (2008) The lateral atlanto-axial joint as a source of headache in congenital atlanto-occipital fusion. *Am J Phys Med Rehabil* 87: 232-237.
- Abuín J, Pichel J, Tomás FP, Amigo J (2015) BigBWA: approaching the Burrows-Wheeler aligner to Big Data technologies. *Bioinformatics* 24: 4003-4005.
- Al-Motabagani MA, Surendra M (2006) Total occipitalization of the atlas. *Anat Sci Int* 81: 173-180.
- Allen M (1972) The Upper Cervical Spine. Regional Anatomy, Pathology and Traumatology. A Systematic Radiological Atlas and Textbook. *J Anat* 113: 264.
- Anderson AJ, Towns GM, Chiverton N (2006) Traumatic Occipitocervical Disruption: A New Technique for Stabilisation. Case Report and Literature Review. *J Bone Joint Surg Am* 88: 1464-1468.
- Bahe S, Stierhof YD, Wilkinson CJ, Leiss F, Nigg EA (2005) Rootletin forms centriole-associated filaments and functions in centrosome cohesion. *J Cell Biol* 171: 27-33.
- Bayrakli F, Guclu B, Yakicier C, Balaban H, Kartal U, Erguner B, Sagiroglu MS, Yuksel S, Ozturk AR, Kazanci B (2013). Mutation in MEOX1 gene causes a recessive Klippel-Feil syndrome subtype *BMC Genets*. 14: 1-7.
- Campos D, Silva TH, Ellwanger JH, Goerck ML, Kipper JF, Piazza JL, Neto LK (2012) Atlanto-occipital Fusion and Its Neurological Complications: A Case Report. *J Morphol* 29: 111-113.
- Conroy PC, Saladino C, Dants TJ, Lalor P, Dockery P, Morrison CG (2012) C-NAP1 and rootletin restrain DNA damage-induced centriole splitting and facilitate ciliogenesis. *Cell cycle* 11: 3769-3778.
- Consortium TIH (2003). The International HapMap Project. *Nature* 426: 789-796.
- Doerks T, Copley RR, Schultz J, Ponting CP, Bork P (2002) Systematic identification of novel protein domain families associated with nuclear functions. *Genome Res* 12: 47-56.
- Hemamalini S (2014) Atlanto-occipital Fusion and Other Variations at the Base of the Skull: A Case Report. *Int J Anat Var* 7: 80-82.
- Hu A, Song BL (2019) The interplay of Patched, Smoothened and cholesterol in Hedgehog signaling. *Curr Opin Cell Biol* 61: 31-38.
- Olivier-Mason A, Kazatskaya A, Kennedy J, McLachlan IG, Blacque OE, Sengupta P (2016) A Conserved role for girdin in basal body positioning and ciliogenesis. *Dev Cell* 38: 493-506.
- Jayanthi V, Kulkarni R, Kulkarni RN (2003) Atlanto-Occipital Fusion - Report of Two Cases. *J anat Soc India* 52: 71-73.
- Kassim NM, Latiff AA, Das S, Ghafar NA, Suhaimi FH, Othman F, Hussan F, Sulaiman IM (2010) Atlanto-occipital Fusion: an Osteological Study with Clinical Implications. *Bratislavské lekárske listy* 111: 562-565.
- Khanmanarong K, Woraputtaporn W, Ratanasuwan S, Namking M, Chaijaroonkhanarak W, Sae-jung S (2013) Occipitalization of the atlas: its incidence and clinical implications. *Acta Med Scand* 42: 41-45.
- Kim KR, Lee YM, Kim YZ, Cho Y, Kim J, Kim K, Lee IC (2013) Cervical Myelopathy Secondary to Atlanto-occipital Assimilation: The Usefulness of the Simple Decompressive Surgery. *Korean J Spine* 10: 189-191.
- Kinnebrew M, Iverson EJ, Patel BB, Pusapati GV, Kong JH, Johnson KA, Luchetti G, Eckert KM, McDonald JG, Covey, DF (2019) Cholesterol accessibility at the ciliary membrane controls hedgehog signaling. *eLife* 8: e50051.
- Lee JE, Gleeson JG (2011) Cilia in the nervous system: linking cilia function and neurodevelopmental disorders. *Curr Opin Neur* 24: 98-105.
- Moore ER, and Jacobs CR (2018) The primary cilium as a signaling nexus for growth plate function and subsequent skeletal development. *J Orthop Res* 36: 533-545.
- Mori RD, Romani M, Arrigo SD, Zaki MS, Loreface E, Tardivo S, Biagini T, Stanley V, Musaeu D, Fluss J (2017) Hypomorphic Recessive Variants in SUFU Impair the Sonic Hedgehog Pathway and Cause Joubert

