TECHNOLOGIES FOR PRENATAL DIAGNOSIS AND ASSESSMENT OF GENETIC DISORDERS

EDITED BY: Fan Jin, Yueqiu Tan, Evica Rajcan-Separovic and Peter C. K. Leung
PUBLISHED IN: Frontiers in Genetics

frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-88963-739-3
DOI 10.3389/978-2-88963-739-3

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

TECHNOLOGIES FOR PRENATAL DIAGNOSIS AND ASSESSMENT OF GENETIC DISORDERS

Topic Editors:
Fan Jin, Zhejiang University, China
Yueqiu Tan, Central South University, China
Evica Rajcan-Separovic, University of British Columbia, Canada
Peter C. K. Leung, University of British Columbia, Canada

Birth defects are one of the major public health concerns in the world, as they cause approximately 20% of infant deaths. Genetic disorders, including chromosome abnormalities and single gene disorders, are the most common causes of birth defects for which there is no efficient treatment. Prenatal genetic screening and diagnosis allow early identification of affected conceptuses and facilitates reproduction planning or counseling.

Molecular technologies have developed rapidly in recent years and have been widely used in screening and diagnosis of genetic disorders at all stages of prenatal development (e.g. pre-implantation, embryonic and fetal). However, their performance still needs to be validated and assessed as the balance between their advantages and disadvantages need to be discussed.

With the ability to detect copy number variations (CNVs), polyploidy, uniparental disomy and maternal cell contamination, SNP-based chromosomal microarray analysis (CMA) is showing the unique importance in diagnosing chromosomal abnormalities. The interpretation of CNVs remains a challenge; however, ultrasound and biochemical screening improve the diagnosis of fetal chromosomal abnormalities. Whole exome sequencing (WES) and whole genome sequencing (WGS) play increasingly significant roles in prenatal and carrier screening for genetic disorders. NGS-based non-invasive prenatal screening (NIPS) is now widely used for detecting common autosomal aneuploidies and has shown the potential of detecting microdeletions and microduplications. However, further investigations of the sensitivity and accuracy are required and large-scale data is necessary to evaluate the performance and clinical applications of current and new methods. Recently, reports of application of newer technologies in prenatal setting became available. Examples include third generation sequencing (reading the nucleotide sequences at the single molecule level), digital PCR (used for direct quantification of DNA) and cell-based NIPT.

In the followed listed papers, the authors showed their successful experiences in identifying novel mutation, detecting low-level mosaicism or de novo mutations limited in germline cells, investigating the association of the CNVs with specific phenotypic alterations by using WES, CMA, digital PCR and some other new-developed molecular techniques. More interesting, the authors also presented a report about the evaluation of diagnostic yield in fetal WES, which suggested a new tendency to apply WES or WGS directly for prenatal diagnosis. We believed that the efficiency of scanning causative mutations and prenatal or preimplantion genetic

diagnosis for genetic disorders will further improved based on the technologies of whole genomic sequencing with further improved output and resolution. **New techniques, such as quick-WES for the newborn in intensive care unit, direct-WGS for prenatal diagnosis and non-invasive test for fetal monogenic disorders, will become available in the near *future*.**

Citation: Jin, F., Tan, Y., Rajcan-Separovic, E., Leung, P. C. K., eds. (2020). Technologies for Prenatal Diagnosis and Assessment of Genetic Disorders. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88963-739-3

Table of Contents

frontiers in Genetics

ORIGINAL RESEARCH
published: 18 June 2019
doi: 10.3389/fgene.2019.00565

Whole Exome Sequencing Identified a Novel Biallelic *SMARCAL1* Mutation in the Extremely Rare Disease SIOD

Jing Jin[1†], Keke Wu[2†], Zhenwei Liu[3], Xiaomin Chen[3,4], Shan Jiang[3], Zhen Wang[1,5] and Weixing Li[1,6*]*

[1] School of Laboratory Medicine and Life Sciences, Wenzhou Medical University, Wenzhou, China, [2] Wenzhou Center for Disease Control and Prevention, Wenzhou, China, [3] Institute of Genomic Medicine, Wenzhou Medical University, Wenzhou, China, [4] Center of Scientific Research, The Second Affiliated Hospital and Yuying Children's Hospital of Wenzhou Medical University, Wenzhou, China, [5] Research Center of Blood Transfusion Medicine, Education Ministry Key Laboratory of Laboratory Medicine, Zhejiang Provincial People's Hospital, People's Hospital of Hangzhou Medical College, Hangzhou, China, [6] Zhejiang Center for Clinical Laboratory, Zhejiang Provincial People's Hospital, People's Hospital of Hangzhou Medical College, Hangzhou, China

OPEN ACCESS

Edited by:
Fan Jin,
Zhejiang University, China

Reviewed by:
Yanan Wang,
UT Southwestern Medical Center, United States
Chenming Xu,
The International Peace Maternity and Child Health Hospital, China

***Correspondence:**
Zhen Wang
yuchen6046@163.com
Weixing Li
ysw967@163.com

[†] *These authors have contributed equally to this work*

Specialty section:
This article was submitted to Genetic Disorders, a section of the journal Frontiers in Genetics

Received: *29 January 2019*
Accepted: *29 May 2019*
Published: *18 June 2019*

Citation:
Jin J, Wu K, Liu Z, Chen X, Jiang S, Wang Z and Li W (2019) Whole Exome Sequencing Identified a Novel Biallelic SMARCAL1 Mutation in the Extremely Rare Disease SIOD. Front. Genet. 10:565. doi: 10.3389/fgene.2019.00565

Schimke immuno-osseous dysplasia (SIOD) is an extremely rare autosomal recessive pleiotropic disease. Although biallelic mutations in *SMARCAL1* gene have been reported to be the genetic etiology of SIOD, its molecular diagnosis has been challenging in a relatively proportion of cases due to the extreme rarity. Here, we made a definitive SIOD diagnosis of a 5-year-old girl with an extremely mild phenotype by applying whole exome sequencing (WES). As a result, a novel maternal mutation (c.2141+5G > A) confirmed to create a novel splice donor site combined with a known paternal mutation (c.1933C > T; p.Arg645Cys) were detected. In addition, previous reported SIOD cases showed excessive enrichment for mutations in the helicase ATP-binding and C-terminal domains of SMARCAL1. Similarly, the novel mutation we identified caused a mutant protein truncated in the SMARCAL1 C-terminus. Interestingly, based on the phenotypic profile, compared to reported cases, the patient in our study exhibited milder symptoms with renal dysfunctions limited to asymptomatic proteinuria, but no neurological signs or recurrent infections. Moreover, we identified 73 *SMARCAL1*-interacting genes, which formed a significant interconnected interaction network with roles in disease-related pathways such as double-strand break repair via homologous recombination, DNA repair, and replication fork processing. Notably, the top 15 *SMARCAL1*-interacting genes all showed a similar renal temporal expression pattern. Altogether, to our knowledge, the case in this study is the first case diagnosed originally based on a genetic test via WES rather than a characteristic phenotype. The identification of the novel allelic mutation (c.2141+5G > A) extends the phenotypic spectrum of *SMARCAL1* mutations and the following bioinformatics analysis presents additional genetic evidence to illustrate the role of *SMARCAL1* in SIOD.

Keywords: Schimke immuno-osseous dysplasia (SIOD), *SMARCAL1*, whole exome sequencing, mild phenotype, bioinformatics

INTRODUCTION

Schimke immuno-osseous dysplasia (SIOD; OMIM #242900) is a rare autosomal recessive disorder with an estimated incidence of 1 in 3 million live births in the United States (Santangelo et al., 2014), which was first described by Schimke et al. in 1971 (Schimke et al., 1971). SIOD is a multisystem disease, the main clinical features of which include growth failure, spondyloepiphyseal dysplasia, progressive nephropathy, and poor cellular immunity (Saraiva et al., 1999; Boerkoel et al., 2000). Hypothyroidism, bone marrow failure, and episodic cerebral ischemia have also been reported in some severely affected patients (Boerkoel et al., 2000). Phenotypic abnormalities are another feature of SIOD patients and usually consist of a broad nose, lumbar lordosis, and a protruding abdomen. Many patients also have microdontia, hypodontia, or malformed deciduous and permanent molars along with hyperpigmented macules located on the trunk (Saraiva et al., 1999; Morimoto et al., 2012).

The SIOD disease is frequently reported to be caused by biallelic mutations in the gene *SMARCAL1* (SWI/SNF-related, matrix associated, actin-dependent regulator of chromatin, subfamily A-like 1), a member of the SNF2 family of proteins that regulates gene transcription, DNA replication, repair and recombination in the context of chromatin (Bansbach et al., 2010). *SMARCAL1* was found to play an important role in replication fork restarts, cell cycle progression, and DNA damage responses (Ciccia et al., 2009). Various mutations have been found to be distributed throughout the whole gene (Lipska-Zietkiewicz et al., 2017). Notably, nonsense or frameshift mutations of *SMARCAL1* can lead to a severe phenotype (Boerkoel et al., 2002), and biallelic missense mutations within the SNF2 domain in other family members were also found to affect protein subcellular localization, enzymatic activity, abundances and chromatin binding, etc. (Elizondo et al., 2009).

Schimke immuno-osseous dysplasia is a phenotypic heterogeneous disease with a marked variation in severity of clinical manifestations. Based on the severity, SIOD has been classified into two different subtypes: severe and mild (Ehrich et al., 1995; Saraiva et al., 1999). Patients with the severe subtype mostly suffer from recurrent infections and develop renal failure or cerebrovascular disease in the first 2 to 5 years of life (Yue et al., 2010). The diagnosis of SIOD mainly relies on characteristic clinical and radiographic features followed by molecular testing (mainly single-gene testing and a multigene panel) if the clinical evidence is insufficient (Lipska-Zietkiewicz et al., 2017), while whole exome sequencing (WES) has seldom been applied to detect *SMARCAL1* mutations (Morimoto et al., 1993).

In this study, we are trying to make a diagnosis of a 5-year-old girl with mild SIOD phenotypes, and in which it was difficult for the pediatrician to make a definitive diagnosis based on phenotype features only. By applying WES technology, one novel mutation (c.2141+5G > A) adjacent to the 5′ donor splice site predicted to disrupt splicing function and one well-known mutation (c.1933C > T; p.Arg645Cys) were identified in the *SMARCAL1* gene. Based on the results of genetic testing, we clarified an SIOD diagnosis of the 5-year-old girl, and the diagnosis was supported by the following series of abnormal clinical symptoms. Compared to previously reported patients who were diagnosed according to characteristic phenotypes with missense mutations at the same position of the Arg645 amino acid, our patient shows the mildest phenotype with no recurrent infections or neurological signs and extremely mild renal dysfunction. Our comprehensive bioinformatics analysis presented in this study further demonstrated the functional role of *SMARCAL1* in SIOD.

MATERIALS AND METHODS

Patient Recruitment

This study was conducted in accordance with the guidelines of the Declaration of Helsinki. It was also approved by the Ethics Committee of the Second Affiliated Hospital and Yuying Children's Hospital of Wenzhou Medical University. Written informed consent was obtained from both parents of the patient. SIOD was diagnosed according to the results of a genetic test ordered by an experienced pediatrician. Routine clinical and laboratory examinations, including hematological and immunological tests and skeletal X-ray, were performed on the patient to confirm the diagnosis.

Whole Exome Sequencing and Variant Calling

Genomic DNA was extracted from the peripheral blood of the patient according to standard procedures using the QIAGEN DNeasy Blood & Tissue Kit (Qiagen, Valencia, CA, United States) and then sheared to 150 bp fragments in length. Whole-exome capture using the Agilent SureSelect Human All Exon v5 Kit (Agilent Technologies, Santa Clara, CA, United States) and high-throughput sequencing by utilizing an Illumina HiSeq 4000 sequencer (Illumina, San Diego, CA, United States) were conducted.

Computationally efficient read preprocessing and quality control for high-throughput sequencing data sets were taken with adaption of a canonical pipeline (Wang et al., 2013). The Trim Galore program was used to remove low-quality reads and adapters. The filtered reads (Phred-scaled quality score $\geq$ 30 and read length $\geq$ 80 bp) were aligned to the human reference genome (GRCH37/hg19) with the Burrows-Wheeler Alignment Tool (BWA) pipeline. Picard was then utilized to realign the reads from the BAM files and label the duplicated reads. In addition, local realignment and map quality score recalibration were performed. All variants, including single-nucleotide variants (SNVs) and InDels, were called according to three incorporated GATK tools: RealignerTargetCreator, IndelRealigner, and BaseRecalibrator. *De novo* and biallelic mutations were identified using mirTrios (Li et al., 2015).

Variant Annotation and Prioritization

The ANNOVAR software tool and in-house codes were used to annotate all of the mutations (Wang et al., 2010). Based on phenotype-related databases including HGMD, any variant was assessed as to whether it was associated with a previously recorded phenotype. To detect rare mutations, variants with a minor allele frequency (MAF) > 0.1% in various publicly available variant databases, including ExAC, 1000 Genome, dbSNP and ESP were filtered out. Subsequently, the effects of the detected variants were predicted by four software tools: STFT (Kumar et al., 2009), Polyphen (Adzhubei et al., 2010), LRT (Chun and Fay, 2009), and CADD (Kircher et al., 2014). After Sanger sequencing confirmation, the remaining variants were thought to be high-confidence causative variants.

RNA Exaction and Exon Junction Detection of cDNA

Total cellular RNAs from the Patient and control were isolated from blood samples using the TRIzol reagent (Invitrogen). And 1 μg of total RNA was used for reverse transcription with Reverse Transcription System (Promega) using oligo-dT primers. *SMARCAL1*-specific PCR primers were designed to amplify the region spanning the exon 13. Primer sets for hypothetical exon junctions were as follows: 5′-GCTGCAGCCAAGGAAATGAC-3′ in exon12-F and 5′-GCGTCCAGGACCACCTTATG-3′ in exon14-R. The final PCR products were visualized with electrophoresis, and submitted for Sanger sequencing as well.

Construction of Physical Interaction Network and Biological Process Enrichment Analysis by DAVID Bioinformatics Resources

The direct protein–protein interaction (PPI) data set used in this study was downloaded from the STRING database. We extracted genes interacting with *SMARCAL1* in the STRING database with interaction scores greater than 400 to construct an internal interaction network. To evaluate the significance of the observed network, we performed random simulations of 100,000 iterations for genes and their connections selected from the STRING database. To further research the function of the genes in the PPI network, biological processes (BPs) analysis of gene functional analysis was conducted on the David website[1].

RESULTS

Clinical Presentation

The index patient, a 5-year-old girl from Han Chinese population, was the first child born to healthy, non-consanguineous, Chinese parents. The patient was referred to the Second Affiliated Hospital and Yuying Children's Hospital of Wenzhou Medical University due to nasosinusitis and adenoid hypertrophy. Physical examination on admission demonstrated a disproportionate, short stature with a collapsed nose, spherical nose tip, short neck and a protruding abdomen. Multiple pigmented nevi were present on the abdomen and bilateral ankle. According to these clinical features, it was difficult to make a diagnosis to explain these abnormal symptoms. Therefore, after the consent of the patient's parents, WES was used to perform genetic testing to detect possible genetic lesions.

After the genetic testing, the child patient was also hospitalized to complete a full physical examination for more phenotype features. Skeletal X-ray showed scoliosis (**Figure 1A**), hypoplasia pelivis (**Figure 1B**), and spondyloepiphyseal dysplasia with ovoid and flat vertebrate (**Table 1**). Although a quantitative test of 24 h urinary protein was 120.4 mg/L, which was in the normal range (<150 mg/24 h), renal dysfunction was evidence with an increased content of urinary microproteins, including microalbuminuria 5.82 mg/dl (<1.9 mg/dl), transferrin 0.45 mg/dl (<0.2 mg/dl), and IgG 1.13 mg/dl (<0.8 mg/dl) (Pomerance, 1997) (**Table 1**). Hematological and immunological studies showed a decreased ratio of lymphocytes 0.167% (0.23–0.53%), along with an increased ratio of neutrophils 0.724% (0.35–0.65%), and monocytes 0.10 (0.03–0.08%). The T lymph ratio was 43% (55–84%), indicating T cell immunodeficiency, including low levels of T helper and T suppressor lymphocytes (**Table 1**). Furthermore, dental examination found enamel dysplasia. The thyroid stimulating hormone content in her serum was 6.0067 mIU/L (0.8–5.0 mIU/L), confirming the diagnosis of hypothyroidism. Nevertheless, electrocardiogram, abdominal ultrasound, and skull MRI plain scans showed no abnormalities.

Sequencing Data Analysis and Identification of Mutations

Although the preliminary examinations showed disproportionate short stature, special facial dysmorphism, hypothyroidism and multiple pigmented nevi, based only on these observations, it was hard to make a diagnosis. Considering the possibility of genetic defects, WES of the patient was performed. After rigorous

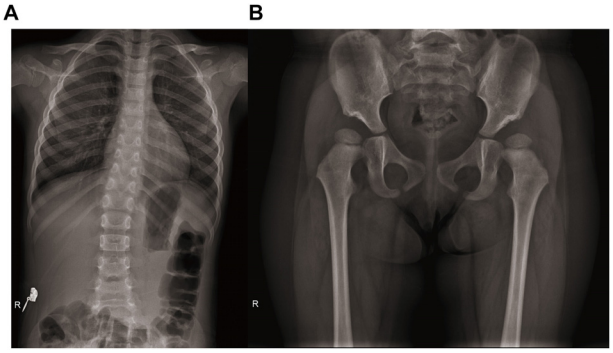

FIGURE 1 | Radiographs of spine and hip joints. **(A)** The spine radiograph showing the patient has scoliosis. **(B)** Bilateral femoral head moved upward. Bilateral acetabular articular surface was flat and shallow. Bilateral Shen Tong's line was discontinuous. Poor alignment of left hip joint.

[1]https://david.ncifcrf.gov/summary.jsp

TABLE 1 | Clinical features of the patient.

Clinical features		Current situation	Characterization and indicators
Development	Intrauterine growth retardation	+	Born at 36 weeks, weight of 1.25 kg, height of 40 cm (<3rd percentile)
	Microsomia	+	103.5 cm at 6-years old
Skeletal feature	Short neck	+	
	Short torso	+	
	Scoliosis	+	
	Oval flat vertebra	+	
	Pelvic hypoplasia	+	Bilateral femoral head moved upward, bilateral acetabular articular surface was flat and shallow, bilateral Shen Tong's line was discontinuous. Poor alignment of left hip joint
	Abnormal femoral head	+	
Kidney dysfunction	Urine protein	+	MA: 5.82 mg/dl (< 1.9 mg/dl)
	Kidney disease	−	
	Focal segmental glomerulosclerosis	−	
Abnormity of blood	T cell deficiency	+	CD3: 17%(62.0–70.0%), CD4: 11% (30.0–40.0%), CD8: 4% (20–27%)
	Lymphopenia	+	LY#: 1.00 × 109/L (0.92–5.3), LY%: 0.167 (0.23–0.53)
	Neutrophil reduction	−	NE#: 4.33 × 10^9/ (1.4–6.5), NE%: 0.724 (0.35–0.65)
	Thrombocytopenia	−	PLT: 350 × 10^9/L
	Anemia	−	HGB: 135 g/L (110–190)
Physical characteristics	Broad nose	+	
	Wide and collapsed bridge of the nose	+	
	Prominent belly	+	
	Multiple pigmented nevi	+	
	Abnormity of hair	+	
	Small or missing teeth	+	
	Corneal opacity	+	A small amount of granular gray-white reflection in the double corneal stroma is located in the posterior stromal layer of the bismuth.
Development	Hypoevolutism	−	
	Academic delay	−	
Vasculature	Headache	−	
	Transient cerebral ischemia stroke		No obvious abnormality of Skull MRI scan
Other inspections	Hypothyroidism	+	TSH:6.0067 mIU/L (0.8–5.0)
	Catarrhal dysentery	−	Fat globule:(+)
	Non-Hodgkin lymphoma	−	

MA, microalbuminuria; LY, lymphocyte; NE, neutrophile granulocyte; PLT, blood platelet; HGB, hemoglobin; TSH, thyroid stimulating hormone. +, positive for the finding; −, negative for the finding.

screening of mutations and Sanger sequencing validation, a known exon mutation (c.1933C > T, p.Arg645Cys) and a new intron mutation (c.2141+5G > A) adjacent to a splicing region (**Figure 2A**) were identified in the *SMRACAL1* gene. Haplotype analysis and sequencing showed that the c.1933C > T mutation was inherited from father and that the c.2141+5G > A mutation was from mother (**Figure 2B**).

The previously reported mutation Arg645Cys was predicted to be pathogenic by four commonly used bioinformatic algorithms, including SIFT (damaging), polyphen-2 (probably damaging), LRT (deleterious), and CADD (damaging). In addition, the amino acid residue was predicted to be conserved. Meanwhile, this variant has shown a low frequency of 8.25×10^{-6} in the ExAC database and was absent in the UK10K, 1000 Genomes and ESP6500. In accordance with the recommended standards of the American College of Medical Genetics (ACMG) (Richards et al., 2015), the variant c.1933C > T is categorized as a "pathogenic variant" because of its classification as PS1, and PS3 (**Supplementary Table S2**). According to the annotation, the c.2141+5G > A mutation has not been previously observed in any public database.