- Syndrome with Cranio-facial and Skeletal Defects. *AM J Hum Genet* 101: 552-563.
- Nimje DA, Wankhede HA (2014) Atlanto-occipital Fusion: A Rare Anomaly of the Craniocervical Junction Edorium. *J Anat Berl* 1: 1-4.
- Okou DT, Mondal K, Faubion WA, Kobrynski LJ, Denson LA, Mulle JG, Ramachandran D, Xiong Y, Svingen P, Patel V (2014) Exome sequencing identifies a novel FOXP3 mutation in a 2-generation family with inflammatory bowel disease. *J Pediatr Gastroenterol Nurt* 58: 561-568.
- Prada CE, Gonzaga-Jauregui C, Tannenbaum R, Penney S, Lupski JR, Hopkin RJ, Sutton VR (2014) Clinical utility of whole-exome sequencing in rare diseases: Galactosialidosis European. *J Med Genet* 57: 339-344.
- Ran FA, Hsu PD, Lin CY, Gootenberg JS, Konermann S, Trevino AE, Scott DA, Inoue A, Matoba S, Zhang, Y (2013) Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* 154: 1380-1389.
- Saini V, Singh R, Bandopadhyay M, Tripathi SK, Shamal SN (2009) Occipitalization of the Atlas: Its Occurrence and Embryological Basis International. *J Anat Var* 2: 65-68.
- Schoichet BK, Baase WA, Kuroki R, Matthews BW (1995) A relationship between protein stability and protein function. *P Natl Acad Sci USA* 92: 452-456.
- Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, Sirotkin K (2001) dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res* 29: 308-311.
- Stevens R (2004) Dictionary of Bioinformatics and Computational Biology. *Brief Bioinform* 2:2.
- Yang J, Liu XQ, Yue GH, Adamian M, Bulgakov O, Li T (2002) Rootletin, a novel coiled-coil protein, is a structural component of the ciliary rootlet. *J Cell Biol* 159: 431-440.
- Zhang W, Taylor SP, Ennis HA, Forlenza KN, Duran I, Li B, Sanchez J, Nevearez L, Nickerson DA, Bamshad M (2017) Expanding the genetic architecture and phenotypic spectrum in the skeletal ciliopathies. *Hum Mutat* 39: 152-166.
- Zhou J, Chen Q, Lanske B, Fleming BC, Terek R, Wei X, Zhang G, Wang S, Li K, Wei L (2014) Disrupting the Indian hedgehog signaling pathway in vivo attenuates surgically induced osteoarthritis progression in Col2a1-CreERT2; Ihhf1/fl mice. *Arthritis Res Ther* 16: R11.

Table S1. Primer sequences for PCR and real-time PCR experiments.

Target Gene	Experiment	Forward primer (5' to 3')	Reverse Primer (5' to 3')
CROCC	PCR	CAGACCAGTGCCCTGAATCG	TCACTCGGGTACACAGATTCC
CROCC	PCR	CAGAGAGAACGGGTAAGCCT	GTACACAGATTCCCCGTCG
KCNJ12	PCR	ATGCTGTCGTCTCTGTTG	TGTTGGCGAACTCAATGT
KCNJ12	PCR	GCATCATCGACTCCTTCAT	TGGCCTCGTCAATCTCAT
CROCC	PCR	TGGAGCTGACACTAGAGACG	GGTTGCGGGTGACAATCTCC
IHH	PCR	TCCGTCAAGTCCGAGCACT	GTCCTGAGTCTCGATGACCTG
SHH	PCR	CCAAGGCACATATCCACTGCT	GTCTCGATCACGTAGAAGACCT
GLI1	PCR	AGCGTGAGCCTGAATCTGTG	CAGCATGTACTGGGCTTTGAA
GPC3	PCR	GAAAGTGAGACTGCGGTGA	GAAGAAGCACACCACCGAGA
PTCH1	PCR	GAAGAAGGTGCTAATGTCCTGAC	GTCCCAGACTGTAATTTCCGC
PTHRP	PCR	ACTTTCCGGAAGCAACCAG	TCTCCGCTCGCGCTC
CROCC	CRISPR-Cas9	CACCGCAGCCCGGCGCAGTGTGGAT	AAACATCCACACTGCGCCGGGCTGC
CROCC	CRISPR-Cas10	CACCGGCTACAGCCCGGCGCAGTG	AAACCACTGCGCCGGGCT GTAGCC