The Novel c.2141+5G > A Allele Affects the Splicing of *SMARCAL1*

The c.2141+5G > A site was adjacent to the 5′ donor splice site in intron 13 (**Figure 2A** and **Supplementary Figure S1**) and was confirmed in the patient and her mother. Though it was a variant of an intron region, we hypothesized that it was the second allele as a compound mutation that caused SIOD, an autosomal recessive multisystem disorder. The mutation was not previously observed in various publicly available databases, including ExAC, UK10K, dbSNP147, 1000 Genomes, or the ESP6500 database. We used four splicing prediction tools dbscSNV (Jian et al., 2014), Human Splicing Finder (Desmet et al., 2009), MaxEntScan (Yeo and Burge, 2004), and SPANR

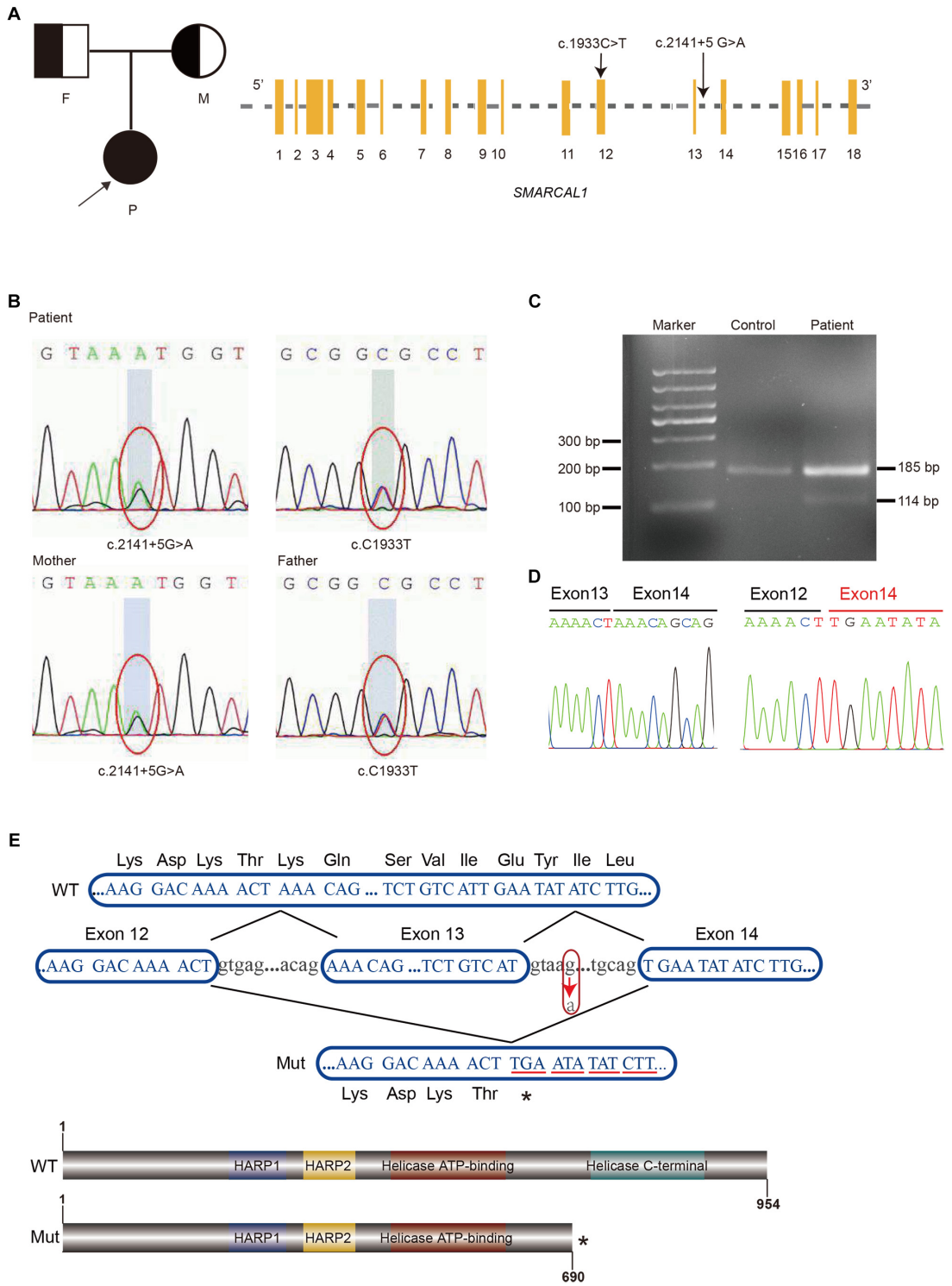

FIGURE 2 | Identification of compound heterozygous mutations in *SMARCAL1*. **(A)** Pedigree of the Family and Schematic of *SMARCAL1*. Yellow boxes represent the coding exons; black characters indicate the mutation identified in this study. **(B)** Sanger sequencing confirmed the heterozygous mutations in *SMARCAL1*. The father was a carrier of the c.C1933T mutation and the mother was a carrier of the c.2141+5G > A mutation. **(C)** Reverse transcriptase PCR analysis on cDNAs from patient and control. An abnormal PCR product (114 bp) was detected in patient. **(D)** Sanger sequencing for the fragment amplified from the cDNAs of both patient and control. **(E)** Nucleotide sequences from exon 12 to exon 14 of *SMARCAL1* and protein schematic of smarcal1. WT, wild type; Mut, mutant. The arrow indicates the c.2141+5G > A variant. The asterisk means a premature stop codon.

(Elliott et al., 1976) to predict splicing changes, which resulted in the creation of a novel splice donor site (**Supplementary Table S3**). Splicing mutations can lead to a mixture of aberrantly and correctly spliced transcripts by partial skipping of exons or inclusion of intronic sequences, or they can change the ratio of programmed alternatively spliced isoforms (Nissim-Rafinia and Kerem, 2005). The position of the mutation immediately adjacent to the 5′ donor splice site suggested it might disrupt 5′-splice site function, result in the skipping of contiguous exon from the mature mRNA and lead to a truncated protein (Montalban et al., 2018).

The exact impact of the c.2141+5 G > A variant on splicing was confirmed by RT-PCR with total RNA extracted from the blood samples from the patient and control, using forward and reverse primers located in exons 12 and 14 respectively. Two different amplification products were detected from the *SMRACAL1* cDNA of the patient: the upper band derived from the full-length transcript which was corresponded to the wild-type allele (**Figure 2C**), whereas the lower band derived from an aberrant transcript lacking the entire exon 13 (c.2070_2141de71) (**Figure 2C**). This exon skipping lead to premature stop codon at position 260 that results in the loss of 247 amino acids from the SMARCAL1 C-terminus, resulting in the loss of Helicase C-terminal functional region (**Figures 2D,E**). According to the guide of ACMG (Richards et al., 2015), the novel variant c.2141+5G > A would be classified as a "pathogenic variant" supported by the classification as PVS1, PM2, and PM3.

Phenotypic Profile and Mutation Summary

To gain a global understanding of the various disease phenotypes and mutations, we collected all of the previously reported cases based on abundant literature. SIOD has a marked variation in severity, ranging from in utero onset with growth retardation and death within the first 5 years of life to onset of symptoms in late childhood. Based on existing phenotypic data, most of the patients had symptoms of spondyloepiphyseal dysplasia (97.5%), T lymphocyte decrease (85.96%) and maculas (62.67%) (**Supplementary Table S1**), while some of the patients diagnosed severe SIOD had blood pancytopenia, recurrent infections and central nervous symptoms (**Supplementary Table S1**).

Schimke immuno-osseous dysplasia is a genetically heterogeneous disease. The pathogenicity of the mutation was usually considered to cause a change in protein function. Therefore, we mapped 49 of the reported amino acids alternative mutations to the schematic representations of SMARCAL1 (**Figure 3**). Helicase ATP-binding and helicase C-terminal domains showed excessive enrichment of mutations. The Arg645 site tends to be a recurrent susceptible locus of SMARCAL1 with diverse amino acid changes, and the other biallelic site (c.2141+5G) is close to exon 13, which encodes a part of the helicase C-terminal subdomain. Besides, we also mapped all the pathogenic mutations in intron regions to the nucleotide structure map of *SMARCAL1* (**Supplementary Figure S1**), among which a reported homozygousmutation c.2244+5 G > A in intron 14 detected in a SIOD patient with mild developmental delay and mild ID was near the novel c.2141+5G > A mutation in this study.

Notably, the variants at residue Arg645 was previously reported as a compound mutation in four SIOD patients (**Table 2**). All the patients were diagnosed with mild SIOD based on mild multisystem clinical symptoms, while the definite diagnosis of our patient was only made after genetic testing. Compared to the four patients that carried missense mutations at the same locus, our patient had milder symptoms without any neurological signs and recurrent infections. Remarkably, the renal dysfunctions of our patient were limited to asymptomatic proteinuria while the kidney

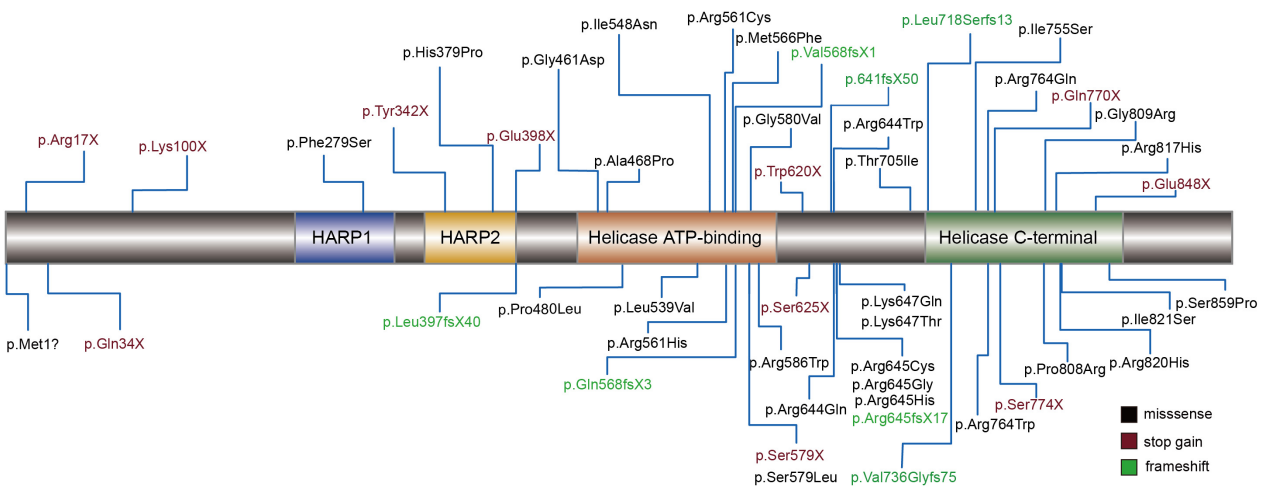

FIGURE 3 | Schematic diagram of mutations assiociated with SIOD in the encoded proteins of *SMARCAL1* gene. Rectangles with different colors indicate specific protein domains. Black font represents missense mutation, vermeil font shows stop-gain mutations, and green font indicates frameshift mutations.

TABLE 2 | Genotype–phenotype analysis of patients with mutations at the same site R645.

Patient ID	1(Lipska-Zietkiewicz et al., 2017)	2-I(Liu et al., 2017)	2-II(Liu et al., 2017)	3(Zivicnjak et al., 2009)	Our patient
Gender	M	F	F	M	M
Country of origin	China	Germany	Germany	France	China
Nucleotide change	c.1933C > T; c.2450G > A	c.1934G > A; c.2542G > T	c.1934G > A; c.2542G > T	c.1933C > G; c.2425G > A	c.1933C > T; c.2141+5G > A
Protein change	p.R645C; p.R817H	p.R645H; p.E848X	p.R645H; p.E848X	p.R645G; p.G809R	p.R645C
Diagnosis	Mild SIOD	Mild SIOD	Mild SIOD	Mild SIOD	Mild SIOD
Diagnosis year	6	11	16	10.8	5
Diagnostic basis	Phenotype	Phenotype	Phenotype	Phenotype	WES
Age at 1st manifestation (years)	5.7	4	6	?	4
1st manifestation phenotype	Proteinuria	Steroid-resistant nephrotic syndrome	Steroid-resistant nephrotic syndrome	?	Hyperthyreosis
Renal function	IgM nephropathy	Renal failure	Terminal renal failure	Renal failure	Asymptomatic proteinuria
Histopathology	Glomerular segmental mesangial matrix hyperplasia	Absence of foot processes and minimal changes in 18 glomeruli	Absence of foot processes and minimal changes in 18 glomeruli	FSGS	Normal
Neurological signs and symptoms	Normal	Normal	Psychosocial problems and reactive depression	Ischemic stroke, demyelinating peripheral neuropathy, epilepsy	Normal
Infections, immunodeficiency	Lymphocytoponia, congenital immune deficiency, decreased blood IgG values	Cyclic lymphopenia, CMV, varicella zoster, helicobacter, and enterococci infections, enterobacteriacea sepsis complicated	High EBV viral load, ocular varicellazoster and herpes encephalitis, sepsis	Catheter-related sepsis SCID	No recurrent infections, decreased ratio of lymphocytes, T cell immunodeficiency
Other extra-renal signs	Puffy eyelids and edema of the lower extremities	Cardiac arrest, papillomas hands	Severe, therapy-resistant papilloma on face, feet and hands	Multiple pigmented nevi, IUGR, preterm delivery	Papilloma on abdomen and bilateral ankle
Thyroid function	Normal	?	?	Normal	Hyperthyreosis
Hypoevolutism	95 cm at 6-years-old (<3rd percentile)	Severe height impairment	Severe height impairment	Height impairment	Short stature
Scoliosis	Scoliosis	Progressive hip dysplasia	Progressive hip dysplasia	Femoral head necrosis	Scoliosis

FSGS, focal glomerular sclerosis; CMV, cytomegalovirus; EBV, Epstein–Barr virus; SCID, severe combined immune deficiency; IUGR, intrauterine growth retardation.

histopathological examination of the other four patients showed varying degrees of abnormalities. In addition to the mildest symptoms, the diagnostic age of our patient was the youngest.

PPI Network, BP Enrichment Analysis, and Renal Expression Profile of *SMARCAL1*-Interacted Genes

Functional interactions between proteins are vital during cellular biological processing and their systematic characterization provides a background of the molecular systems biology. To investigate the PPI network of *SMARCAL1*, we applied relevant data from the STRING v10 database (Szklarczyk et al., 2015). As a result, a total of 73 genes were included in the interconnected PPI network, with an cut-off interaction score greater than 400 (**Figure 4A**). The PPI network with varying degrees of interactions showed an excess number of interacting proteins among those encoded by the 73 genes ($P = 1 \times 10^{-5}$) and an excess of pairwise connections ($P = 1 \times 10^{-5}$) compared with 100,000 randomly simulated datasets (**Figure 4A**). In this network, two genes *RPA2* and *RAP1* which encode subunits of the heterotrimeric replication protein A (RPA) complex showed the strongest interaction with *SMARCAL1*, with an interaction score of 997 and 974 respectively. In addition, genes encoding ATR–ATRIP protein kinase complex which is crucial for the cellular response to replication stress and

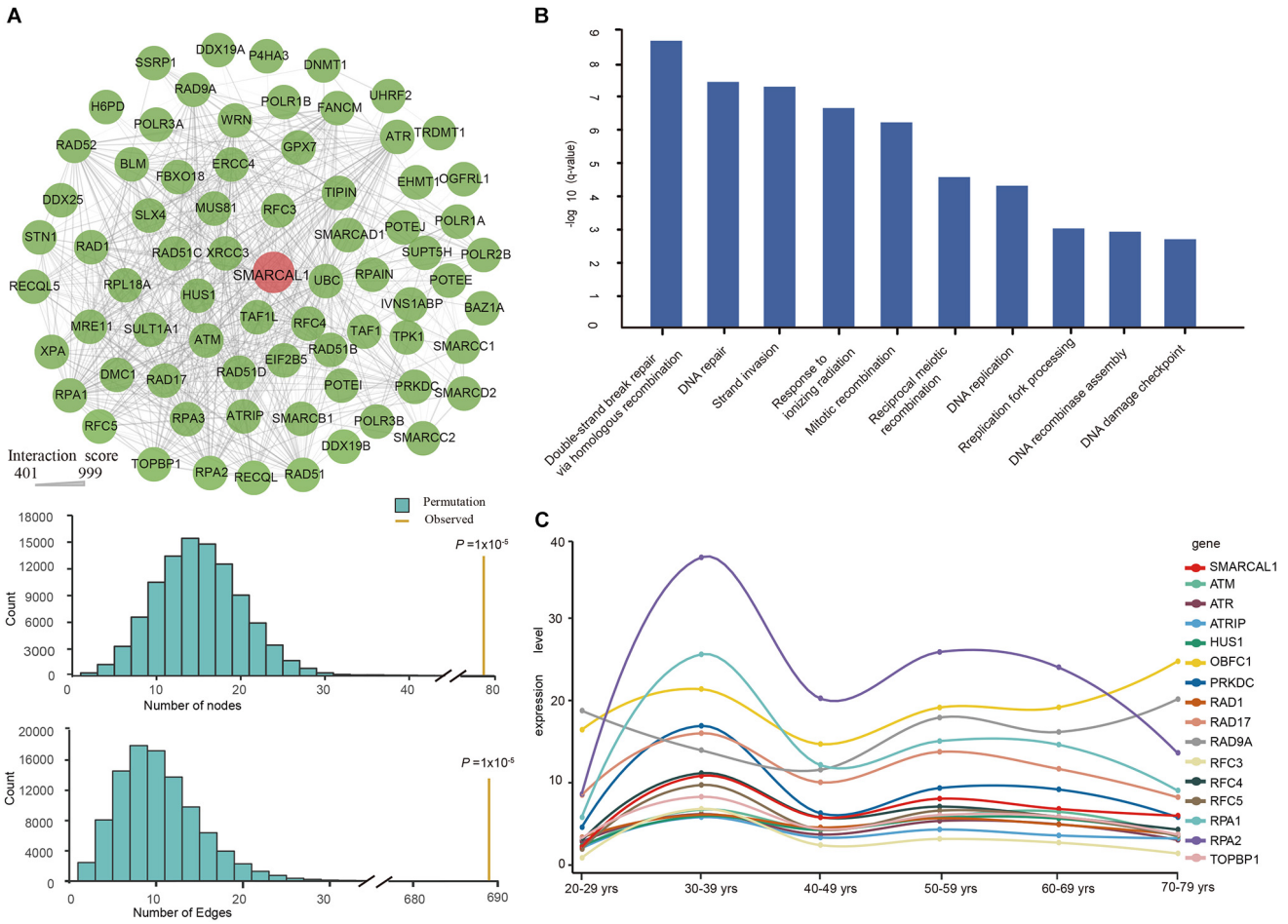

FIGURE 4 | Protein–protein interaction (PPI) Network, Gene Ontology enrichment analysis, and renal expression profile of *SMARCAL1*-interacted genes. **(A)** Interconnected PPI network of the 73 genes with interaction score more than 400 with *SMARCAL1*. A permutation test for genes and connections was conducted with 100,000 iterations. Nodes denote genes, and edges denote interactions between two genes. The thickness of an edge denotes the interaction scores between gene pairs. **(B)** The BP enrichment analysis of the 73 genes in the PPI network. The x-axis shows the GO terms, and the y-axis represents the value of –log10 (*q*-value), indicating the enrichment scores of the GO terms. All *P*-values are corrected for multiple testing using FDR. **(C)** Renal expression profile of 15 genes with the strongest genetic interaction effect with *SMARCAL1*. Different color represents special gene, while red indicates *SMARCAL1*.

DNA damage were also found to have tightly connection to *SMARCAL1*.

To further understand the biological function of these 73 genes, we performed an enrichment analysis of biological processes by DAVID Bioinformatics Resources. The enrichment results showed multiple significant terms with FDR values < 0.05, such as "double-strand break repair via homologous recombination (GO:0000724)," "DNA repair (GO:0006281)," and "replication fork processing (GO:0031297)" (**Figure 4B**).

Defining the normal expression of *SMARCAL1* gene is an essential first step in determining the cells that contribute to the pathogenesis of FSGS in SIOD patients. To detect the expression profile of *SMARCAL1*-interacting genes in the kidney, we extracted relevant transcriptional data from the GTEx database. The top 15 *SMARCAL1*-interacting genes except *RFC5* showed a similar renal expression profile to *SMARCAL1* (**Figure 4C**).

DISCUSSION

Schimke immuno-osseous dysplasia is a rare autosomal recessive disorder caused by compound mutations of *SMARCAL* (Boerkoel et al., 2002) characterized by dysmorphism (Saraiva et al., 1999; Boerkoel et al., 2000), spondyloepiphysial dysplasia (Spranger et al., 1991; Lucke et al., 2006), T cell immunodeficiency (Boerkoel et al., 2000; Lucke et al., 2006), and nephrotic syndrome (Boerkoel et al., 2000; Lucke et al., 2006). SIOD shows phenotypic heterogeneity from mild to severe phenotypes (Simon et al., 2014) and the severity is proportionate to the degree of *SMARCAL1* inactivation (Elizondo et al., 2009). Infantile-onset, early lethal severe patients would carry at least one *SMARCAL1* loss-of-function mutation (deletion, nonsense, or frameshift), whereas patients with a less severe phenotype and survival into the second 10 years would likely carry compound missense mutations

(Lou et al., 2002). However, Elizondo et al. reported missense mutations that alter subcellular location, enzymatic activity, protein levels or chromatin binding could result in severe phenotypes (Elizondo et al., 2009). From all of previously reported patients (summarized in **Supplementary Table S1**), we found that severe SIOD could arise from missense mutations, while frameshift and truncation mutations could lead to a mild phenotype, further suggesting that the genotype–phenotype correlations remain largely unpredictable (Elizondo et al., 2009).

Compared to patients initially diagnosed as SIOD mainly based on a characteristic phenotype, those with SIOD probands diagnosed unexpectedly by genetic screening (SRNS-gene panel) showed a milder extrarenal phenotype during follow-up and a better 10-year patient survival (Lipska-Zietkiewicz et al., 2017). It is known that WES can effectively identify the pathogenic mutations for those whose clinical symptoms are too mild to diagnose the disease, especially rare diseases (Wu et al., 2012; Sarin et al., 2015; Krabbenborg et al., 2016). Notably, the extremely mild and non-characteristic phenotype of our patient did not allow for a definite diagnosis, and thus our patient was the first diagnosed based on a genetic test applying WES before the disease was confirmed. The only previously reported SIOD patient employing WES as genetic testing was conducted after an initial phenotype-based diagnosis accompanied with severe neurodevelopmental delay, facial dysmorphism, and spondyloepiphyseal bone dysplasia (Simon et al., 2014). Our study proves once again that WES is a high-performance tool in the early diagnosis of the extremely rare diseases.

Interestingly, four additional SIOD patients carried missense mutations at the same residue Arg645 (e.g., Arg645Cys, Arg645His, Arg645Gly) (Zivicnjak et al., 2009; Lipska-Zietkiewicz et al., 2017; Liu et al., 2017) accompanied by another different mutation, and two of these were male siblings from a German family (Zivicnjak et al., 2009) (**Table 2**). All four patients were diagnosed as mild type according to the combined consideration of phenotypes. When reviewing their medical history, an obvious organic lesion of the kidney was observed in the four patients before diagnosis. In contrast to those, our affected patient exhibited asymptomatic proteinuria without recurrent infections or neurological signs and symptoms. Furthermore, the phenotype of the five patients differ from each another, including the two siblings, suggesting sensitive phenotypic pleiotropy. The same *SMARCAL1* mutations have been discovered in both mild and severe cases of the disease, and evidence of environmental factors shaping SIOD phenotypes in fruit flies, mice, and humans (Lucke et al., 2005; Baradaran-Heravi et al., 2012) has been found. Thus, the effect of a *SMARCAL1* mutation cannot only be determined by the *SMARCAL1* genotype but also by the stochastic, environmental, and genetic background variations (Lucke et al., 2005; Carroll et al., 2013).