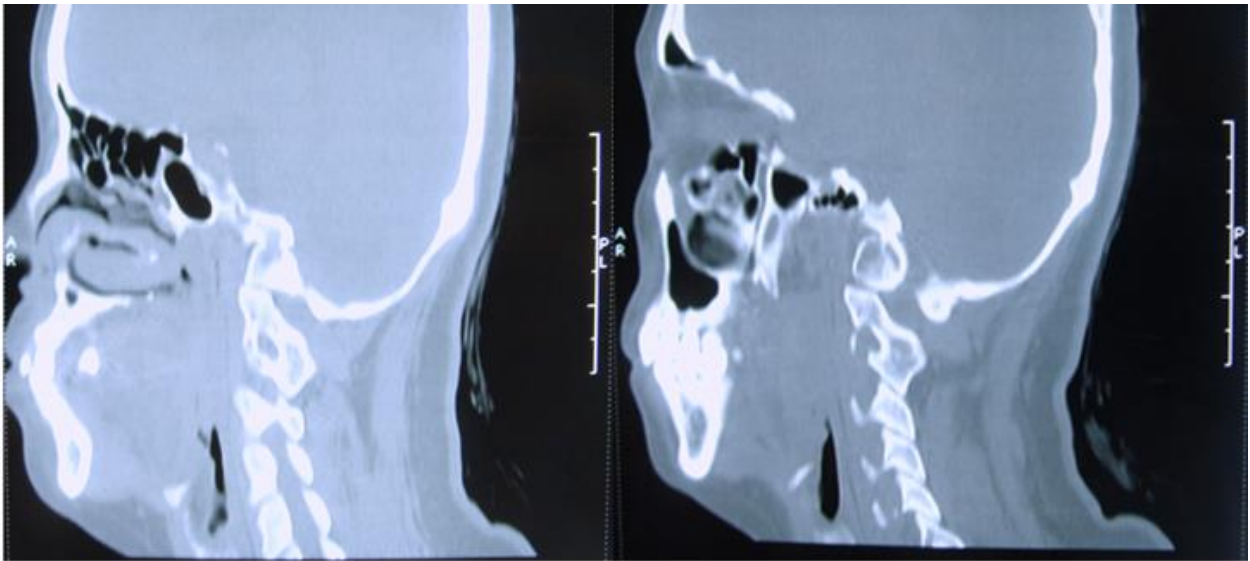


Figure S1. The two-dimensional sagittal reformatted computed tomographic reconstructions of the patient before the operation treatment in 2014. CT scans results showed malformation statement of the proband when he was diagnosed as atlanto-occipital fusion in 2014. The disease has developed into complete fusion of the lateral atlantoaxial muscle and occipital condyle.

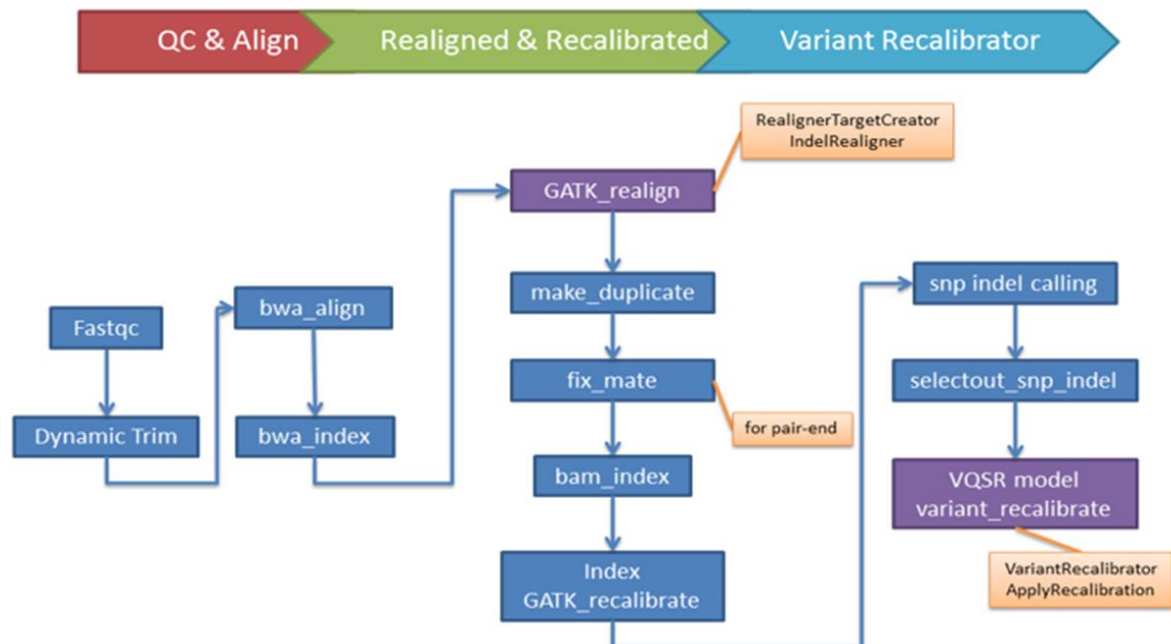


Figure S2. The GATK analysis process for the WES data which were from the sequence for all the proband family members.