The patient described in this study harbored a novel splicing mutation (c.2141+5G > A) confirmed to be an alternate donor splice site mutation, resulting in the loss of C-terminal functional region of SMARCAL1 protein. Moreover, the novel mutation was classified as a "pathogenic variant" according to the guide of ACMG. Similarly, a reported SIOD patient with mild developmental delay and mild ID carried a homozygous mutation c.2244+5 G > A (Lipska-Zietkiewicz et al., 2017), within the donor site of intron 14. Both the two splicing sites are very close to the initial amino acid encoding the helicase C-terminal region, in which most disease-related mutations are accumulated. Phosphorylation of subdomain in the C-terminus can change SMARCAL1 activity, resulting in either overexpression or silencing of *SMARCAL1* that then causes the accumulation of replication-associated DNA damage (Carroll et al., 2014).

SMARCAL1 showed the strongest interaction with genes encoding RAP complex and ATR–ATRIP protein kinase complex in the PPI network (**Figure 4A**). SMARCAL1 is an ATP-driven annealing helicase that catalyzes the formation of double-stranded DNA from complementary single-strand DNA strands by binding to *RPA* (Yusufzai and Kadonaga, 2008; Yusufzai et al., 2009). RPA complex, which is involved in DNA replication, repair, recombination, telomere maintenance, and coordinating the cellular response to DNA damage, plays an important role in DNA metabolism (Liaw et al., 2011; Daniels et al., 2012). The recruitment of ATR–ATRIP complex at sites of DNA damage is based on RPA-coated ssDNA, which facilitates complex recognition of substrates for phosphorylation and the initiation of checkpoint signaling (Zou and Elledge, 2003). In the developing fetal kidney, SMARCAL1 is expressed in the ureteric epithelium, stroma, metanephric mesenchyme, and in all stages of the developing nephron, including the maturing glomerulus, suggesting an important, functional role for the SMARCAL1 helicase during kidney development (Dekel et al., 2008). The absence or disruption of SMARCAL1 protein function in SIOD patients might cause abnormalities in kidney development that manifested during the perinatal or adolescent period. In the postnatal adult kidney, SMARCAL1 is expressed in the glomerulus, tubules of the nephron, and collecting ducts, demonstrating a role for SMARCAL1 protein in the maintenance and integrity of podocytes and endothelial cells (Sarin et al., 2015). Most of the genes showed strongest interaction with *SMARCAL1* showing a similar renal mRNA expression profile to *SMARCAL1* that indicates a similar role these genes play in renal function maintenance.

CONCLUSION

We reported a 5-year-old girl with an extremely mild phenotype based on genetic testing by applying WES. We identified a new allelic mutation (c.2141+5G > A) of *SMARCAL1,* which is confirmed to create a novel splice donor site. Our results provide new insights into the phenotypic spectrum of *SMARCAL1*. The biological function and renal expression analyses of *SMARCAL1*-interacting genes provide additional genetic evidence supporting the potential role of *SMARCAL*1 in SIOD.

ETHICS STATEMENT

This study was conducted in accordance with the guidelines of the Declaration of Helsinki. It was also approved by the Ethics Committee of the Second Affiliated Hospital and Yuying Children's Hospital of Wenzhou Medical University. Written informed consent of all participants was collected from their parents or guardians at the time of recruitment for participation in the study and publication of the study.

AUTHOR CONTRIBUTIONS

JJ was involved in the study design, data analysis, and manuscript drafting. ZL and XC performed exome sequencing. SJ performed Sanger sequencing. KW participated in the data analysis and performed clinical evaluation of the patients. ZW and WL were involved in the study design and critical evaluation of the manuscript. All authors read and approved the final manuscript.

FUNDING

This study was supported by the National Natural Science Foundation of China (Grant Nos. 81772664 and 81501824).

ACKNOWLEDGMENTS

We thank the family for participating in this study. We also thank Prof. Jinyu Wu and Feng Gu (Wenzhou Medical University) for their critical and insightful suggestions about our work.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2019.00565/full#supplementary-material

REFERENCES

Adzhubei, I. A., Schmidt, S., Peshkin, L., Ramensky, V. E., Gerasimova, A., Bork, P., et al. (2010). A method and server for predicting damaging missense mutations. *Nat. Methods* 7, 248–249. doi: 10.1038/nmeth0410-248

Bansbach, C. E., Boerkoel, C. F., and Cortez, D. (2010). SMARCAL1 and replication stress: an explanation for SIOD? *Nucleus* 1, 245–248. doi: 10.4161/nucl.1.3.11739

Baradaran-Heravi, A., Cho, K. S., Tolhuis, B., Sanyal, M., Morozova, O., Morimoto, M., et al. (2012). Penetrance of biallelic SMARCAL1 mutations is associated with environmental and genetic disturbances of gene expression. *Hum. Mol. Genet.* 21, 2572–2587. doi: 10.1093/hmg/dds083

Boerkoel, C. F., O'Neill, S., Andre, J. L., Benke, P. J., Bogdanovic, R., Bulla, M., et al. (2000). Manifestations and treatment of schimke immuno-osseous dysplasia: 14 new cases and a review of the literature. *Eur. J. Pediatr.* 159, 1–7.

Boerkoel, C. F., Takashima, H., John, J., Yan, J., Stankiewicz, P., Rosenbarker, L., et al. (2002). Mutant chromatin remodeling protein SMARCAL1 causes schimke immuno-osseous dysplasia. *Nat. Genet.* 30, 215–220. doi: 10.1038/ng821

Carroll, C., Badu-Nkansah, A., Hunley, T., Baradaran-Heravi, A., Cortez, D., and Frangoul, H. (2013). Schimke immunoosseous dysplasia associated with undifferentiated carcinoma and a novel SMARCAL1 mutation in a child. *Pediatr. Blood Cancer* 60, E88–E90. doi: 10.1002/pbc.24542

Carroll, C., Bansbach, C. E., Zhao, R., Jung, S. Y., Qin, J., and Cortez, D. (2014). Phosphorylation of a C-terminal auto-inhibitory domain increases SMARCAL1 activity. *Nucleic Acids Res.* 42, 918–925. doi: 10.1093/nar/gkt929

Chun, S., and Fay, J. C. (2009). Identification of deleterious mutations within three human genomes. *Genome Res.* 19, 1553–1561. doi: 10.1101/gr.092619.109

Ciccia, A., Bredemeyer, A. L., Sowa, M. E., Terret, M. E., Jallepalli, P. V., Harper, J. W., et al. (2009). The SIOD disorder protein SMARCAL1 is an RPA-interacting protein involved in replication fork restart. *Genes Dev.* 23, 2415–2425. doi: 10.1101/gad.1832309

Daniels, J. P., Gull, K., and Wickstead, B. (2012). The trypanosomatid-specific N terminus of RPA2 is required for RNA polymerase I assembly, localization, and function. *Eukaryot. Cell* 11, 662–672. doi: 10.1128/EC.00036-12

Dekel, B., Metsuyanim, S., Goldstein, N., Pode-Shakked, N., Kovalski, Y., Cohen, Y., et al. (2008). Schimke immuno-osseous dysplasia: expression of SMARCAL1 in blood and kidney provides novel insight into disease phenotype. *Pediatr. Res.* 63, 398–403. doi: 10.1203/PDR.0b013e31816721cc

Desmet, F. O., Hamroun, D., Lalande, M., Collod-Beroud, G., Claustres, M., and Beroud, C. (2009). Human splicing finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Res.* 37:e67. doi: 10.1093/nar/gkp215

Ehrich, J. H., Burchert, W., Schirg, E., Krull, F., Offner, G., Hoyer, P. F., et al. (1995). Steroid resistant nephrotic syndrome associated with spondyloepiphyseal dysplasia, transient ischemic attacks and lymphopenia. *Clin. Nephrol.* 43, 89–95.

Elizondo, L. I., Cho, K. S., Zhang, W., Yan, J., Huang, C., Huang, Y., et al. (2009). Schimke immuno-osseous dysplasia: smarcal1 loss-of-function and phenotypic correlation. *J. Med. Genet.* 46, 49–59. doi: 10.1136/jmg.2008.060095

Elliott, M., Janes, N. F., Pulman, D. A., Gaughan, L. C., Unai, T., and Casida, J. E. (1976). Radiosynthesis and metabolism in rats of the 1R isomers of the insecticide permethrin. *J. Agric. Food Chem.* 24, 270–276. doi: 10.1021/jf60204a007

Jian, X., Boerwinkle, E., and Liu, X. (2014). In silico prediction of splice-altering single nucleotide variants in the human genome. *Nucleic Acids Res* 42, 13534–13544. doi: 10.1093/nar/gku1206

Kircher, M., Witten, D. M., Jain, P., O'Roak, B. J., Cooper, G. M., and Shendure, J. (2014). A general framework for estimating the relative pathogenicity of human genetic variants. *Nat. Genet.* 46, 310–315. doi: 10.1038/ng.2892

Krabbenborg, L., Vissers, L. E., Schieving, J., Kleefstra, T., Kamsteeg, E. J., Veltman, J. A., et al. (2016). Understanding the psychosocial effects of wes test results on parents of children with rare diseases. *J. Genet. Couns.* 25, 1207–1214. doi: 10.1007/s10897-016-9958-5

Kumar, P., Henikoff, S., and Ng, P. C. (2009). Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat. Protoc.* 4, 1073–1081. doi: 10.1038/nprot.2009.86

Li, J., Jiang, Y., Wang, T., Chen, H., Xie, Q., Shao, Q., et al. (2015). mirTrios: an integrated pipeline for detection of de novo and rare inherited mutations from trios-based next-generation sequencing. *J. Med. Genet.* 52, 275–281. doi: 10.1136/jmedgenet-2014-102656

Liaw, H., Lee, D., and Myung, K. (2011). DNA-PK-dependent RPA2 hyperphosphorylation facilitates DNA repair and suppresses sister chromatid exchange. *PLoS One* 6:e21424. doi: 10.1371/journal.pone.0021424

Lipska-Zietkiewicz, B. S., Gellermann, J., Boyer, O., Gribouval, O., Zietkiewicz, S., Kari, J. A., et al. (2017). Low renal but high extrarenal phenotype variability in Schimke immuno-osseous dysplasia. *PLoS One* 12:e0180926. doi: 10.1371/journal.pone.0180926

Liu, S., Zhang, M., Ni, M., Zhu, P., and Xia, X. (2017). A novel compound heterozygous mutation of the SMARCAL1 gene leading to mild Schimke immune-osseous dysplasia: a case report. *BMC Pediatr.* 17:217. doi: 10.1186/s12887-017-0968-8

Lou, S., Lamfers, P., McGuire, N., and Boerkoel, C. F. (2002). Longevity in schimke immuno-osseous dysplasia. *J. Med. Genet.* 39, 922–925. doi: 10.1136/jmg.39.12.922

Lucke, T., Billing, H., Sloan, E. A., Boerkoel, C. F., Franke, D., Zimmering, M., et al. (2005). Schimke-immuno-osseous dysplasia: new mutation with weak genotype-phenotype correlation in siblings. *Am. J. Med. Genet. A* 135, 202–205. doi: 10.1002/ajmg.a.30691

Lucke, T., Franke, D., Clewing, J. M., Boerkoel, C. F., Ehrich, J. H., Das, A. M., et al. (2006). Schimke versus non-schimke chronic kidney disease: an anthropometric approach. *Pediatrics* 118, e400–e407. doi: 10.1542/peds.2005-2614

Montalban, G., Fraile-Bethencourt, E., Lopez-Perolio, I., Perez-Segura, P., Infante, M., Duran, M., et al. (2018). Characterization of spliceogenic variants located in regions linked to high levels of alternative splicing: BRCA2 c.7976+5G > T as a case study. *Hum. Mutat.* 39, 1155–1160. doi: 10.1002/humu.23583

Morimoto, M., Kerouredan, O., Gendronneau, M., Shuen, C., Baradaran-Heravi, A., Asakura, Y., et al. (2012). Dental abnormalities in schimke immuno-osseous dysplasia. *J. Dent. Res.* 91(7 Suppl.), 29S–37S. doi: 10.1177/0022034512450299

Morimoto, M., Lewis, D. B., Lucke, T., and Boerkoel, C. F. (1993). "Schimke immunoosseous dysplasia," in *GeneReviews((R))*, eds M. P. Adam, H. H. Ardinger, R. A. Pagon, S. E. Wallace, L. J. H. Bean, K. Stephens, et al. (Seattle WA: University of Washington).

Nissim-Rafinia, M., and Kerem, B. (2005). The splicing machinery is a genetic modifier of disease severity. *Trends Genet.* 21, 480–483. doi: 10.1016/j.tig.2005.07.005

Pomerance, H. H. (1997). Nelson textbook of pediatrics. *Arch. Pediatr. Adolesc. Med.* 151:324. doi: 10.1001/archpedi.1997.02170400110025

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., et al. (2015). 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.* 17, 405–424. doi: 10.1038/gim.2015.30

Santangelo, L., Gigante, M., Netti, G. S., Diella, S., Puteo, F., Carbone, V., et al. (2014). A novel SMARCAL1 mutation associated with a mild phenotype of Schimke immuno-osseous dysplasia (SIOD). *BMC Nephrol.* 15:41. doi: 10.1186/1471-2369-15-41

Saraiva, J. M., Dinis, A., Resende, C., Faria, E., Gomes, C., Correia, A. J., et al. (1999). Schimke immuno-osseous dysplasia: case report and review of 25 patients. *J. Med. Genet.* 36, 786–789. doi: 10.1136/jmg.36.10.786

Sarin, S., Javidan, A., Boivin, F., Alexopoulou, I., Lukic, D., Svajger, B., et al. (2015). Insights into the renal pathogenesis in schimke immuno-osseous dysplasia: a renal histological characterization and expression analysis. *J. Histochem. Cytochem.* 63, 32–44. doi: 10.1369/0022155414558335

Schimke, R. N., Horton, W. A., and King, C. R. (1971). Chondroitin-6-sulphaturia, defective cellular immunity, and nephrotic syndrome. *Lancet* 298, 1088–1089. doi: 10.1016/S0140-6736(71)90400-4

Simon, A. J., Lev, A., Jeison, M., Borochowitz, Z. U., Korn, D., Lerenthal, Y., et al. (2014). Novel SMARCAL1 bi-allelic mutations associated with a chromosomal breakage phenotype in a severe SIOD patient. *J. Clin. Immunol.* 34, 76–83. doi: 10.1007/s10875-013-9957-3

Spranger, J., Hinkel, G. K., Stoss, H., Thoenes, W., Wargowski, D., and Zepp, F. (1991). Schimke immuno-osseous dysplasia: a newly recognized multisystem disease. *J. Pediatr.* 119(1 Pt 1), 64–72. doi: 10.1016/s0022-3476(05)81040-6

Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., et al. (2015). STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* 43, D447–D452. doi: 10.1093/nar/gku1003

Wang, K., Li, M., and Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* 38:e164. doi: 10.1093/nar/gkq603

Wang, T., Liu, Q., Li, X., Wang, X., Li, J., Zhu, X., et al. (2013). RRBS-analyser: a comprehensive web server for reduced representation bisulfite sequencing data analysis. *Hum. Mutat.* 34, 1606–1610. doi: 10.1002/humu.22444

Wu, J., Shen, E., Shi, D., Sun, Z., and Cai, T. (2012). Identification of a novel Cys146X mutation of SOD1 in familial amyotrophic lateral sclerosis by whole-exome sequencing. *Genet. Med.* 14, 823–826. doi: 10.1038/gim.2012.50

Yeo, G., and Burge, C. B. (2004). Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *J. Comput. Biol.* 11, 377–394. doi: 10.1089/1066527041410418

Yue, Z., Xiong, S., Sun, L., Huang, W., Mo, Y., Huang, L., et al. (2010). Novel compound mutations of SMARCAL1 associated with severe schimke immuno-osseous dysplasia in a chinese patient. *Nephrol. Dial. Transplant.* 25, 1697–1702. doi: 10.1093/ndt/gfq071

Yusufzai, T., and Kadonaga, J. T. (2008). HARP is an ATP-driven annealing helicase. *Science* 322, 748–750. doi: 10.1126/science.1161233

Yusufzai, T., Kong, X., Yokomori, K., and Kadonaga, J. T. (2009). The annealing helicase HARP is recruited to DNA repair sites via an interaction with RPA. *Genes Dev.* 23, 2400–2404. doi: 10.1101/gad.1831509

Zivicnjak, M., Franke, D., Zenker, M., Hoyer, J., Lucke, T., Pape, L., et al. (2009). SMARCAL1 mutations: a cause of prepubertal idiopathic steroid-resistant nephrotic syndrome. *Pediatr. Res.* 65, 564–568. doi: 10.1203/PDR.0b013e3181998a74

Zou, L., and Elledge, S. J. (2003). Sensing DNA damage through ATRIP recognition of RPA-ssDNA complexes. *Science* 300, 1542–1548. doi: 10.1126/science.1083430

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Genetics

ORIGINAL RESEARCH
published: 25 June 2019
doi: 10.3389/fgene.2019.00425

Evaluation of Diagnostic Yield in Fetal Whole-Exome Sequencing: A Report on 45 Consecutive Families

Lior Greenbaum[1,2,3], Ben Pode-Shakked[1,3], Shlomit Eisenberg-Barzilai[1], Michal Dicastro-Keidar[1], Anat Bar-Ziv[1], Nurit Goldstein[1], Haike Reznik-Wolf[1], Hana Poran[1], Amihai Rigbi[4], Ortal Barel[5], Aida M. Bertoli-Avella[6], Peter Bauer[6], Miriam Regev[1,3], Annick Raas-Rothschild[1,3], Elon Pras[1,3] and Michal Berkenstadt[1,3]**

[1] The Danek Gertner Institute of Human Genetics, Sheba Medical Center, Tel Hashomer, Israel, [2] The Joseph Sagol Neuroscience Center, Sheba Medical Center, Tel Hashomer, Israel, [3] Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel, [4] Faculty of Education, Beit Berl College, Kfar Saba, Israel, [5] Sheba Cancer Research Center, Sheba Medical Center, Tel Hashomer, Israel, [6] Centogene AG, Rostock, Germany

OPEN ACCESS

Edited by:
Fan Jin,
Zhejiang University, China

Reviewed by:
Nelson L. S. Tang,
The Chinese University of Hong Kong, China
Qing Zhou,
Zhejiang University, China

***Correspondence:**
Michal Berkenstadt
m.berken@sheba.health.gov.il
Lior Greenbaum
lior.greenbaum@sheba.health.gov.il

Specialty section:
This article was submitted to Genetic Disorders, a section of the journal Frontiers in Genetics

Received: *17 November 2018*
Accepted: *17 April 2019*
Published: *25 June 2019*

Citation:
Greenbaum L, Pode-Shakked B, Eisenberg-Barzilai S, Dicastro-Keidar M, Bar-Ziv A, Goldstein N, Reznik-Wolf H, Poran H, Rigbi A, Barel O, Bertoli-Avella AM, Bauer P, Regev M, Raas-Rothschild A, Pras E and Berkenstadt M (2019) Evaluation of Diagnostic Yield in Fetal Whole-Exome Sequencing: A Report on 45 Consecutive Families. Front. Genet. 10:425. doi: 10.3389/fgene.2019.00425

Prenatal ultrasound (US) abnormalities often pose a clinical dilemma and necessitate facilitated investigations in the search of diagnosis. The strategy of pursuing fetal whole-exome sequencing (WES) for pregnancies complicated by abnormal US findings is gaining attention, but the reported diagnostic yield is variable. In this study, we describe a tertiary center's experience with fetal WES from both terminated and ongoing pregnancies, and examine the clinical factors affecting the diagnostic rate. A total of 45 consecutive families of Jewish descent were included in the analysis, for which clinical fetal WES was performed under either single (fetus only), trio (fetus and parents) or quatro (two fetuses and parents) design. Except one, all families were non-consanguineous. In 41 of the 45 families, WES was sought following abnormal fetal US findings, and 18 of them had positive relevant family history (two or more fetuses with US abnormalities, or single fetus with US abnormalities and an affected parent). The overall diagnostic yield was 28.9% (13/45 families), and 31.7% among families with fetal US abnormalities (13/41). It was significantly higher in families with prenatal US abnormalities and relevant family history (10/18, 55.6%), compared to families with prenatal US abnormal findings and lack of such history (3/23, 13%) ($p = 0.004$). WES yield was relatively high (42.9–60%) among families with involvement of brain, renal or musculoskeletal US findings. Taken together, our results in a real-world setting of genetic counseling demonstrates that fetal WES is especially indicated in families with positive family history, as well as in fetuses with specific types of congenital malformation.

Keywords: whole-exome sequencing (WES), prenatal diagnosis, ultrasound abnormalities, clinical genetics, congenital anomalies

INTRODUCTION

Recent years have seen the growing implementation of next-generation sequencing (NGS) techniques into widespread clinical use, revolutionizing the diagnostic odyssey for many families with monogenic disorders (Yang et al., 2013; Stranneheim and Wedell, 2016). While whole-exome sequencing (WES) and whole-genome sequencing (WGS) are more commonly utilized as a tool

for molecular diagnosis of affected pediatric and adult patients (Lee et al., 2014; Taylor et al., 2015; Sawyer et al., 2016), data regarding their utility in routine use at the prenatal setting is still relatively limited (Jelin and Vora, 2018).

Initially, several preliminary reports described the successful use of fetal WES/WGS in single families (Talkowski et al., 2012; Filges et al., 2014). Since then, several groups have reported their experience with exome sequencing in the context of fetuses detected by ultrasound (US) to have structural anomalies (for example, Carss et al., 2014; Drury et al., 2015; Fu et al., 2018; Normand et al., 2018). The widely variable diagnostic rates in literature (ranging from around 6.2% to as high as 80%) may be attributable to differences in inclusion criteria, single (fetus only) versus trio sequencing and sample size (Best et al., 2018). Diagnostic yield was higher in cases of, for instance, parental consanguinity, previous normal karyotype and chromosomal microarray analysis (CMA), fetuses with multiple congenital anomalies or fetal features suggestive of a specific syndrome, related to variants in multiple known causative genes (Vora et al., 2017; Best et al., 2018; Daum et al., 2019). Recently, two large scales studies of fetal WES were published. In the Prenatal Assessment of Genomes and Exomes (PAGE) study, which included exome sequencing in 610 fetuses (and 1202 parental samples) with structural anomalies detected by US, a diagnostic genetic variant was found in 8.5% of the fetuses, and additional 3.9% of fetuses had variant of unknown significance (VUS) with potential clinical usefulness (Lord et al., 2019). In another WES study of fetuses with structural anomalies, diagnostic rate of 10.3% among 234 fetus-parents trios was reported (Petrovski et al., 2019).

We describe herein our experience with clinical fetal WES in 45 consecutive families from either terminated or ongoing pregnancies, and discuss the added value and challenges associated with this diagnostic strategy. The study represents a real-world setting in genetic counseling service, in which mixture of various indications for WES is encountered, and WES is performed for clinical purposes and paid by families.

MATERIALS AND METHODS

Sample Recruitment

In this consecutive series study, we included all families who were consulted at the Danek Gertner Institute of Human Genetics at Sheba Medical Center (SMC) between January 2015 and September 2018, and for whom at least one fetal WES was performed, from a terminated or an ongoing pregnancy. A total of 45 families were retrospectively included in the current study.

All families were of Jewish ancestry, 44 were non-consanguineous and in one family the parents were first degree cousins. Prior to pregnancy or at its first stages, all families were instructed to undertake prenatal genetic carrier screening for prevalent diseases, according to their ethnic origin. All couples received genetic counseling at the SMC genetic institute and a detailed clinical evaluation, including at least one prenatal US examination, and family medical history were obtained. Fetal magnetic resonance imaging (MRI) was carried out in several pregnancies, in addition to US. We included in this study both fetuses with or without abnormal US findings.

WES was performed as a clinical service, paid for directly by the families (for two families, internal funding of the genetic institute was used), and final WES report was issued by the performing lab. Both parents provided written informed consent for the clinical WES, after they had received an explanation regarding the benefits and limitations of the test.

Fetal DNA was extracted following amniocentesis, chorionic villus sampling (CVS) or from fetal material taken during termination of pregnancy (TOP) procedure, and was stored for genetic testing. All TOPs were approved by the institutional committee. Parental DNA was extracted from peripheral blood samples. In addition, prior to WES, CMA was completed in 42 of the families (in at least single fetus) at the SMC cytogenetic laboratory, as a routine clinical service. In all these cases, no cytogenetic findings that explained the phenotype were detected. In the three families with missing CMA, molecular diagnosis was found by WES.

Whole-Exome Sequencing

Thirty three families performed trio (fetus-mother-father) WES, 7 families quatro (two fetuses-mother-father) WES and in 5 families a singleton WES was sent, depending on availability of fetal DNA from previously terminated pregnancies, as well as financial considerations.

The majority of WES (32 out of 45) were performed at Centogene laboratories[1]. Briefly, commercial capture kits for exome sequencing were used, and sequencing was performed on Ilumina platform, to obtain an average coverage depth of approximately 100X. Bioinformatics analysis included alignments of reads to reference genome, filtering out low quality reads and artifacts, and annotation of variants as described previously (Trujillano et al., 2017). Disease causing variants within the Human Gene Mutation Database (HGMD)[2], ClinVar[3] or in CentoMD[4] as well as variants with a minor allele frequency (MAF) of less than 1% in the Exome Aggregation Consortium (ExAC)[5] were considered, focusing on exons and flanking intronic bases. The family history and clinical description provided were used to evaluate the identified variants.

Eight of WES were performed at the bioinformatics unit of SMC Cancer Research Center, using similar WES and bioinformatics methodology, and five in CeGaT laboratories[6]. In five cases, the WES raw data was further transferred to an independent bioinformatitian for re-analysis, paid for directly by the family.

Exome data were interpreted according to the American College of Medical Genetics and Genomics (ACMG) guidelines (Richards et al., 2015), and variants were classified as either

[1]https://www.centogene.com/

[2]http://www.hgmd.cf.ac.uk/

[3]https://www.ncbi.nlm.nih.gov/clinvar/

[4]https://www.centomd.com/

[5]http://exac.broadinstitute.org/

[6]https://www.cegat.de/

pathogenic, likely pathogenic, uncertain significance, likely benign or benign.

Molecular Diagnosis

The final molecular diagnosis to each family was given by the SMC genetic institute team, based on the WES report variant classification and consistency between suggested genes and phenotype. A family was considered as molecularly diagnosed when pathogenic or likely pathogenic variants were found in genes that were pertinent to the clinical phenotype and in line with the suspected inheritance pattern. When variants were classified as VUS in the original WES report, re-evaluation was performed based on clinical considerations, bioinformatics and segregation analysis in additional family members. Team decision was made accordingly.

Data Analysis

Families were divided into three groups:

(1) Abnormal prenatal US findings and positive relevant family history: defined as families with two or more fetuses with abnormal prenatal US findings (in current and previous pregnancies), or families with a single fetus with US abnormalities and an affected parent, mother or father (with a resembling phenotype, or potentially manifesting carrier mother). This relates to options of autosomal recessive, autosomal dominant and X-linked models of inheritance, and even of gonadal mosaicism.
(2) Abnormal prenatal US findings without relevant family history: referring to families with a single fetus affected by abnormal US findings, but without any previous fetuses or family members (including adult siblings) with abnormal prenatal US or a medical history relevant to the findings.
(3) Fetuses without any US abnormalities (with or without first degree relative affected by a severe genetic disorder that was excluded in the fetus prior to WES, by Sanger sequencing).

For the diagnostic rate calculations, each family was counted as a single case. To evaluate the effect of fetal US abnormalities type on WES referral and diagnostic yield, we classified the families into five phenotypic groups (based on findings in at least one fetus): brain, renal, cardiac or musculoskeletal system abnormalities and increased nuchal translucency (NT)/edema/hydrops signs. Part of the families matched to more than one group, while others to none.

For statistical analysis, we used *T*-test, Pearson chi-square or Fisher's exact test (two sided), as appropriate. *P*-value < 0.05 was considered statistically significant.

RESULTS

Characteristics of 45 fetal WES families included in the presented study and type of WES (single/trio/qautro) are presented in **Table 1**. In 41 of the 45 families, WES was performed due to a wide range of prenatal US abnormalities. Average maternal age at the pregnancy of the index fetus was 33.5 ± 4.2 years (range, 23–43 years). The pregnant women were at an average gestational age of 23 ± 7.5 weeks (range, 12.5–36 weeks) at first genetic counseling in our clinic (for the index fetus). For 23 families, WES was sent after TOP (51.1%), and in 22 families during an ongoing pregnancy (48.9%). Turnaround time from WES submission to results was 15 to 20 working days for ongoing pregnancies, and up to 2 months following TOP.

Overall, 13 of the 45 families met criteria for a molecular diagnosis, reaching diagnostic yield of 28.9% (**Table 1**). Among families with US abnormalities, the yield was 13/41 (31.7%). We detected three *de novo* heterozygous causative variants, one autosomal dominant causative variant inherited from the mother, eight autosomal recessive cases (three compound heterozygous and five with homozygous variants) and one X-linked causative variant. Osteogenesis imperfecta due to *de novo* causative variant in the *COL1A2* gene was found in two families, while all other conditions occurred once in a single family only. All diagnosed families are described in **Table 2**, which includes two cases that were previously reported by our group: diaphanospondylodysostosis (*BMPER* gene) (Greenbaum et al., 2019) and *LMOD3*-associated Nemaline Myopathy (Berkenstadt et al., 2018). Description of US findings in families without molecular diagnosis is presented in **Supplementary Table 1**.

In none of the four families without prenatal abnormal US findings was a molecular diagnosis reached. Among the 18 families with prenatal US abnormalities and positive relevant family history, a molecular diagnosis was found in 10 (55.6%). This was significantly higher compared to the 23 families with abnormal prenatal US findings and lack of relevant family history, in which the molecular cause was detected in only 3 (13%) ($\chi^2(1) = 8.43$, $p = 0.004$). There was no significant difference in maternal age between the two groups.

In a sub-analysis, the diagnostic rate in families with two or more fetuses with abnormal prenatal US findings was 8 out of 15 (53.3%), significantly higher than in families with prenatal US abnormalities in a mere single fetus, regardless of a relevant family history (5 out of 26, 19.2%; $p = 0.038$, Fisher's exact test).

Looking at specific organ/body system involvement, higher diagnostic rates (more than the overall rate) were observed among families with fetal US abnormalities in the brain (60%), renal (42.9%) and musculoskeletal system (55.6%), compared to cardiac US abnormalities (16.7%) or increased NT/edema/hydrops signs (27.3%) (**Table 3**).

Among ongoing pregnancies with US abnormalities, molecular diagnosis was found in 2 of 18 families (11.1%), both with positive relevant family history (the pregnancies were carried on). Interestingly, among 41 families with US abnormalities, significantly more families without relevant history performed WES during ongoing pregnancy (14/23, 60.9%), compared to families with positive history (4/18, 22.2%) ($\chi^2(1) = 6.12$, $p = 0.013$).

For all families with molecular diagnosis, options for future family planning were discussed, including prenatal diagnosis by amniocentesis or pregestational diagnosis (PGD) before next pregnancy (in autosomal dominant, recessive or X-linked inheritance) and amniocentesis to rule out recurrence due to gonadal mosaicisim (for *de novo* mutations).

TABLE 1 | Characteristics of 45 fetal WES families included in the presented study, type of WES and rate of molecular diagnosis.

Families included	Total number	Maternal age at pregnancy (years mean, SD)	Fetal WES performed during ongoing pregnancy (N,%)	Single/Trio/Quatro WES	Families with molecular diagnosis (N,%)
Overall	45	33.5 (4.2)	22 (48.9%)	5/33/7	13 (28.9%)
(1) Abnormal prenatal US findings and positive relevant family history	18	32.4 (4)	4 (22.2%)	2/9/7	10 (55.6%)
- Two or more fetuses with abnormal prenatal US findings (*)	15	32.1 (4.1)	2 (13.3%)	1/7/7	8 (53.3%)
- Single fetus with US abnormalities and an affected parent	3	34.3 (3.8)	2 (66.7%)	1/2/0	2 (66.7%)
(2) Single fetus with abnormal prenatal US findings and lack of relevant family history	23	34.3 (4.5)	14 (60.9%)	1/22/0	3 (13%)
(3) Fetuses without any US abnormalities	4	33.8 (2.4)	4 (100%)	2/2/0	0

In families with two or more affected fetuses, maternal age at pregnancy relates to pregnancy of the later fetus included in the WES.
**In one family, the father and two fetuses were affected.*

Representative Clinical Cases

Illustrative cases of two families with molecular diagnosis are presented below (numbered according to **Table 2**), emphasizing the challenges in the genetic counseling and the necessity of accurate US phenotyping. In the first family, with two affected fetuses (suspected recessive inheritance model), the list of potential disease causing genes in the differential diagnosis was long, and WES shortened the time of investigation. In the second family, without relevant family history, no specific diagnosis was suspected by the clinicians, and WES was probably the most cost-effective method for genetic work-up.

Family 2

A healthy couple of Jewish Ethiopian origin was referred for genetic counseling at 33 weeks and 2 days of gestation in their second pregnancy. A previous pregnancy was terminated at week 25, due to abnormal US findings, including encephalocele, large multicystic kidneys, oligohydramnios, suspected polydactyly and lack of urinary bladder and stomach demonstration. In the second pregnancy, NT was normal. US at early stage of pregnancy showed brain and renal malformations, but the couple did not continue follow up. At 33 week of gestation, US demonstrated posterior fossa abnormality (suspected dandy-walker malformation), a short and malformed corpus callosum, intrauterine growth retardation (IUGR), single umbilical artery (SUA), small dysgenetic kidneys, hypertelorism, and oligohydramnios. Urinary bladder was visualized. In follow up fetal brain MRI, a large cystic finding in posterior fossa was noticed, in addition to ventriculomegaly, heterotopic foci at the ventricular wall and short corpus callosum. From this deceased fetus, DNA was extracted, CMA was normal, and the couple sent a trio WES few months later. The combination of brain malformations, polydactcyly and kidneys abnormalities suggested diagnosis of Meckle–Gruber syndrome. However, since many genes are related to this disorder, and to exclude other potential disorders which were on the differential diagnosis, WES was regarded as optimal diagnostic approach. A homozygous NM_024809.4: c.1506-2A>G variant in the *TCTN2* gene (intron 13) was found in the fetus. This variant is predicted to disrupt a highly conserved acceptor splice site, and was previously described as disease causing for Meckel–Gruber type 8 syndrome in consanguineous Arab family (Shaheen et al., 2011). Other mutations in *TCTN2* were found to cause Joubert syndrome (Huppke et al., 2015). Both parents were heterozygous for this variant. The couple was advised to consider PGD in following pregnancies.

Family 13

A Healthy Jewish couple, both 31 years old (father from Iraq origin and mother of Ashkenazi origin) was consulted at 31 weeks (w)+5 days (d) of gestation of their first pregnancy. NT was 2.9 mm. On 16w+3d of gestation, US demonstrated elevated nuchal fold (5.7 mm), bilateral mild hypoplastic 5th finger and bilateral pyelectasis. Fetal echocardiogram was normal. Amniocentesis performed and CMA was normal. At 22w+3d, relative short and thick corpus callosum was noticed, as well as nuchal fold of 5.1–5.5 mm and mild pyelectasis. Further multiple US were carried by several experts, indicating short corpus callosum, hydronephrosis (10–13 mm), bulbous nose and mild polyhydramnios. Fetal brain MRI (at 30 weeks of gestation) demonstrated dysmorphic and short corpus callosum (<3rd percentile). The couple decided to terminate the pregnancy. In post mortem observation, facial dysmorphism, including coarseness and a bulbous nose, was noticed. Trio WES revealed a *de novo* heterozygous variant in *TCF4* gene (NM_001243226.2: c.2032C>T, p.Arg678*) in the fetus, consistent with the diagnosis

TABLE 2 | Resolved 13 families- summary of main US findings, WES results and final molecular diagnosis.

Family number	Main US findings (according to fetuses)	WES type	Gene	Causative variants	Inheritance and zygosity	Diagnosis (relevant phenotype MIM number)
1	1st: Shortening of long bones (femur, humerus, tibia), IUFD 2nd: Narrow thorax, bowed femur, shortening of long bones (mostly fibula)	Quatro	*EVC2*	NM_147127.4: c.572A>T, p.Asn191Ile; NM_147127.4: c.3265C>T, p.Gln1089*	AR (compound het)	Ellis-van Creveld syndrome (MIM: 225500)
2	1st: Encephalocele, large multicystic kidneys, oligohydrmanios, suspected polydactyly, lack of urinary bladder and stomach demonstration 2nd: Posterior fossa abnormality (suspected dandy-walker malformation), short and malformed corpus callosum, IUGR, SUA, small dysgenetic kidneys, urinary bladder was not visualized, oligohydramnios, hypertelorism	Trio	*TCTN2*	NM_024809.4: c.1506-2A>G	AR (hom)	Meckel syndrome type 8 (MIM: 613885)
3	1st: Fetal akinesia, mild polyhydramnios, small stomach, suspected right clubfoot, extended lower limbs, clenched hands, neck hyper-extension 2nd: Arthrogryposis, hypotonic features, abnormal posture	Trio	*LMOD3*	NM_198271.4: c.723_733del, p.Asp242Glufs*4; NM_198271.4: c.360dupA, p.Glu121Argfs*5	AR (compound het)	Nemaline Myopathy 10 (MIM: 616165)
4	1st: Abnormal spine and chest, unusual skull shape, echogenic cystic and horseshoe like kidneys 2nd: Increased NT (8 mm), generalized edema, spine distortion, bilateral clubfoot, absent nasal bone 3rd: Reduced/lack ossification in the skull, ribs and vertebrae, protruding abdomen, short trunk	Single	*BMPER*	NM_133468.5: c.410T>A, p.Val137Asp	AR (hom)	Diaphanospondylodysostosis (MIM: 608022)
5	1st: Distal arthrogryposis (hands), probably unilateral clubfoot, IUFD 2nd: Bilateral clubfoot	Quatro	*FKBP14*	NM_017946.3: c.568_570del, p.Lys190del	AR (hom)	Ehlers-Danlos syndrome, kyphoscoliotic type, 2 (MIM: 614557)
6	1st: Posterior urethral valve, cystic finding in kidney, suspected omphalocele 2nd: Increased NT (10 mm), cystic lesion near umbilical cord insertion site 3rd: Septated cystic hygroma, partial vermian agenesis, ARSA, omphalocele, echogenic and multicystic kidneys 4th: Increased NT (4.5 mm), facial dysmorphism (hypoplastic nasal bridge and micrognathia), echogenic kidneys, omphalocele, postaxial polydactyly, clubfoot, complex heart malformation (VSD, double outlet right ventricle, tricuspid valve regurgitation)	Quatro	*PIGN*	NM_176787.5: c.163C>T, p.Arg55*; NM_176787.5: c.2283G>C, p.Lys761Asn	AR (compound het)	Multiple congenital anomalies-hypotonia-seizures syndrome 1 (MIM: 614080)

(Continued)

TABLE 2 | Continued

Family number	Main US findings (according to fetuses)	WES type	Gene	Causative variants	Inheritance and zygosity	Diagnosis (relevant phenotype MIM number)
7	1st: large polycystic kidneys, oligohydramnios, moderate bilateral ventriculomegaly 2nd: Polycystic kidneys, hydrocephalus, mega cysterna magna, macrocephaly 3rd: Enlarged echogenic kidneys, severe oligohydramnios, hydrocephalus, mega cysterna magna, thin corpus callosum	Quatro	*CPT2*	NM_001330589.1: c.1239_1240del, p.Lys414Thrfs*7	AR (hom)	CPT II deficiency, lethal neonatal (MIM: 608836)
8	1st: Occipital encephalocele, ventriculomegaly, mild to moderate hydronephrosis 2nd: Occipital encephalocele, ventriculomegaly, microphthalmia, cataract	Trio	*B3GALNT2*	NM_001277155.2: c.236-1G>C	AR (hom)	Muscular dystrophy-dystroglycanopathy (congenital with brain and eye anomalies, type A, 11) (MIM: 615181)
9	Short corpus callosum, suspected unilateral cataract and coloboma, IUGR (*)	Single	*MED12*	NM_005120.3; c.6388C>T, p.Gln2130*	XL hemizygous (maternally inherited)	Opitz-Kaveggia syndrome (MIM: 305450); Ohdo syndrome, X-linked (MIM: 300895)
10	Preaxial polydactyly of foot, syndactyly of hands (**)	Trio	*GLI3*	NM_000168.6: c.1445G>A, p.Cys482Tyr	AD het (maternally inherited)	Greig cephalosyndactyly (MIM: 175700); Polydactyly, preaxial, type IV (MIM: 174700)
11	Shortening and bowing of long bones, poor bone mineralization, reduced skull ossification, small/narrow thorax	Trio	*COL1A2*	NM_000089.3: c.1829G>T, p.Gly610Val	*de novo* het	Osteogenesis imperfecta type 2-3 (MIM: 166210, 259420)
12	Poor ossification of skull, tibial bowing, fractures of femur, shortening of long bones	Trio	*COL1A2*	NM_000089.3: c.2260G>T, p.Gly754Cys	*de novo* het	Osteogenesis imperfecta type 2-3 (MIM: 166210, 259420)
13	Increased NT (5.7 mm), bilateral mild hypoplastic 5th finger, bilateral pyelectasis/hydronephrosis, short and thick corpus callosum, bulbous nose, mild polyhydramnios	Trio	*TCF4*	NM_001243226.2: c.2032C>T, p.Arg678*	*de novo* het	Pitt-Hopkins syndrome (MIM: 610954)

All fetuses had abnormal US findings. Cases 1–8 refer to families with two or more affected fetuses, 9–10 to families with single fetus and relevant family medical history (affected parent) and 11–13 to families with single fetus and lack of relevant family history. Cases 3 and 4 were previously reported by Berkenstadt et al. (2018) and Greenbaum et al. (2019), respectively.

**Mother – Prominent facial dysmorphism (synophrys, small mouth and ears, mid-facial hypoplasia, prominent philtrum, dental crowding), head circumference in 10th percentile, hirsutism, menstural abnormalities and Intact cognition. She is probably a manifesting carrier. The fetus was male.*

***Mother – syndactyly (hands), hypotonia at infancy.*

Abbreviations: AD, autosomal dominant; AR, autosomal recessive; ARSA, aberrant right subclavian artery; het, heterozygous; hom, homozygous. IUFD, intrauterine fetal death; IUGR, intrauterine growth restriction; MIM, Mendelian inheritance in man; NT, nuchal translucency, SUA, single umbilical artery; VSD, ventricular septal defect; XL, X-linked.

TABLE 3 | Referral indications and WES yield, according to US findings in specific organ or body systems (among 41 families with US abnormalities).

Organ/body system	Number of families referred due to indication (N,%)	Families with molecular diagnosis (N,%)
Brain	10 (24.4%)	6/10 (60%)
Renal	14 (34.1%)	6/14(42.9%)
Musculoskeletal (*)	18 (43.9%)	10/18 (55.6%)
Cardiac malformation	6 (14.6%)	1/6 (16.7%)
Increased NT(**)/edema/hydrops signs	11 (26.8%)	3/11 (27.3%)

Part of the families were classified to more than single organ/body system group, while others to none.
**Distal limbs anomalies (e.g., polydactyly, syndactyly, clubfoot etc.) were included in this group, but not orofacial malformations.*
***Increased NT was defined as above 3 mm.*

of Pitt-Hopkins syndrome. The fetus facial dysmorphism and the short corpus callosum supported the diagnosis. This variant creates a premature stop codon, and was previously reported in this disorder (Maduro et al., 2016). In a following pregnancy, amniocentesis was performed to exclude mutation in the fetus, due to option of germline mosaicism.

DISCUSSION

In this retrospective study, representing a clinical experience and practice within a tertiary referral hospital in Israel, we investigated the yield of fetal WES, from terminated and ongoing pregnancies.

The overall yield of molecular diagnosis was 28.9% (13 out of 45 families), and in families with US abnormalities, the yield was 31.7%. We noticed a significantly higher diagnostic rate in families with abnormal prenatal US findings and positive relevant family history (mainly an affected fetus from previous pregnancies), compared to families without a relevant history. This shows the importance of fetal WES indication for pregnancies with US abnormalities and positive family history, suggesting an underlining genetic disease.

Our diagnostic rate is within the range of 21–32%, seen in several studies of diagnostic yield for exome sequencing performed on fetal samples (Drury et al., 2015; Fu et al., 2018; Normand et al., 2018; Daum et al., 2019), and close to a diagnostic rate of 36.7% from exome sequencing of 278 infants in an intensive care unit (Meng et al., 2017). However, it is substantially higher than 8.5% and 10.3% found in two recent large scale fetal WES studies (Lord et al., 2019; Petrovski et al., 2019). This gap may be explained by enrichment of cases with positive family history in our cohort, and indeed our diagnostic rate in families without such history was only 13%. Interestingly, a study of 146 fetal exomes (Normand et al., 2018) has not found difference in diagnostic rate between sporadic cases and cases with significant family history. Therefore, the effect of medical family history on fetal WES yield should be further studied.