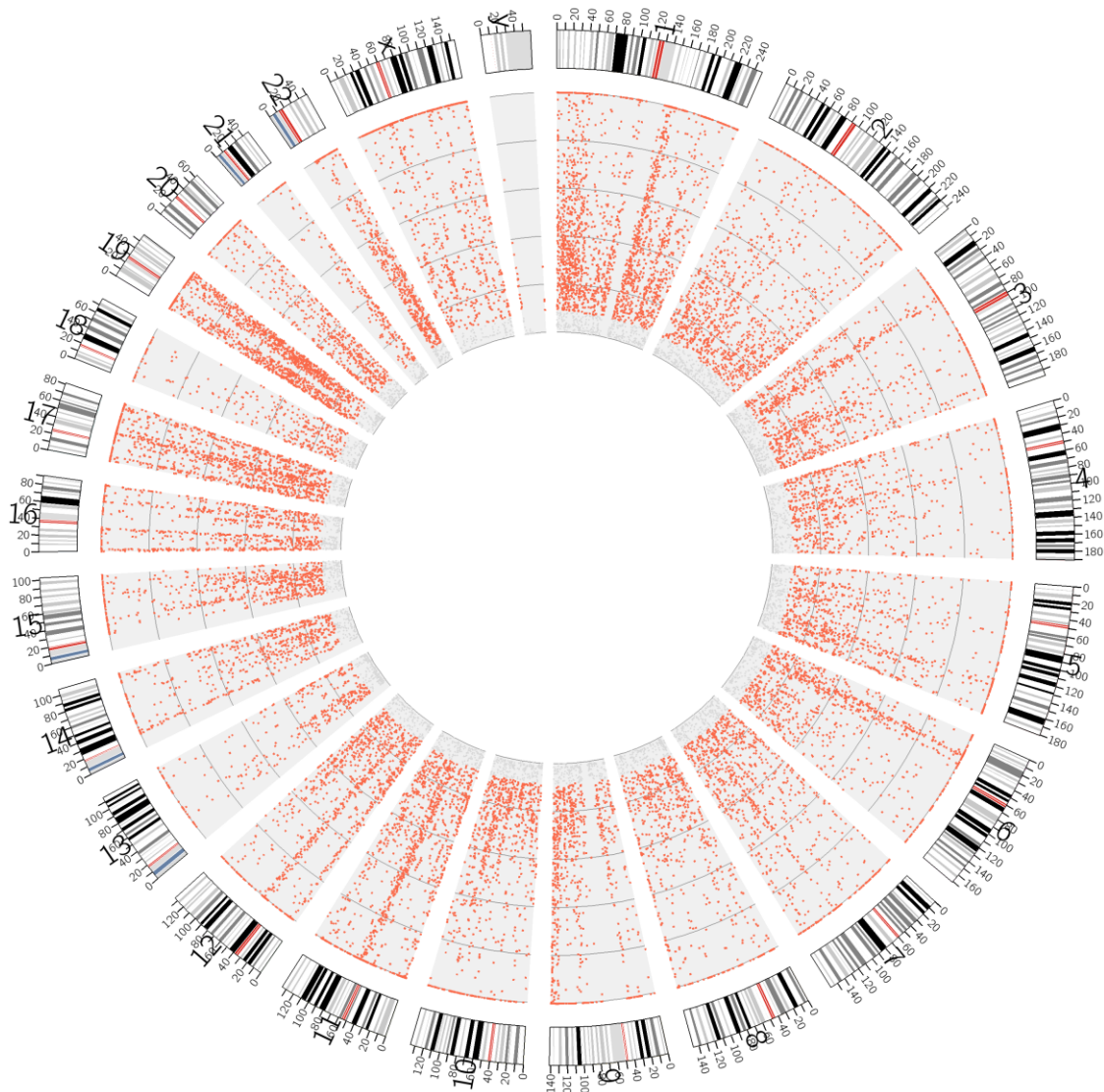


Figure S3. The gene coverage on the chromosomes of the proband, including all genes containing all variants mapped from the WES data.