In addition, WES yield was in the range of 42.9–60% among families with involvement of brain, renal or musculoskeletal US abnormalities (higher than the overall diagnostic rate), and lower among families with cardiac malformations or increased NT/edema/hydrops signs (16.7 and 27.3%, respectively). However, since our sample size is relatively small, any interpretation of this sub-analysis should be cautious.

The resolved cases resulted in identification of causative variants in 12 genes. Some of the diagnoses are relatively common in the setting of prenatal diagnosis, such as osteogenesis imperfecta, while others are rare and include novel variants neither previously reported nor found in the databases. The latter are challenging to interpretate, due to the uncertain impact of the variants. As expected in the clinical setting, due to time and budget restrictions, no functional analysis at the molecular level was available for our cases. Thus, segregation of the variant in other family members was crucial to support diagnosis. In light of these limitations, we cannot definitely exclude that some of the molecular diagnoses are false positive. Two illustrative cases, families with molecular diagnosis of Meckel–Gruber type 8 and Pitt-Hopkins syndromes, demonstrate the clinical and genetic work-up during pregnancy and after termination, the utility of fetal WES, and its contribution for future family planning.

Well-described in the postnatal setting, prenatal WES also harbors special challenges, including numerous ethical issues which should be contemplated and addressed (Hillman et al., 2015; Jelin and Vora, 2018). The potential identification of secondary findings (also designated as incidental) and/or VUS, may have profound effect when found in ongoing pregnancies, either on parental decision making with regard to termination of pregnancy, or on the parental psychological burden following the birth of an affected child (Hillman et al., 2015; van den Veyver and Eng, 2015). One suggested strategy to circumvent some of these concerns in the prenatal setting is targeted exome testing, in which a premeditated and limited list of genes is analyzed (Pangalos et al., 2016; Chandler et al., 2018). On the other hand, according to the ACMG guidelines, no incidental findings in its selected list of actionable genes should be reported in prenatal samples.

Other concerns include issues related to turnaround time, lack of full coverage of potentially relevant genes, and inaccurate and limited phenotyping capabilities for fetuses, resulting in insufficient data for the WES interpretation (Aarabi et al., 2018; Best et al., 2018). Postnatal clinical evaluations or autopsies may be of great value, improving the diagnostic accuracy and yield, with new data not observed in US examination (Aarabi et al., 2018).

Our study is based on a cohort of consecutive (non-selected) series of families for which WES was performed in certified clinical laboratories, on a clinical base (not for research purpose as the main goal), and fully paid for, usually by families. In this regard, the study represents a real-world situation in prenatal

counseling practicality, where the diagnosis is pursued due to a clinical need, and not within a research setup. Of note, most families with US abnormalities and lack of relevant family history asked to perform WES during ongoing pregnancy (14/23, 60.9%), while most families with positive history sent it after TOP (14/18, 77.8%). Four families asked to perform fetal WES during ongoing pregnancy even when no US abnormalities were found, mainly for detection of *de novo* pathogenic variants or recessive disorders. Three of the four families had offsprings affected by a severe single-gene condition. In these cases, WES was sent even when the specific pathogenic variant in the family was excluded in the fetus prior to WES submission (by Sanger sequencing), since families asked to search for other potential genetic disorders in the fetus.

The current study population is relatively homogenous in terms of demographic characteristics, composed of 45 families from Jewish ancestry, without consanguinity (except one family), all living in Israel and followed by a single prenatal genetic clinic. For most families in our cohort, genetic evaluation in addition to WES included CMA as well as parental screening for common genetic diseases in the Jewish population. In this aspect, the reported cohort might not be representative of other patient populations. For example, it is plausible that diagnostic yield may be even higher in WES pursued for fetuses of consanguineous couples.

The majority of families with fetal abnormal US findings seen in our clinic (several hundred each year), decided not to perform fetal WES, due to multiple considerations, including financial issues. Therefore, our study is biased toward families with high motivation to reach molecular diagnosis and available financial resources. An additional limitation of the present study is that WES was sent to three different clinical services. Therefore, sequencing was performed using different protocols and kits, and diverse bioinformatics pipelines were used for data analysis, potentially affecting the diagnostic yield due to inter-laboratory differences. Uniform sequencing and analysis in one laboratory, applied to all families, may have been an advantage. In addition, revision of WES data was not done for most unresolved cases. Together with our relative small sample size, the results should be considered as preliminary, and warrant further study.

To conclude, our study demonstrates that fetal WES is especially indicated when positive relevant family history is present, or in fetuses with specific types of congenital malformation. Following careful patient selection, clinical fetal WES is an important tool for evaluating fetuses with abnormal US findings, with important implications for both ongoing and future pregnancies.

ETHICS STATEMENT

The study was carried out in accordance with the recommendations of the Institutional Review Board (IRB) committee of SMC, and was approved by it.

AUTHOR CONTRIBUTIONS

LG and MB designed the study, provided genetic counseling to families, researched the data, and wrote the manuscript. BP-S and EP contributed to study design and interpretations of data. SE-B, MD-K, AB-Z, NG, HP, MR, and AR-R provided genetic counseling to families, collected the data, and revised the manuscript. AR assisted in statistical analysis. HR-W, AB-A, PB, and OB participated in bioinformatics analysis and revised the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2019.00425/full#supplementary-material

REFERENCES

Aarabi, M., Sniezek, O., Jiang, H., Saller, D. N., Bellissimo, D., Yatsenko, S. A., et al. (2018). Importance of complete phenotyping in prenatal whole exome sequencing. *Hum. Genet.* 137, 175–181. doi: 10.1007/s00439-017-1860-1

Berkenstadt, M., Pode-Shakked, B., Barel, O., Barash, H., Achiron, R., Gilboa, Y., et al. (2018). LMOD3-associated nemaline myopathy: prenatal ultrasonographic, pathologic, and molecular findings. *J. Ultrasound Med.* 37, 1827–1833. doi: 10.1002/jum.14520

Best, S., Wou, K., Vora, N., Van Der Veyver, I. B., Wapner, R., and Chitty, L. S. (2018). Promises, pitfalls and practicalities of prenatal whole exome sequencing. *Prenat. Diagn.* 38, 10–19. doi: 10.1002/pd.5102

Carss, K. J., Hillman, S. C., Parthiban, V., Mcmullan, D. J., Maher, E. R., Kilby, M. D., et al. (2014). Exome sequencing improves genetic diagnosis of structural fetal abnormalities revealed by ultrasound. *Hum. Mol. Genet.* 23, 3269–3277. doi: 10.1093/hmg/ddu038

Chandler, N., Best, S., Hayward, J., Faravelli, F., Mansour, S., Kivuva, E., et al. (2018). Rapid prenatal diagnosis using targeted exome sequencing: a cohort study to assess feasibility and potential impact on prenatal counseling and pregnancy management. *Genet. Med.* 20, 1430–1437. doi: 10.1038/gim.2018.30

Daum, H., Meiner, V., Elpeleg, O., and Harel, T. (2019). Fetal exome sequencing: yield and limitations in a single tertiary center. *Ultrasound Obstet Gynecol.* 53, 80–86. doi: 10.1002/uog.19168

Drury, S., Williams, H., Trump, N., Boustred, C., Lench, N., Scott, R. H., et al. (2015). Exome sequencing for prenatal diagnosis of fetuses with sonographic abnormalities. *Prenat. Diagn.* 35, 1010–1017. doi: 10.1002/pd.4675

Filges, I., Nosova, E., Bruder, E., Tercanli, S., Townsend, K., Gibson, W. T., et al. (2014). Exome sequencing identifies mutations in KIF14 as a novel cause of an autosomal recessive lethal fetal ciliopathy phenotype. *Clin. Genet.* 86, 220–228. doi: 10.1111/cge.12301

Fu, F., Li, R., Li, Y., Nie, Z. Q., Lei, T., Wang, D., et al. (2018). Whole exome sequencing as a diagnostic adjunct to clinical testing in fetuses with structural abnormalities. *Ultrasound Obstet Gynecol.* 51, 493–502. doi: 10.1002/uog.18915

Greenbaum, L., Gilboa, Y., Raas-Rothschild, A., Barel, O., Kol, N., Wolf, H. R., et al. (2019). Diaphanospondylodysostosis: refining the prenatal diagnosis of a rare skeletal disorder. *Eur. J. Med. Genet.* 62, 167–171. doi: 10.1016/j.ejmg.2018.07.004

Hillman, S. C., Willams, D., Carss, K. J., Mcmullan, D. J., Hurles, M. E., and Kilby, M. D. (2015). Prenatal exome sequencing for fetuses with structural abnormalities: the next step. *Ultrasound Obstet. Gynecol.* 45, 4–9. doi: 10.1002/uog.14653

Huppke, P., Wegener, E., Bohrer-Rabel, H., Bolz, H. J., Zoll, B., Gartner, J., et al. (2015). Tectonic gene mutations in patients with joubert syndrome. *Eur. J. Hum. Genet.* 23, 616–620. doi: 10.1038/ejhg.2014.160

Jelin, A. C., and Vora, N. (2018). Whole exome sequencing: applications in prenatal genetics. *Obstet. Gynecol. Clin. North Am.* 45, 69–81. doi: 10.1016/j.ogc.2017.10.003

Lee, H., Deignan, J. L., Dorrani, N., Strom, S. P., Kantarci, S., Quintero-Rivera, F., et al. (2014). Clinical exome sequencing for genetic identification of rare Mendelian disorders. *JAMA* 312, 1880–1887. doi: 10.1001/jama.2014.14604

Lord, J., Mcmullan, D. J., Eberhardt, R. Y., Rinck, G., Hamilton, S. J., Quinlan-Jones, E., et al. (2019). Prenatal exome sequencing analysis in fetal structural anomalies detected by ultrasonography (PAGE): a cohort study. *Lancet* 393, 747–757. doi: 10.1016/S0140-6736(18)31940-8

Maduro, V., Pusey, B. N., Cherukuri, P. F., Atkins, P., Du Souich, C., Rupps, R., et al. (2016). Complex translocation disrupting TCF4 and altering TCF4 isoform expression segregates as mild autosomal dominant intellectual disability. *Orphanet. J. Rare Dis.* 11:62. doi: 10.1186/s13023-016-0439-6

Meng, L., Pammi, M., Saronwala, A., Magoulas, P., Ghazi, A. R., Vetrini, F., et al. (2017). Use of exome sequencing for infants in intensive care units: ascertainment of severe single-gene disorders and effect on medical management. *JAMA Pediatr.* 171:e173438. doi: 10.1001/jamapediatrics.2017.3438

Normand, E. A., Braxton, A., Nassef, S., Ward, P. A., Vetrini, F., He, W., et al. (2018). Clinical exome sequencing for fetuses with ultrasound abnormalities and a suspected mendelian disorder. *Genome Med.* 10:74. doi: 10.1186/s13073-018-0582-x

Pangalos, C., Hagnefelt, B., Lilakos, K., and Konialis, C. (2016). First applications of a targeted exome sequencing approach in fetuses with ultrasound abnormalities reveals an important fraction of cases with associated gene defects. *PeerJ* 4:e1955. doi: 10.7717/peerj.1955

Petrovski, S., Aggarwal, V., Giordano, J. L., Stosic, M., Wou, K., Bier, L., et al. (2019). Whole-exome sequencing in the evaluation of fetal structural anomalies: a prospective cohort study. *Lancet* 393, 758–767. doi: 10.1016/S0140-6736(18)32042-7

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., et al. (2015). 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.* 17, 405–424. doi: 10.1038/gim.2015.30

Sawyer, S. L., Hartley, T., Dyment, D. A., Beaulieu, C. L., Schwartzentruber, J., Smith, A., et al. (2016). Utility of whole-exome sequencing for those near the end of the diagnostic odyssey: time to address gaps in care. *Clin. Genet.* 89, 275–284. doi: 10.1111/cge.12654

Shaheen, R., Faqeih, E., Seidahmed, M. Z., Sunker, A., Alali, F. E., Alqahtani, K., et al. (2011). A TCTN2 mutation defines a novel meckel gruber syndrome locus. *Hum. Mutat.* 32, 573–578. doi: 10.1002/humu.21507

Stranneheim, H., and Wedell, A. (2016). Exome and genome sequencing: a revolution for the discovery and diagnosis of monogenic disorders. *J. Intern. Med.* 279, 3–15. doi: 10.1111/joim.12399

Talkowski, M. E., Ordulu, Z., Pillalamarri, V., Benson, C. B., Blumenthal, I., Connolly, S., et al. (2012). Clinical diagnosis by whole-genome sequencing of a prenatal sample. *N. Engl. J. Med.* 367, 2226–2232. doi: 10.1056/NEJMoa1208594

Taylor, J. C., Martin, H. C., Lise, S., Broxholme, J., Cazier, J. B., Rimmer, A., et al. (2015). Factors influencing success of clinical genome sequencing across a broad spectrum of disorders. *Nat. Genet.* 47, 717–726. doi: 10.1038/ng.3304

Trujillano, D., Bertoli-Avella, A. M., Kumar Kandaswamy, K., Weiss, M. E., Koster, J., Marais, A., et al. (2017). Clinical exome sequencing: results from 2819 samples reflecting 1000 families. *Eur. J. Hum. Genet.* 25, 176–182. doi: 10.1038/ejhg.2016.146

van den Veyver, I. B., and Eng, C. M. (2015). Genome-wide sequencing for prenatal detection of fetal single-gene disorders. *Cold Spring Harb. Perspect Med.* 5:a023077. doi: 10.1101/cshperspect.a023077

Vora, N. L., Powell, B., Brandt, A., Strande, N., Hardisty, E., Gilmore, K., et al. (2017). Prenatal exome sequencing in anomalous fetuses: new opportunities and challenges. *Genet. Med.* 19, 1207–1216. doi: 10.1038/gim.2017.33

Yang, Y., Muzny, D. M., Reid, J. G., Bainbridge, M. N., Willis, A., Ward, P. A., et al. (2013). Clinical whole-exome sequencing for the diagnosis of mendelian disorders. *N. Engl. J. Med.* 369, 1502–1511. doi: 10.1056/NEJMoa1306555

Conflict of Interest Statement: AB-A and PB are employees of Centogene AG.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Genetics

CASE REPORT
published: 02 July 2019
doi: 10.3389/fgene.2019.00630

Paternal Low-Level Mosaicism-Caused *SATB2*-Associated Syndrome

*Yeqing Qian[1,2], Jiao Liu[3], Yanmei Yang[1,2], Min Chen[1,2], Chunlei Jin[3], Penglong Chen[3], Yongliang Lei[3], Hangyi Pan[1,2] and Minyue Dong[1,2]**

[1] Women's Hospital, School of Medicine Zhejiang University, Hangzhou, China, [2] Key Laboratory of Reproductive Genetics (Zhejiang University), Ministry of Education, Hangzhou, China, [3] Prenatal Diagnosis Center, Lishui Maternity and Child Health Care Hospital, Lishui, China

Mutations of *SATB2* (OMIM#608148) gene at 2q33.1 have been associated with the autosomal dominant *SATB2*-associated syndrome (SAS), which is still short of comprehensive diagnosis technologies for small deletions and low-level mosaicism. In this Chinese Han family, single nucleotide polymorphism array identified a 4.9-kb deletion in the *SATB2* gene in two consecutive siblings exhibiting obvious developmental delay and dental abnormalities but failed to find so in their parents. Prenatal diagnosis revealed that their third child carried the same deletion in *SATB2* and the pregnancy was terminated. To determine the genetic causes behind the inheritance of *SATB2* deletion, gap-PCR was performed on peripheral blood-derived genomic DNA of the family and semen-derived DNA from the father. Gap-PCR that revealed the deletions in the two affected siblings were inherited from the father, while the less intense mutant band indicated the mosaicism of this mutation in the father. The deletion was 3,013 bp in size, spanning from chr2: 200,191,313-200,194,324 (hg19), and covering the entire exon 9 and part of intron 8 and 9 sequences. Droplet digital PCR demonstrated mosaicism percentage of 13.2% and 16.7% in peripheral blood-derived genomic DNA and semen-derived DNA of the father, respectively. Hereby, we describe a family of special AT-rich sequence-binding protein 2-associated syndrome caused by paternal low-level mosaicism and provide effective diagnostic technologies for intragenic deletions.

Keywords: *SATB2*-associated syndrome, chromosome microarray analysis, mosaicism, droplet digital PCR, gap-PCR

OPEN ACCESS

Edited by:
Peter C.K. Leung,
University of British Columbia, Canada

Reviewed by:
Tao Yang,
Shanghai Jiao-Tong University School of Medicine, China
Bufang Xu,
Shanghai Jiao Tong University, China

****Correspondence:***
Minyue Dong
dongmy@zju.edu.cn

Specialty section:
This article was submitted to Genetic Disorders, a section of the journal Frontiers in Genetics

Received: *27 March 2019*
Accepted: *17 June 2019*
Published: *02 July 2019*

Citation:
Qian Y, Liu J, Yang Y, Chen M, Jin C, Chen P, Lei Y, Pan H and Dong M (2019) Paternal Low-Level Mosaicism-Caused SATB2-Associated Syndrome. Front. Genet. 10:630. doi: 10.3389/fgene.2019.00630

INTRODUCTION

Special AT-rich sequence-binding protein 2 (SATB2) encodes a nuclear matrix DNA-binding protein that is involved in transcription regulation and chromatin remodeling. The *SATB2* gene spans 195.6 kb of genomic DNA (chr2:200,134,223-200,329,831, hg19) and is located at chromosome 2q33.1. SATB2 protein is composed of 733 amino acids, weighing approximately 82.5 kDa. The functional domains of SATB2 consist of two CUT domains (CUT1 and CUT2) that bind to the matrix attachment region and a homeodomain at the C-terminus (Fitzpatrick et al., 2003). All these domains are highly conserved across the vertebrate taxa (Sheehan-Rooney et al., 2010).

The *SATB2*-associated syndrome (SAS), which is a relatively newly described syndrome, is an autosomal dominant disorder also caused by alterations in the *SATB2* gene (Docker et al., 2014).

SAS is characterized by neurodevelopmental disabilities (intellectual disability and absent or severely impaired speech), behavioral abnormalities (hyperactivity, sleeping difficulties, autistic features, obsessive tendencies, and/or aggressiveness), craniofacial anomalies (cleft palate or high-arched palate, dental abnormalities), and skeletal anomalies (tibial bowing, osteomalacia, osteopenia, or osteoporosis) (Docker et al., 2014; Bengani et al., 2017; Zarate and Fish, 2017; Zarate et al., 2017; Zarate et al., 2018b).The major features of SAS can be summarized by the acronym S.A.T.B.2: S, severe speech anomalies; A, abnormalities of the palate; T, teeth anomalies; B, behavioral issues with or without bone or brain MRI anomalies; and age of onset before 2 years of age (Docker et al., 2014). Dr. Yuri Zarate, one of the most famous experts in the SAS area, gave recommended diagnostic evaluation and health surveillance for SAS at the website www.satb2gene.com.

In this study, we report the first occurrence of phenotypically normal father with low-level mosaicism of intragenic deletion in the *SATB2* gene, which was identified in his two children with SAS. This deletion was initially hinted by chromosomal microarray (CMA) in affected children but failed to diagnose in their father, while mosaicism in father was ascertained by gap-PCR, and the percentages of mosaicism in both the peripheral blood and semen were determined by droplet digital PCR (ddPCR). Our study provided a novel diagnostic method for those intragenic deletions in families having more than two consecutive births carrying similar abnormalities.

MATERIALS AND METHODS

Clinical Report

The 33-year-old pregnant woman was at 16 weeks of gestation. Her husband, 38 years old, has a normal phenotype as his wife does. No family history of mental or behavioral abnormalities or any other intellectual disability were reported. They came to our hospital for prenatal diagnosis because their two children had severe developmental delay. The husband's mother was diagnosed with esophageal cancer, and she died of multiple tumors at the age of 50.

The proband, male, 11 years old, weighed 28.6 kg (10th percentile), 136.6 cm tall (10th percentile), and had a low bone density (1st percentile). Main clinical manifestations consist of intellectual disability, no speech, high-arched palate, dental abnormalities (crowding and abnormal shape), small jaw, short tongue strap (surgically corrected), hypotonia, poor appetite, salivation, behavior abnormality (aggressiveness), and hydrocele. The head MRI examination showed brain demyelinating disease at the age of 3. He started walking at 4 years of age.

His sister, female, 8 years old, weighed 19.0 kg (3th percentile), 115.2 cm tall (3th percentile), and low bone density (11th percentile). Main clinical manifestations consist of intellectual disability, severely impaired speech (can only call dad and mom), high-arched palate, dental abnormalities (crowding and abnormal shape), small jaw, short tongue strap (received surgery), hypotonia, poor appetite, and salivation. She started walking at the age of 2. She started showing epilepsy since 6 years of age, even though her head MRI showed normal results at the age of 6. Her intellectual abilities were below an average 7 years old.

The karyotype analysis of G-banding in the peripheral blood of two siblings and parents showed a standard chromosomal pattern (data not shown). CMA result of both the proband and his sister using CytoScan™ HD whole genome SNP array (Affymetrix, USA) showed two copy number variations (CNVs): arr[hg19] 2q33.1(200,192,328-200,197,269)x1,2q35(218,105,663-218,816,675)x3. They have a 4.9-kb deletion in the 2q33.1 region of chromosome 2 (**Figure S1**). The deletion spans the exon 9 of the *SATB2* (OMIM: 608148) gene. They also have a 711-kb duplication (arr[hg19] 2q35(218,105,663-218,810,908)x3) in the 2q35 region, which contains *DIRC3* (OMIM: 608262) and *TNS1* (OMIM: 600076) genes. There has been no clear disease-related report on the duplication of this fragment or on these two genes. The mother's CMA result: arr[hg19] 5q31.3q33.1 (141,383,108-151,756,016)x2 hmz. The father's CMA results: arr[hg19] 2q35 (218,105,663-218,810,908)x3. The couple did not carry the 4.9-kb deletion at 2q33.1, while the 711-kb duplication at 2q35 of the children should have been inherited from their father. However, the possibility of a low-level somatic or gonadal mosaicism was not excluded. Meanwhile, the exact breakpoints were yet to be determined.

The pregnant woman received amniocentesis at 19 weeks of gestation. The karyotype of the fetus was normal, but CMA result was the same as those of the affected children. In addition, the B-ultrasound showed that the fetus had ventricular septal defect. After genetic counseling and through consideration, the couple chose to terminate the pregnancy at 23 weeks of gestation.

This study was carried out in accordance with the recommendations of the Ethics Committee of Women's Hospital, School of Medicine Zhejiang University. All subjects gave written informed consent in accordance with the Declaration of Helsinki, and written informed consent was also obtained from the guardians of the two affected siblings for the publication of this case report. The protocol was approved by the Ethics Committee of Women's Hospital, School of Medicine Zhejiang University.

Peripheral Blood DNA Extraction

Peripheral blood samples were collected from the family members including the husband's father (I1), the couple (II1 and II2), and their two children bearing SAS (III1 and III2) (**Figure 1A**). Genomic DNA was extracted from peripheral blood samples using the QIAGEN spin columns on a QIACube (QIAGEN GmbH) according to the manufacturer's instructions. All DNA samples were dissolved in water and stored at −20°C.

Gap-PCR and Primer Walking

Gap-PCR was chosen for locating exact breakpoints of the deletion. Three sets of primers (Gap2500-F/R, Gap1000-F/R, Gap500-F/R, **Table 1**) were designed according to the CMA results. Gap2500-F/R primer set worked well using TaKaRa LA Taq™ (TaKaRa Bio Inc.) and two-step procedure. The procedure of the PCR was as follows: 94°C for 1 min followed by 30 cycles at 98°C for 10 s, 68°C for 8 min, and a final extension step at 72°C

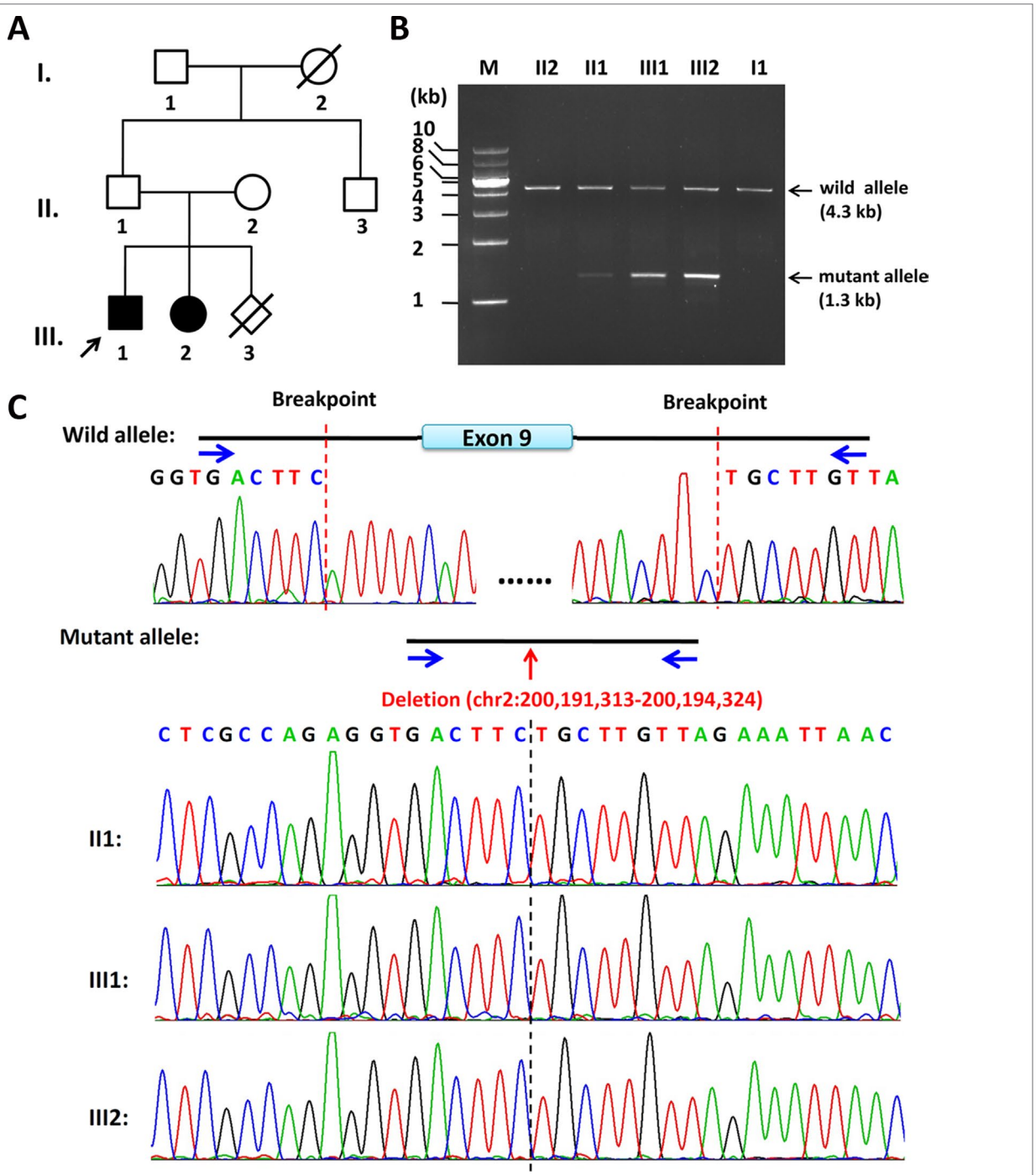

FIGURE 1 | Pedigree of the family and the characterization of the deletion breakpoints by gap-PCR. **(A)** Pedigree of this family. The arrow refers to the proband. The solid square (male) and circle (female) represent the two affected siblings; **(B)** Gap-PCR of family members using the peripheral blood-derived genomic DNA and primer sets (4334-F/R). Lanes 1–7: Marker, II2, II1, III1, III2, II3, and I1. The expected *SATB2* wild allele band of ~4.3 kb is seen in members of II2, II1, III1, III2, II3, and I1. A lower band, consistent with the size of mutant allele, is seen in II1, III1, and III2 but not in II2 or I1. Notice that the father (II1) has a much less intense band of about 1.3 kb in size, which is the same size as his two children (III1 and III2). **(C)** Sequencing results of the ~4.3-kb band and 1.3-kb band. The intragenic deletion border is chr2: 200,191,313-200,194,324 (hg19), which includes the entire exon 9 and its franking intronic sequences of the *SATB2* gene. Blue arrows refer to gap-PCR primer set (4334-F/R) used to detect the breakpoints of the 3kb deletion.

TABLE 1 | Gap-PCR primers.

Primer	Direction	Sequence (5'-3')	Product length (bp)
Gap2500-F	F	ATCAAATTGTCATTGTTGTGCC	9,376
Gap2500-R	R	CCATCTAAACCTCAGTTCCCTC	
Gap1000-F	F	ATAATACCTTCTCCTTCCCATC	6,575
Gap1000-R	R	TTAACAACTTGCCCAACTTACT	
Gap500-F	F	GTTAAAGGCCCATCAAAGGTAA	5,667
Gap500-R	R	GTTCGAGACCAGCATGGACAAC	
6666-F	F	GATAATCAATGGGAGATAATGG	6,666
6666-R	R	CCTTGTAAGTCAGTCTGGCACT	
5335-F	F	TCTTTCCCATAATAAACTCCAC	5,335
5335-R	R	TCTAAACCTCAGTTCCCTCATC	
4334-F	F	AAATGTCTTTGGCATCTGTTC	4,334
4334-R	R	CCATCTAAACCTCAGTTCCCT	
3442-F	F	ACAATTCTTCCCAAGTGCCTAC	3,443
3442-R	R	CATCTAAACCTCAGTTCCCTCA	

for 10 min. Based on this result, another four primer sets (6666-F/R, 5335-F/R, 4334-F/R, 3442-F/R, **Table 1**) were designed to amplify more suitable size of fragments for sequencing. The procedure of the PCR was as follows: 94°C for 1 min followed by 30 cycles at 98°C for 10 s, 68°C for 4 min, and a final extension step at 72°C for 10 min.

Sperm Preparation and Genomic DNA Extraction

The semen sample of the father was obtained by masturbation following 5 days of sexual abstinence and then was allowed to liquefy for 60 min at 37°C before processing. Semen sample was diluted 1:2 in Earle's solution, centrifuged at 500×g for 5 min. Then sperm pellets were washed twice with Earle's solution, and the final pellets were preserved immediately in liquid nitrogen for DNA extraction. Pellets were homogenized, and genomic DNA was extracted using the QIAGEN spin columns on a QIACube (QIAGEN GmbH) according to the manufacturer's instructions.

Real-Time Quantitative PCR

Three pairs of primers (**Figure 2A**) were designed to detect potential deletions in this family, including two primer sets outside the deletion region and one primer set inside the deletion. The design of primers is in accordance with the following standard: the length of the amplified product is less than 300 bp, and annealing temperature is close to 60°C. The three pairs of primers were designed by Primer-BLAST software (http://www.ncbi.nlm.nih.gov/tools/primer-blast/) (see **Table 2**).

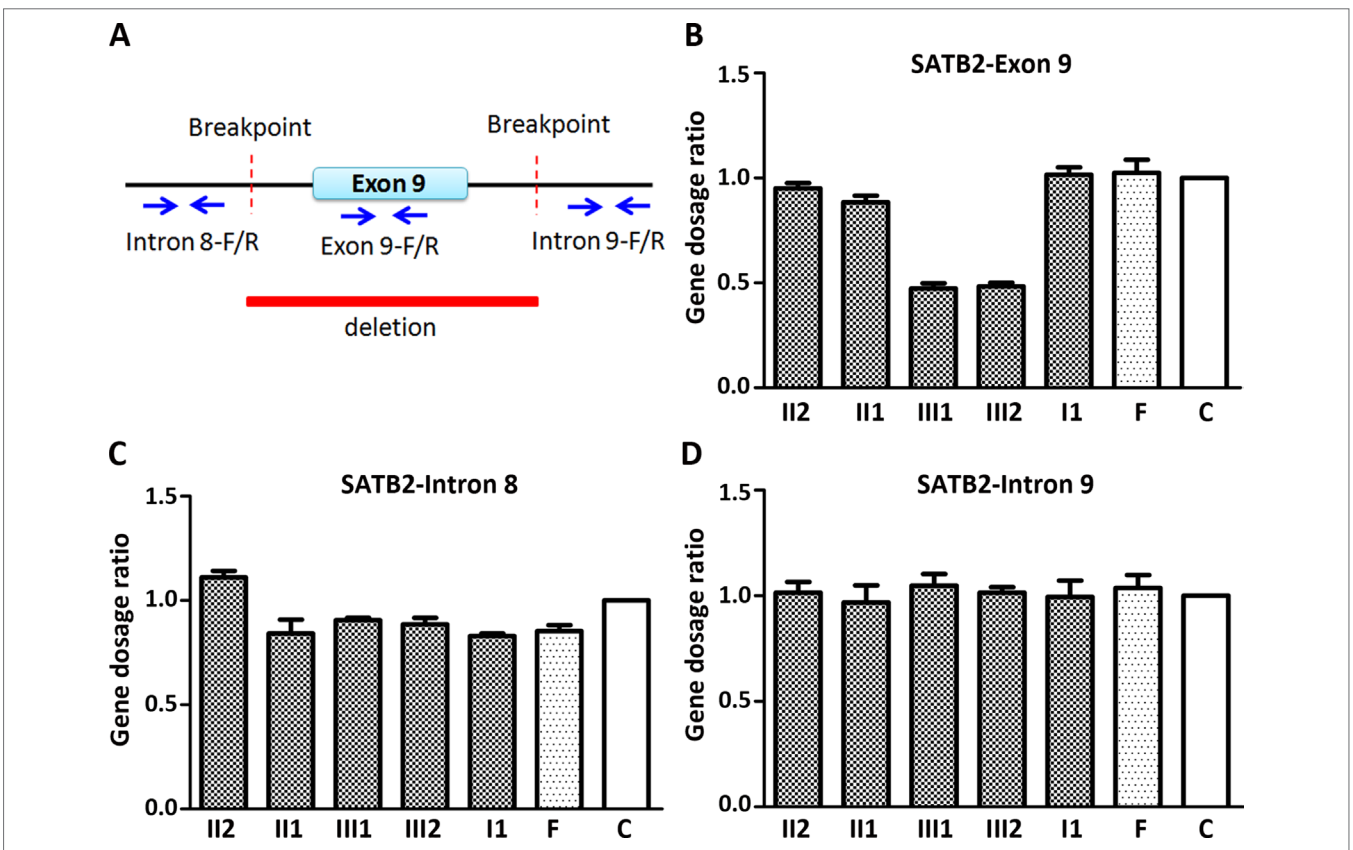

FIGURE 2 | Quantitative PCR (qPCR) of family members using the peripheral blood-derived genomic DNA. **(A)** Schematic diagram of location of qPCR primers. Blue arrows refer to qPCR primers; **(B)** Comparison of gene dosage ratios for the family members using the SATB2-exon 9 primers; **(C)** Comparison of gene dosage ratios for the family members using the SATB2-intron 8 primers; **(D)** Comparison of gene dosage ratios for the family members using the SATB2-intron 9 primers. F, normal female; C, control (normal male).

TABLE 2 | Quantitative PCR (qPCR) primers.

Primer	Direction	Sequence (5'-3')	Product length (bp)
Intron 8-F	F	ATGGAAATAGACCTGCACCTAC	186
Intron 8-R	R	TGACTCACCCAAATAGAAAGAT	
Exon 9-F	F	GATTCTGCGTAAGGAAGAAGACC	158
Exon 9-R	R	AGACCATGCTCACATTGGGATT	
Intron 9-F	F	ACCTAATAGCTTTCAGTGCCAGAC	130
Intron 9-R	R	TTCCCTGCTACACCTATCCCTA	

Specificity and effectiveness have been examined in preliminary experiment. Furthermore, β-globin was chosen as reference gene (forward primer: 5'-ACACAACTGTGTTCACTAGC-3'; reverse primer: 5'-CAACTTCATCCACGTTCACC-3'). The reaction was performed on a LightCycler 480 II real-time PCR system (Roche) in a final volume of 20 μl containing 10 μl of SYBR Premix Ex Taq Polymerase (TaKaRa), 0.4 μl each of the forward and reverse primers (10 μmol/l), and 2 μl of the genomic DNA (20 ng). Each sample was tested in triplets. The procedure of the PCR was as follows: 95°C for 10 s, followed by 40 cycles at 95°C for 5 s, 58°C for 30 s, and a final extension step at 72°C for 10 min. The dosage ratio of the tested samples was calculated by the ΔΔCt method for each test sample according to Livak and Schmittgen (2001).

Droplet Digital PCR

The somatic and gonadal mosaicism levels of the father were estimated using the Pilot Droplet Digital PCR system [Pilot Gene Technologies (Hangzhou) Co., Ltd. China]. The primer set SATB2Del-F:CAGAAAGTATCTGGGCCTATCAT/SATB2Del-R: GTACAAAAGAGCTGAAAACAAATACA, which was located on each side of the breakpoint, were used along with the 6-carboxyl-x-rhodamine (ROX)-labeled fluorescence probe SATB2Del-P: ROX-AGGTGACTTCTGCTTGTT-MGB, which spanned the deletion region to detect the copy number of the mutation. Another primer–probe set PARP2-F: GCGGAGGGAAGCTCATCAGTG/ PARP2-R: CCCTAGTCTCAGACCTTCCCAA/PARP2-P: 6-FAM-ACATGGGAGTGGAGTGACAGG-BHQ1 was employed to detect the single copy reference gene Poly (ADP-ribose) polymerase 2 (PARP2) in this study. The mosaic ratio was calculated as the copy number of mutation/the copy number of the reference gene.

The droplet generation, PCR amplification, and chip scanning were performed according to the manufacturer's instructions. Briefly, 20 μl of PCR mixture containing 1× PCR buffer, 1 μl each of the genomic DNA, 150 nM of 6-FAM-labeled probe, 600 nM of ROX-labeled probe, and 600 nM of each primer were loaded into Pilot droplet chips. Droplets were automatically generated in the Pilot droplet generator. Chips were then transferred into the Pilot iThermal1.0, and samples were amplified using the following cycling settings: initial denaturation at 95°C for 10 min, followed by 40 cycles of 95°C for 30 s, and 55°C for 60 s. Post-PCR chips were scanned using Pilot iScanner5 chip scanner, and results were analyzed using the GenePMSv1.1 software.

RESULTS

The Deletion Is of Paternal Origin and Includes Part of Intron 8, the Entire Exon 9, and Part of Intron 9 of *SATB2*

By the clinical diagnosis and CMA, we narrowed down the target of interest to *SATB2*. To make sure the exact breakpoints of *SATB2* mutation, three primer sets were designed outside the region indicated by the CMA result. One of the three gap-PCR primer sets (Gap2500-F/R) worked well, and gel electrophoresis results showed both father, and the two siblings share two amplified bands (**Figure S2**). The larger band is about 10 kb in size, and the smaller band is about 7 kb in size. The 7-kb band of the father is less intense than that of the two siblings. Based on this result, another four primer sets were designed to amplify fragments with proper lengths for Sanger sequencing. Gap-PCR products using primers (4334-F/R) of all the family members shared the 4.3-kb band, while gap-PCR products of II2 (father), III1 (the proband), and III2 (the sister) have another smaller band (~1.3kb) (**Figure 1B**). The Sanger sequencing result of the smaller band of the father and the two siblings revealed that the breakpoints were located at chr2: 200,191,313-200,194,324 (hg19) when blasted with genomic sequence (NM_001172517). The deleted region includes part of intron 8, the entire exon 9, and part of intron 9 (**Figure 1C**). Given that the father has no phenotype and the father carries a smaller band of lower intensity comparing with those of the two siblings, the father might have a low level of mosaicism.

Deletion Is Present in Peripheral Blood and Semen of the Father

In order to know the likelihood of recurrence of the deletion in the next pregnancy of the mother, gap-PCR using primers 4334-F/R was performed on semen-derived DNA of the father. Gel electrophoresis results that showed both amplified products of semen-derived DNA and peripheral-blood DNA of the father share two amplified bands with same sizes (**Figure S3**). We compared the smaller mutant band between the amplified products of both DNA samples by Sanger sequencing, and the sequencing results showed they were all deleted mutant (data not shown). We found that the father has not only somatic deletion but also germline deletion (since it is found in the semen); hence, the risk of the fetus carrying the same deletion in the next pregnancy still exists.

The Two Siblings Had One Copy in the Deletion Region by Quantitative PCR

Since CMA and gap-PCR are having conflicting results regarding the deletion of father, quantitative PCR (qPCR) was carried out using three primer sets (**Figure 2A**) in the family members to determine which test result is correct. Because of the deletion region including part of intron 8, the entire exon 9, and part of intron 9, we examined the exon 9 (representative of deletion region) and intron 8 and 9 (representative of non-deletion region) of the *SATB2* gene by the qPCR. The level of exon 9 in

the two siblings is about half of the normal level in other family members (**Figure 2B**). The result of intron 8 and 9 showed no obvious differences in the family members (**Figure 2C** and **D**). The above results confirmed that both CMA and gap-PCR were right about the level of deletion in the two siblings but failed to justify the gap-PCR results about the father's deletion level. Further confirmation on father's deletion level is required.

The Father Was Low-Level Somatic and Gonadal Mosaicism in the *SATB2* Gene

In order to check the level of somatic and gonadal mosaicism of the *SATB2* gene in the father, ddPCR was used to estimate the frequencies of the mutant allele in this study. As shown in **Figure 3** and **Table 3**, compared with the frequencies of no-template control (**Figure 3A**), the mutant frequencies of the proband and his sister were 52.88% (**Figure 3C**) and 52.11% (**Figure 3D**), respectively, indicating heterozygosity on the detected site for both of them, while their mother has zero mutant frequency (**Figure 3B**), signifying wild type on the detected site. Two DNA samples from peripheral blood and semen of the proband's father were analyzed by ddPCR in this study, which showed the mutant frequencies of 13.16% (**Figure 3E**) and 16.68% (**Figure 3F**), respectively. Above results confirmed the low-level mosaicism in the *SATB2* gene of the father. Gap-PCR and ddPCR together provide an effective diagnostic technique for this intragenic deletion.

DISCUSSION

Glass syndrome or 2q33.1 microdeletion syndrome (OMIM#612313) was first described by Glass et al. (1989) in a 16-year-old male with

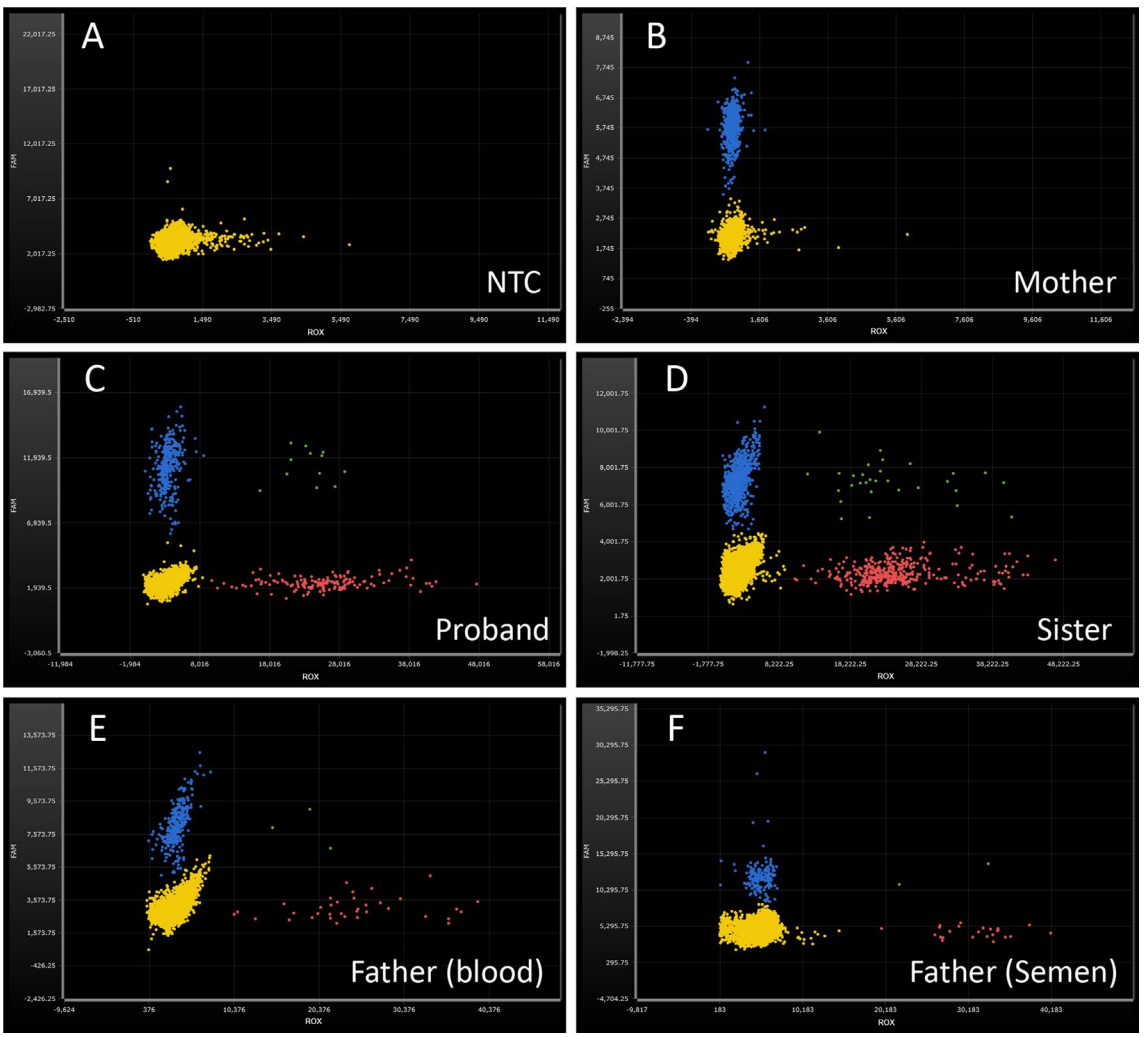

FIGURE 3 | Droplet digital PCR analysis for the mutant frequencies of *SATB2* gene. ROX-labeled probe detects the delete mutation of *SATB2* gene (red dots). 6-FAM-labeled probe detects the reference gene PARP2 (blue dots). Green dots, ROX and 6-FAM double positive. Yellow dots, no amplification. **(A)** No-template control, NTC; no amplification on both fluorescence channels. **(B)** wild type, DNA from the proband's mother; normal amplification on FAM channel, while no amplification on ROX channel. **(C, D)** heterozygosity, DNA from the proband and his sister, respectively; the ratios of ROX/FAM were close to 0.50. **(E, F)** mosaicism, DNA from the proband's peripheral blood and semen of the father; the ratios of ROX/FAM were 13.16% and 16.68%, respectively.

TABLE 3 | Mutant frequencies of *SATB2* gene detected by Droplet Digital PCR.

Subject	PARP2 concentration (FAM) copies/μl	SATB2 concentration (ROX) copies/μl	Mutant frequency (ROX/FAM, %)
Father (blood)	48.080	6.326	13.16
Father (semen)	21.978	3.667	16.68
Mother	147.06	0	0
Proband	65.415	34.595	52.88
Sister	62.813	32.733	52.11

cytogenetically visible deletion of 2q32.2-q33.1, and he had severe intellectual disability, cleft palate, short stature, and craniofacial dysmorphism (Glass et al., 1989). This syndrome has a large deletion of 2q33.1 that includes the disease-causing gene *SATB2* and other genes. To reduce confusion and unify the nomenclature, a new term SATB2-associated syndrome (SAS) emerged. To date, intragenic deletions and duplications, point mutations involving *SATB2* gene have been reported as pathogenic SAS mutations in the literature (Zarate and Fish, 2017; Zarate et al., 2017; Zarate et al., 2018a).

To our knowledge, deletions have been reported in 22 patients thus far and ranged from 10 to 317 kb in size and with only *SATB2* included in the region (Rosenfeld et al., 2009; Balasubramanian et al., 2011; Zarate et al., 2018a; Zarate et al., 2019). Herein, the 3-kb intragenic deletion in the family of this study might be the smallest reported deletion in the *SATB2* gene thus far, and it is also covered within the 10-kb deletion reported in case ID 49 by Zarate et al. (Zarate et al., 2018a). ID 49 was an 18-year-old male carrying a 10-kb deletion (chr2:200,190,560-200,200,832, hg19) including the entire exon 9 and part of intron 8 and 9 sequences within the *SATB2* gene. He had growth retardation, cleft palate, feeding difficulties during infancy, sialorrhea, and abnormal MRI, and he started walking at the age of 12. We note that cleft palate and feeding difficulties during infancy were absent in our case, indicating that these characteristics may be variable in individuals affected by heterozygous deletion of *SATB2* exon 9. Sequence of exon 9 (coding 393–461 amino acids) covers C-terminal part of CUT1 domain (355–434). Among all SAS families with trio analysis by whole-exome sequencing or panels, Zarate and colleagues found that the most frequent pathogenic variant were missense mutations, and most of these variants localized to the CUT1 domain (Zarate et al., 2018a). That might suggest the importance of CUT1 (including exon 8 and 9) for the function of SATB2 protein. On the other hand, Bengani et al. constructed several mutant SATB2 plasmids and then transfected to Hela and human fibroblast cell lines. They found that cells with p.Arg389Cys change in CUT1 domain displayed more diffuse patterns of nuclear localization than the wild type, and the authors proposed that CUT1 domain may be required to initiate interaction with chromatin or matrix (Bengani et al., 2017). Therefore, considering the significance of CUT1 domain, exon 9 may be crucial to SAS symptoms.

Of the SAS-affected families reported thus far with trio analysis, two instances of mosaicism have now been documented in the literature. One case was presumed to be gonadal mosaicism because two affected siblings had the same mutation (c.582-2A>G), but none was detected in the leukocyte DNA of either parent (Bengani et al., 2017). Another case was also suspected to be gonadal mosaicism for the two affected siblings share an identical 2.9-Mb micro-duplication that was not present in either parent (Bengani et al., 2017; Zarate et al., 2018a). In the present study, the father with normal phenotype in this family was proved to have a low-level somatic and gonadal mosaicism rather than both siblings carrying *de novo* mutations in the *SATB2* gene. The deletion in our case failed to detect due to low level of resolution in this specific region by SNP array (**Figure S1**); more details were only revealed with the help of gap-PCR and ddPCR. We recommend paternal semen sample testing as the first step because the acquisition of semen sample is easy and safe. If the results are negative, mutation from maternal germline origin could not be ruled out.

The real-time qPCR is another common method to determine copy number of target genes in addition to CMA; however, it does have several limitations such as failure to detect low-level mosaicism variations. This limitation can be overcome by ddPCR due to its ability to precisely quantify mosaic genomic copy number variations when compared with qPCR (Fujiki et al., 2016; Taylor et al., 2017; Zhou et al., 2018). The proband's father was mosaic in *SATB2*, estimated to be 13.16% in the peripheral blood and 16.68% in the semen by ddPCR. With mosaicism detected in semen as well, recurrence risks are higher compared with those phenotypes occurring from *de novo* variants. It is difficult to estimate an accurate recurrence risk if the percent mosaicism in the gonad is unknown.

Our study further validates the need to perform trio analysis if two phenotypically normal parents had two or more consecutive pregnancies with similar SAS phenotypes. In terms of parental testing, CMA and traditional qPCR are not sensitive enough as exemplified in our study. Gap-PCR and ddPCR are more sensitive for detecting low-level mosaicism, and ddPCR may further be used for estimation of mutant frequencies, if needed. The different results have greatly different reproductive recurrence risks, that is, apparently, *de novo SATB2* pathogenic variants confer very low recurrence risk versus paternal mosaicism that confers up to a 50% recurrence risk.

In conclusion, CMA is recommended as the first-tier test followed by gap-PCR and ddPCR if a similar situation is arising. We highlighted the importance of detailed genetic testing, testing method, and counseling for cases of somatic or gonadal mosaicism in an unaffected parent of children with SAS syndrome. Due to inconclusiveness of CMA results alone, we suspect that this inheritance might be underreported and the results would have a direct impact on reproductive planning and prenatal diagnosis.

DATA AVAILABILITY STATEMENT

All datasets are included in the manuscript or supplementary files.

ETHICS STATEMENT

This study was carried out in accordance with the recommendations of Ethics Committee of Women's Hospital, School of Medicine Zhejiang University, with written informed consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the Ethics Committee of Women's Hospital, School of Medicine Zhejiang University.

AUTHOR CONTRIBUTIONS

MD, YQ, and JL conceived of the study, participated in its design, interpreted the molecular results, and drafted the manuscript; YY carried out the qPCR; CJ, PC, and YL collected the samples and clinical data; YQ and MC carried out the CMA, gap-PCR, and ddPCR; HP helped to revise the manuscript. All authors have read and approved the final manuscript.

FUNDING

The work was supported by the National Key Research and Development Program of China (2018YFC1004900, 2016YFC1000703), Key Research and Development Program of Zhejiang Province (2019C03025), Zhejiang University—Lishui Cooperation Project (2018zdhz08), the National Natural Science Foundation of China (81300532, 81801441), and Zhejiang Key Laboratory of Organ Development and Regeneration (ZX15001017014).

ACKNOWLEDGMENTS

We are grateful to Dr. Jiang Xia and Dr. De-Kun Dong (Pilot Gene Technologies, Hangzhou 310006, China) and Dr. Jiong Gao (BGI Genomics, BGI-Shenzhen 518083, China) for analyzing the data and modifying the article. We also would like to thank our patients for agreeing to donate their personal data to our study and have these published.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2019.00630/full#supplementary-material

FIGURE S1 | CMA analysis of the proband, his sister, and their father and mother. The red arrows refer to the deletion. The black arrows refer to the normal copy in the deletion region.

FIGURE S2 | Gap-PCR of the core family members using primer set (Gap2500-F/R). Lanes 1–5: Marker, II2, II1, III1, III2. Notice that II2 has one band about 10 kb in size, while the other three members have another band about 7 kb in size.

FIGURE S3 | Gap-PCR of the father (II2) revealed heterozygous deletion in peripheral blood-derived and semen-derived DNA using primer set (4334-F/R). Lanes 1–3: Marker, blood, sperm. Notice that lane 3 has a lower band about 1.3 kb in size as lane 2. The PCR condition was as the same with that described in "MATERIALS AND METHODS" except the amplification cycles was changed to 34 cycles to obtain more PCR products for sequencing.

REFERENCES

Balasubramanian, M., Smith, K., Basel-Vanagaite, L., Feingold, M. F., Brock, P., Gowans, G. C., et al. (2011). Case series: 2q33.1 microdeletion syndrome–further delineation of the phenotype. *J. Med. Genet.* 48, 290–298. doi: 10.1136/jmg.2010.084491

Bengani, H., Handley, M., Alvi, M., Ibitoye, R., Lees, M., Lynch, S. A., et al. (2017). Clinical and molecular consequences of disease-associated de novo mutations in SATB2. *Genet. Med.* 19, 900–908. doi: 10.1038/gim.2016.211

Docker, D., Schubach, M., Menzel, M., Munz, M., Spaich, C., Biskup, S., et al. (2014). Further delineation of the SATB2 phenotype. *Eur. J. Hum. Genet.* 22, 1034–1039. doi: 10.1038/ejhg.2013.280

Fitzpatrick, D. R., Carr, I. M., Mclaren, L., Leek, J. P., Wightman, P., Williamson, K., et al. (2003). Identification of SATB2 as the cleft palate gene on 2q32-q33. *Hum. Mol. Genet.* 12, 2491–2501. doi: 10.1093/hmg/ddg248

Fujiki, K., Shirahige, K., Kaur, M., Deardorff, M. A., Conlin, L. K., Krantz, I. D., et al. (2016). Mosaic ratio quantification of isochromosome 12p in Pallister-Killian syndrome using droplet digital PCR. *Mol. Genet. Genomic Med.* 4, 257–261. doi: 10.1002/mgg3.200

Glass, I. A., Swindlehurst, C. A., Aitken, D. A., Mccrea, W., and Boyd, E. (1989). Interstitial deletion of the long arm of chromosome 2 with normal levels of isocitrate dehydrogenase. *J. Med. Genet.* 26, 127–130. doi: 10.1136/jmg.26.2.127

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262

Rosenfeld, J. A., Ballif, B. C., Lucas, A., Spence, E. J., Powell, C., Aylsworth, A. S., et al. (2009). Small deletions of SATB2 cause some of the clinical features of the 2q33.1 microdeletion syndrome. *PLoS One* 4, e6568. doi: 10.1371/journal.pone.0006568

Sheehan-Rooney, K., Palinkasova, B., Eberhart, J. K., and Dixon, M. J. (2010). A cross-species analysis of Satb2 expression suggests deep conservation across vertebrate lineages. *Dev. Dyn.* 239, 3481–3491. doi: 10.1002/dvdy.22483

Taylor, S. C., Laperriere, G., and Germain, H. (2017). Droplet Digital PCR versus qPCR for gene expression analysis with low abundant targets: from variable nonsense to publication quality data. *Sci. Rep.* 7, 2409. doi: 10.1038/s41598-017-02217-x

Zarate, Y. A., Bosanko, K. A., Caffrey, A. R., Bernstein, J. A., Martin, D. M., Williams, M. S., et al. (2019). Mutation update for the SATB2 gene. *Hum. Mutat.* (in press). doi: 10.1002/humu.23771

Zarate, Y. A., and Fish, J. L. (2017). SATB2-associated syndrome: Mechanisms, phenotype, and practical recommendations. *Am. J. Med. Genet. A* 173, 327–337. doi: 10.1002/ajmg.a.38022

Zarate, Y. A., Kalsner, L., Basinger, A., Jones, J. R., Li, C., Szybowska, M., et al. (2017). Genotype and phenotype in 12 additional individuals with SATB2-associated syndrome. *Clin. Genet.* 92, 423–429. doi: 10.1111/cge.12982

Zarate, Y. A., Smith-Hicks, C. L., Greene, C., Abbott, M. A., Siu, V. M., Calhoun, A., et al. (2018a). Natural history and genotype-phenotype correlations in 72

individuals with SATB2-associated syndrome. *Am. J. Med. Genet. A* 176, 925–935. doi: 10.1002/ajmg.a.38630

Zarate, Y. A., Steinraths, M., Matthews, A., Smith, W. E., Sun, A., Wilson, L. C., et al. (2018b). Bone health and SATB2-associated syndrome. *Clin. Genet.* 93, 588–594. doi: 10.1111/cge.13121

Zhou, B., Haney, M. S., Zhu, X., Pattni, R., Abyzov, A., and Urban, A. E. (2018). Detection and quantification of mosaic genomic DNA variation in primary somatic tissues using ddPCR: Analysis of mosaic transposable-element insertions, copy-number variants, and single-nucleotide variants. *Methods Mol. Biol.* 1768, 173–190. doi: 10.1007/978-1-4939-7778-9_11

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Genetics

ORIGINAL RESEARCH
published: 10 July 2019
doi: 10.3389/fgene.2019.00605

De Novo Germline Mutations in *SEMA5A* Associated With Infantile Spasms

Qiongdan Wang[1,2†], Zhenwei Liu[3†], Zhongdong Lin[4], Ru Zhang[2], Yutian Lu[1,2], Weijue Su[5], Feng Li[4], Xi Xu[1,2], Mengyun Tu[1,2], Yongliang Lou[2,6] Junzhao Zhao[5*] and Xiaoqun Zheng[1,2,6*]*

[1] Department of Laboratory Medicine, The Second Affiliated Hospital and Yuying Children's Hospital of Wenzhou Medical University, Wenzhou, China, [2] School of Laboratory Medicine and Life Sciences, Wenzhou Medical University, Wenzhou, China, [3] Institute of Genomic Medicine, Wenzhou Medical University, Wenzhou, China, [4] Department of Pediatric Neurology, The Second Affiliated Hospital and Yuying Children's Hospital, Wenzhou Medical University, Wenzhou, China, [5] Department of Obstetrics and Gynecology, The Second Affiliated Hospital and Yuying Children's Hospital, Wenzhou Medical University, Wenzhou, China, [6] Key Laboratory of Laboratory Medicine, Ministry of Education, Wenzhou, Zhejiang, China

OPEN ACCESS

Edited by:
Fan Jin,
Zhejiang University, China

Reviewed by:
Yueqiu Tan,
Central South University, China
X Frank Yang,
Indiana University Bloomington, United States
Inês Barroso,
University of Cambridge, United Kingdom

***Correspondence:**
Xiaoqun Zheng
jszhengxq@163.com
Junzhao Zhao
z.joyce08@163.com
Yongliang Lou
louyongliang2013@163.com

†These authors have contributed equally to this work.

Specialty section:
This article was submitted to Genetic Disorders, a section of the journal Frontiers in Genetics

Received: 04 December 2018
Accepted: 07 June 2019
Published: 10 July 2019

Citation:
Wang Q, Liu Z, Lin Z, Zhang R, Lu Y, Su W, Li F, Xu X, Tu M, Lou Y, Zhao J and Zheng X (2019) De Novo Germline Mutations in SEMA5A Associated With Infantile Spasms. Front. Genet. 10:605. doi: 10.3389/fgene.2019.00605

Infantile spasm (IS) is an early-onset epileptic encephalopathy that usually presents with hypsarrhythmia on an electroencephalogram with developmental impairment or regression. In this study, whole-exome sequencing was performed to detect potential pathogenic *de novo* mutations, and finally we identified a novel damaging *de novo* mutation in *SEMA5A* and a compound heterozygous mutation in *CLTCL1* in three sporadic trios with IS. The expression profiling of *SEMA5A* in the human brain showed that it was mainly highly expressed in the cerebral cortex, during the early brain development stage (8 to 9 post-conception weeks and 0 to 5 months after birth). In addition, we identified a close protein-protein interaction network between *SEMA5A* and candidate genes associated with epilepsy, autism spectrum disorder (ASD) or intellectual disability. Gene enrichment and function analysis demonstrated that genes interacting with *SEMA5A* were significantly enriched in several brain regions across early fetal development, including the cortex, cerebellum, striatum and thalamus ($q < 0.05$), and were involved in axonal, neuronal and synapse-associated processes. Furthermore, *SEMA5A* and its interacting genes were associated with ASD, epilepsy syndrome and developmental disorders of mental health. Our results provide insightful information indicating that *SEMA5A* may contribute to the development of the brain and is associated with IS. However, further genetic studies are still needed to evaluate the role of *SEMA5A* in IS to definitively establish the role of *SEMA5A* in this disorder.

Keywords: epileptic encephalopathy, infantile spasms, *de novo* mutations, *SEMA5A*, whole-exome sequencing

INTRODUCTION

Epileptic encephalopathies (EEs) are a group of complex brain disorders characterized by intractable and early-onset epilepsy with or without developmental delays, which are highly genetically heterogeneous (Berg et al., 2010; Scheffer et al., 2017). Infantile spasms (IS, also known as West syndrome) are considered a subset of EEs that are characterized by the early onset of epileptic spasms, typically in the first year of life, and are always accompanied by a hypsarrhythmia pattern on the electroencephalogram (EEG) and developmental impairment (Nabbout and Dulac, 2003; Pavone et al., 2014). Moreover, EEs are caused by a variety of etiologies that are not yet fully

understood. However, recent studies have provided evidence that genetic factors have a tremendous impact on the pathogenesis of EEs (McTague et al., 2016; Shbarou and Mikati, 2016). The discovery of candidate genes will further our understanding of the mechanisms underlying epileptogenesis.

The development of next-generation sequencing techniques, such as whole-exome sequencing (WES), has greatly facilitated gene discovery in EEs, and these approaches have been widely used to detect pathogenic mutations (Dimassi et al., 2016; Jin et al., 2018). Furthermore, recent genetic studies have used trio exome sequencing to confirm that many pathogenetic *de novo* mutations (DNMs) are critical genetic components in the pathogenesis underlying EEs (Claes et al., 2001; Carvill et al., 2013; Epi4K Consortium and Epilepsy Phenome/Genome Project, 2013). DNMs are the most damaging form of rare genetic mutations and occur mainly in the germline (Veltman and Brunner, 2012). Most genes with pathogenetic DNMs in EE encode voltage-gated ion channels or receptors associated with neurotransmitter, including sodium channels, potassium channels, GABA receptors, glutamate receptors and NMDA receptors (Fukata and Fukata, 2017; He et al., 2019), and are often involved in a variety of functional pathways related to neuronal excitability or synaptic and neuronal connectivity (Pardo et al., 2014).

Axon guidance proteins, including semaphorins, ephrins, slits, repulsive guidance molecules, and netrins, can act as attractants or repellents during axon branching, synapse formation and plasticity (Shen and Cowan, 2010; Van Battum et al., 2015). In the central nervous system, semaphorins play a role in synaptic plasticity including the regulation of synaptic structures and synaptic transmission (Pasterkamp and Giger, 2009). Moreover, both NMDA and AMPA receptors participate in synaptic plasticity (Song and Huganir, 2002; Lau and Zukin, 2007). Strikingly, DNMs have been identified in the *GRIN1* and *GRIN2B* genes, which encode the proteins that form the subunits of the NMDA receptor, in individuals with West syndrome and a severe developmental delay (Epi4K Consortium and Epilepsy Phenome/Genome Project, 2013; Lemke et al., 2014). Because appropriate brain function and mental health rely on the accurate regulation of the central nervous system synapse density, an imbalance between excitatory and inhibitory synaptic transmission is partly responsible for neurodevelopmental disorders such as schizophrenia and autism spectrum disorder (ASD) characterized by impairments in social interactions and restricted behaviors and interests (Penzes et al., 2011).

In this study, we performed WES on three individuals with IS as well as their unaffected parents and identified a novel damaging DNM in *SEMA5A*. As a member of the semaphorin gene family with bifunctional axon guidance activities, the discovery of a mutation in *SEMA5A* may further strengthen the role of axon guidance proteins in EEs. In addition, we also detected a DNM in *PLEKHG4B* and a compound heterozygous mutation in *CLTCL1* in patients with IS.

MATERIALS AND METHODS

Subjects

Three probands with IS and their healthy parents were enrolled in the study from the Second Affiliated Hospital and Yuying Children's Hospital of Wenzhou Medical University, after study approval was provided by the Hospital Ethics Committee. Moreover, written informed consent of all participants was obtained from their parents or guardians at the time of recruitment. All probands were referred with typical seizure presentation and were diagnosed with IS by an experienced pediatric neurologist.

Whole-Exome Sequencing

Peripheral blood (2 ml) was drawn from the three affected probands and their unaffected family members. Genomic DNA was isolated using a QIAGEN DNeasy Blood & Tissue Kit (Qiagen, Valencia, CA, USA) from the peripheral blood of each included individual. The DNA was then subjected to an additional quality and quantity evaluation step using a NanoDrop 2000 spectrophotometer (Thermo Scientific, Wilmington, DE, USA). Subsequently, exome-coding DNA was captured with an Agilent SureSelect Human All Exon v6 Kit (Agilent Technologies, Santa Clara, CA, USA), and the libraries were sequenced on an Illumina HiSeq2000 sequencer (San Diego, CA, USA), which produced 150-bp paired-end reads.

All raw sequencing data obtained from these three trios were analyzed in a similar manner according to a customized bioinformatics pipeline (Wang et al., 2013). After quality filtering was performed on the raw sequence data using the Trim Galore program, the cleaned sequences were aligned to the human reference genome (GRCH37/hg19) using Burrows-Wheeler Aligner (BWA). Picard was performed to realign the reads and remove any reads that were duplicated or mapped to multiple genome locations. Variant and genotype calling were performed using the Genome Analysis Toolkit (GATK) (McKenna et al., 2010), and DNMs were detected by two software tools [ForestDNM (Michaelson et al., 2012) and mirTrios (Li et al., 2015)].

Mutation Annotation

ANNOVAR was used to annotate all called variants. The minor allele frequency (MAF) of the detected sequence variants was estimated in various publicly available databases, including ExAC, UK10K, dbSNP147, 1000 Genomes, and ESP6500. If a detected variant was presented with a MAF > 0.1% in any database, it was eliminated. Subsequently, the effects of the detected variants were predicted according to SIFT (https://sift.bii.a-star.edu.sg/, a variant with SIFT score < 0.05 predicted damaging), VEST3 (https://karchinlab.org/apps/appVest.html, a variant with VEST3 score ≥ 0.5 indicated damaging), Polyphen2 (http://genetics.bwh.harvard.edu/pph2/, a variant with a score of 0.909 to 1.0 indicated probable damage, while those with scores between 0.0 to 0.446 meant benign), and GERP++ (http://mendel.stanford.edu/SidowLab/downloads/gerp/, a variant with a GERP++ score ≥ 2 indicated conserved). All potential damaging DNMs were visualized by Splicing Viewer (Liu et al., 2012) and confirmed by Sanger sequencing.

Gene Expression

The RNA sequencing data for human tissue-derived *SEMA5A* were obtained from the human protein atlas (HPA) (https://www.proteinatlas.org). In addition, data were obtained from the human

brain transcriptome at HBT (http://hbatlas.org/) to evaluate the spatial and temporal expression pattern of *SEMA5A* in human brain tissues. Furthermore, a digital atlas of gene expression patterns in mouse tissues at embryonic day 14.5 was selected and assessed in GenePaiant (www.genepaint.org). Expression patterns were defined by non-radioactive *in situ* hybridization (ISH), and selected images were annotated in detail.

Construction of the Protein–Protein Interaction Network and Enrichment Analysis of Genes in the Network

A critical assessment and integration of protein–protein interaction (PPI) information was obtained from the STRING database to create the PPI networks. Then, we selected genes that interacted with *SEMA5A* and that had interaction scores greater than 200. Specific expression analysis (SEA) was performed on a developed online tool (http://genetics.wustl.edu/jdlab/csea-tool-2/) to explore whether *SEMA5A* and the identified interacting genes were highly enriched in a particular human brain region. In addition, R package clusterProfiler and DOSE were used to further explore the functions of the interacting genes in the PPI network. There were 43 *SEMA5A*-interacting genes that were shared by neuropsychiatric disorders and used to construct the interconnected PPI network. Random simulations of 100,000 permutations for genes and connections were performed to verify the non-random nature of the networks. Moreover, fisher's exact test was used to evaluate the enrichment of 43 genes in the postsynaptic density (PSD) proteins (Collins et al., 2006) and genes under evolutionary constraint (Samocha et al., 2014).

RESULTS

Identification of Potential Mutations

To detect candidate pathogenic DNMs, we recruited three trios with IS for WES. After low-quality reads and adapters were removed, approximately 5.2–11.03 GB of clean reads were obtained for each individual. All samples passed the quality control measures, and more than 99% of the qualified bases were aligned to the human reference genome (hg19) with an average sequencing depth of 74.91-fold. It is important to note that a mean of 87.49% of the target regions was covered at 20× and that more than 90% of the exonic regions were sequenced at 10× coverage (**Supplementary Table S1**). Following alignment and variant annotation, we detected a total of five *de novo* SNVs and three *de novo* indel mutations in coding regions. We confirmed that two DNMs and a compound heterozygous mutation were present among the three trios. In addition, all of the identified mutations were confirmed by Sanger sequencing, and the deleteriousness of mutations were predicted by SIFT, VEST3, Polyphen2, and GERP++.

Candidate Genes for Infantile Spasms

Patient 1, who was born at term following an uneventful pregnancy to healthy non-consanguineous parents with no pertinent family history, had seizure symptom onset at the age of 7 months with the EEG displaying hypsarrhythmia (**Figure 1A**), a characteristic feature of IS. Additionally, the magnetic resonance imaging (MRI) of patient 1 at 7 months old, revealed brain dysplasia and agenesis of the corpus callosum (**Figures 1B**, **C**). The patient carried one missense mutation (c.1201C > T) in the *SEMA5A* gene that was identified by WES and confirmed as a DNM by Sanger sequencing (**Figure 2A**). At the structural level, the *de novo* missense mutation caused an arginine to cysteine substitution at amino acid 401 in the Sema domain [p.(Arg401Cys)], a conserved domain containing 450 amino acids (**Figure 2B**). The Sema domain serves as a repulsive guidance cue and might prevent axons from branching into surrounding myotome regions (Hilario et al., 2009). Then we predicted the structural model of the SEMA5A and there was no hydrogen bond between Arg-401 of SEMA5A and other residues. However, Cys-401 forms three hydrogen bonds with Phe-402, Asp-425, and Gln-292, respectively (**Supplementary Figure S1**). Hence, the substitution of arginine by cystine at amino acid 401 of SEMA5A may affect the structural stability.

In addition, we found that the identified DNM is evolutionarily conserved across various vertebrates (**Figure 2C**). Moreover, in silico prediction programs (SIFT, VEST3, Polyphen2 and GERP++) indicated that the *de novo* missense mutation in *SEMA5A* identified in this study is likely to be damaging and conserved in the protein (**Table 1**). Furthermore, the missense Z score of *SEMA5A* is 1.6794, which means that it tended to be intolerant of functional genetic variation (Lek et al., 2016). It should be mentioned that five different mutations, including a *de novo* missense mutation [c.2852C > G, p.(Ser951Cys)], a nonsense mutation [c.586C > T p.(Arg196*)] and three missense mutations, have been found in *SEMA5A* in individuals diagnosed with ASD (**Figure 2B**) (Iossifov et al., 2014; D'Gama et al., 2015; Mosca-Boidron et al., 2016). Additionally, *SEMA5A* has not been reported to be involved in epilepsy and is absent in any gene database related to epilepsy (Ran et al., 2015).

We additionally identified a compound heterozygous mutation in *CLTCL1* in patient 2 (**Supplementary Figure S2A**). In this patient, one missense mutation was inherited from the mother [c.3946A > G, p.(Met1316Val)], and was predicted to be damaging and conserved by VEST3, Polyphen2 and GERP++. Another missense mutation [c.3493C > T, p.(Arg1165Cys)] that was inherited from the father was predicted by four prediction tools (SIFT, VEST3, Polyphen2, and GERP++) to be damaging and conserved (**Supplementary Figure S2B**, **Table 1**). Both mutations were located in a highly conserved domain (**Supplementary Figure S2C**).

In patient 3, we identified a missense mutation in *PLEKHG4B* [c.2740G > A, p.(Ala914Thr)] that was regarded as a DNM and confirmed by Sanger sequencing. Nevertheless, the SIFT, VEST3 and Polyphen2 tools predicted that this mutation is tolerable and benign, and GERP++ predicted that it is non-conserved. Therefore, it was not considered as a possible candidate mutation in further analyses.

The Expression Profiles of SEMA5A

Considering that IS can give rise to severe cognitive and behavioral impairments, we attempted to determine the

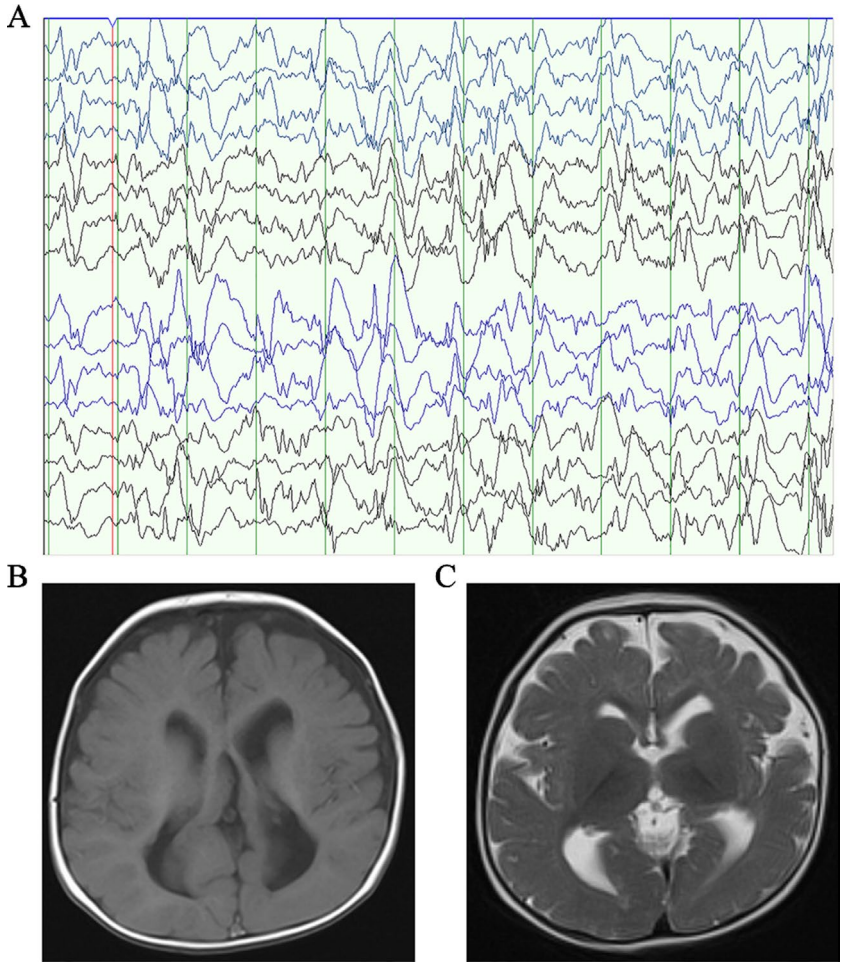

FIGURE 1 | Clinical presentation. **(A)** The electroencephalogram obtained in patient 1 when he was 7 months old. **(B)** Axial T1-weighted brain MRI obtained in patient 1 when he was 7 months old. **(C)** Axial T2-weighted brain MRI obtained in patient 1 when he was 7 months old.

role of *SEMA5A* in the development of brain tissues by evaluating its expression pattern. RNA sequencing data were available for 36 tissues in the HPA database and indicated that *SEMA5A* acts in a variety of tissues. It is important to note that *SEMA5A* is highly expressed in the cerebral cortex, in which it showed its fourth-highest expression level in the 36 tissues (**Figure 3A**). To affirm this finding and to gain a higher-resolution spatiotemporal view of *SEMA5A* expression in the human brain, we analyzed RNA-seq data across various brain regions and developmental stages in HBT. The results showed that *SEMA5A* expression is extensively scattered across different developmental periods and regions of the human brain. Interestingly, *SEMA5A* is preferentially highly expressed in 11 areas of the neocortex, as well as several other brain regions such as the hippocampus (HIP), amygdala (AMY), and the striatum (STR), during early embryonic development (8 to 9 post-conception weeks, **Figure 3B**). In addition, *SEMA5A* was also highly expressed in almost all human brain tissues at 0 to 5 months after birth, which also represents an important stage of human brain development (**Figure 3B**).

Furthermore, we investigated the spatiotemporal expression pattern of *SEMA5A* in the mouse embryo at embryonic day 14.5, on images of ISH from GenePaint. Consistent with the findings mentioned above, it showed that *SEMA5A* was strongly expressed in the neocortex of the embryonic mouse brain (**Figure 3C**). The expression profiles of *SEMA5A* in the brain indicates that this gene may play an essential role in the development of the brain and that dysfunctional *SEMA5A* may lead to human neurodevelopmental disorders such as EEs.

PPI Network and Gene Function Analyses

To gain insight into the biological function of *SEMA5A* and those genes that interact with *SEMA5A*, we extracted genes with interaction scores greater than 200 from the STRING database, and then constructed a PPI network. In the resulting PPI network, we found 196 nodes, among which 43 genes were known candidate genes in neurodevelopmental disorders. These 43 genes included 16 ASD candidate genes, six epilepsy candidate genes, six intellectual disability (ID) candidate genes, three candidate genes shared by

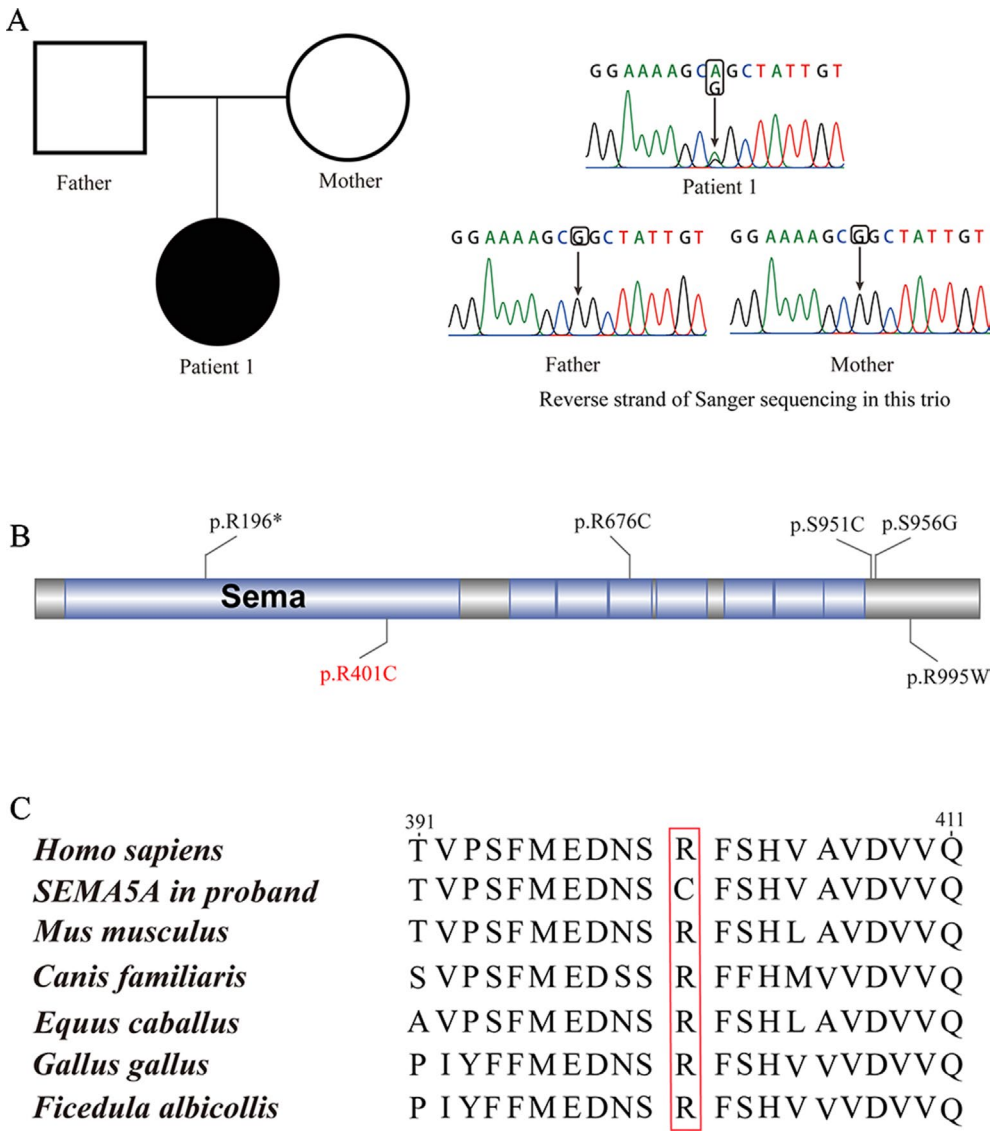

FIGURE 2 | *De novo* mutations (DNMs) in *SEMA5A*. **(A)** Sanger sequencing of patient 1. The filled symbol indicates the affected individual. **(B)** Protein schematic of SEMA5A. Red indicates the DNM identified in this study, and black indicates mutations detected in ASD. **(C)** The conservation of the DNM in *SEMA5A* among various vertebrates.

TABLE 1 | Mutations detected by WES in this study in patients with IS.

Trio	Chr	Gene	Mutation	Inheritance	Protein change	SIFT (score)	VEST3 (score)	Polyphen2 (score)	GERP++ (score)
Patient1	chr5	*SEMA5A*	missense	*De novo*	p.(Arg401Cys)	Damaging (0)	Damaging (0.944)	Probably damaging (0.998)	Conserved (5.75)
Patient 2	chr22	*CLTCL1*	missense	inherited	p.(Met1316Val)	Tolerable (0.14)	Damaging (0.634)	Probably damaging (0.994)	Conserved (3.5)
Patient2	chr22	*CLTCL1*	missense	inherited	p.(Arg1165Cys)	Damaging (0)	Damaging (0.765)	Probably damaging (0.961)	Conserved (3.16)
Patient3	chr5	*PLEKHG4B*	missense	*De novo*	p.(Ala914Thr)	Tolerable (0.51)	Tolerable (0.118)	Benign (0.364)	Non-conserved (0.11)

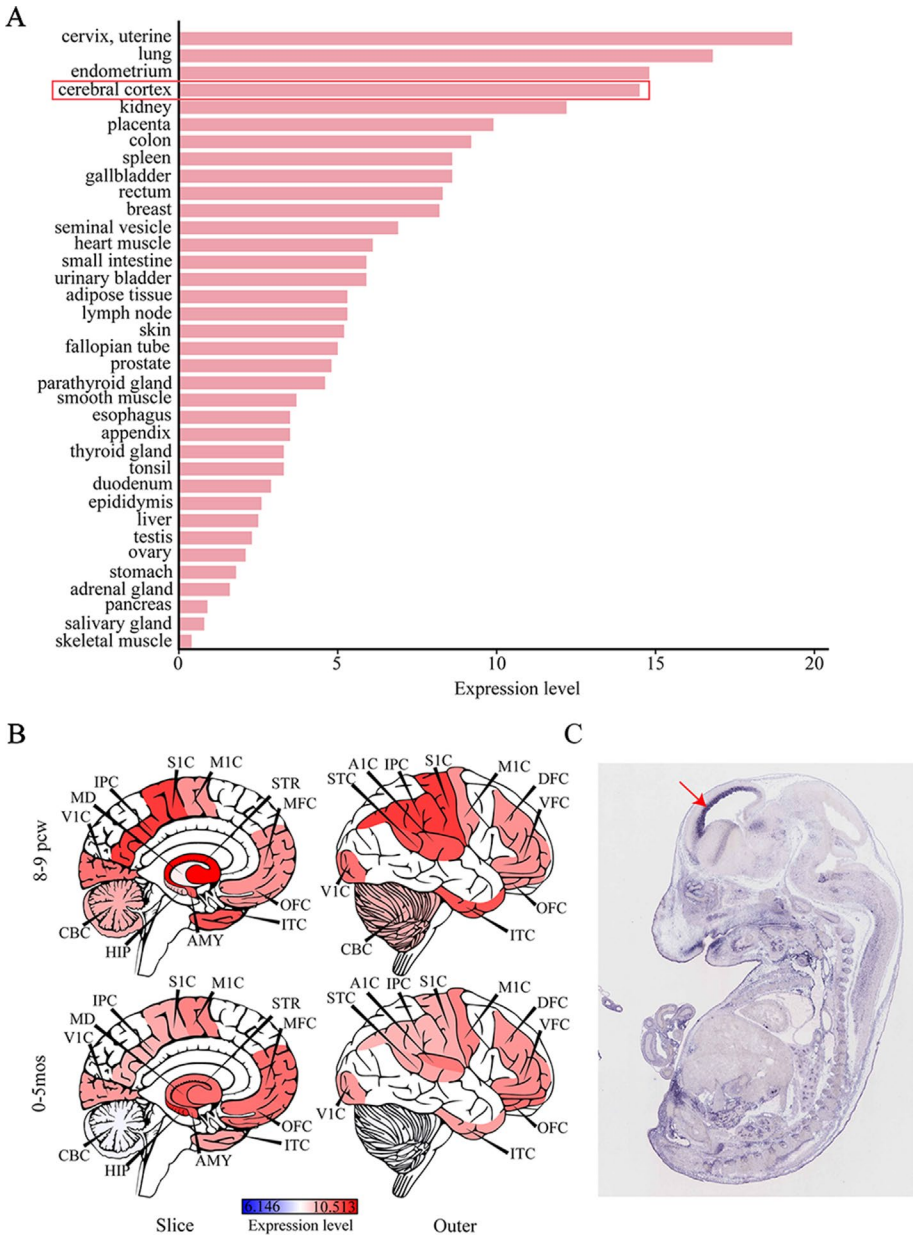

FIGURE 3 | The expression profile of *SEMA5A*. **(A)** The expression of *SEMA5A* in 36 tissues in the human protein atlas (HPA) database. **(B)** Spatiotemporal expression profile of *SEMA5A* in HBT. A1C, primary auditory cortex; AMY, amygdala; CBC, cerebellar cortex; DFC, dorsolateral prefrontal cortex; HIP, hippocampus; IPC, posterior inferior parietal cortex; ITC, inferior temporal cortex; M1C, primary motor cortex; MD, mediodorsal nucleus of the thalamus; MFC, medial prefrontal cortex; OFC, orbital prefrontal cortex; S1C, primary somatosensory cortex; STC, posterior superior temporal cortex; STR, striatum; V1C, primary visual cortex; VFC, ventrolateral prefrontal cortex. **(C)** The expression of *SEMA5A* in a mouse embryo at embryonic day 14.5 was analyzed by situ hybridization. The arrow indicates the neocortex.

ASD and epilepsy, nine shared by ASD and ID, one shared by ID and epilepsy and three shared by ASD, ID and epilepsy (**Figure 4A**). To evaluate the interconnectivity among these 43 genes, we performed an interconnected PPI analysis and observed a non-random interaction connectivity with 100,000 permutation tests ($q = 1 \times 10^{-5}$ for genes, $q = 1 \times 10^{-5}$ for connections). Moreover, we discovered that these genes were significantly enriched in PSD ($q = 7.50 \times 10^{-6}$) and genes under evolutionary constraint ($q = 2.04 \times 10^{-5}$) (**Supplementary Figure S3 and S4**).

Furthermore, when we performed a specific spatiotemporal expression analysis to further explore the expression patterns of the interacting genes, as expected, we observed that *SEMA5A* and its interacting genes were significantly enriched in the cortex, cerebellum, striatum and thalamus from early fetal to early infancy periods ($q < 0.05$) (**Figure 4B**). Next, a Gene Ontology (GO) enrichment analysis was performed to determine whether the genes in the PPI network were enriched in neurodevelopment-associated GO terms. The results showed that *SEMA5A* and its interacting

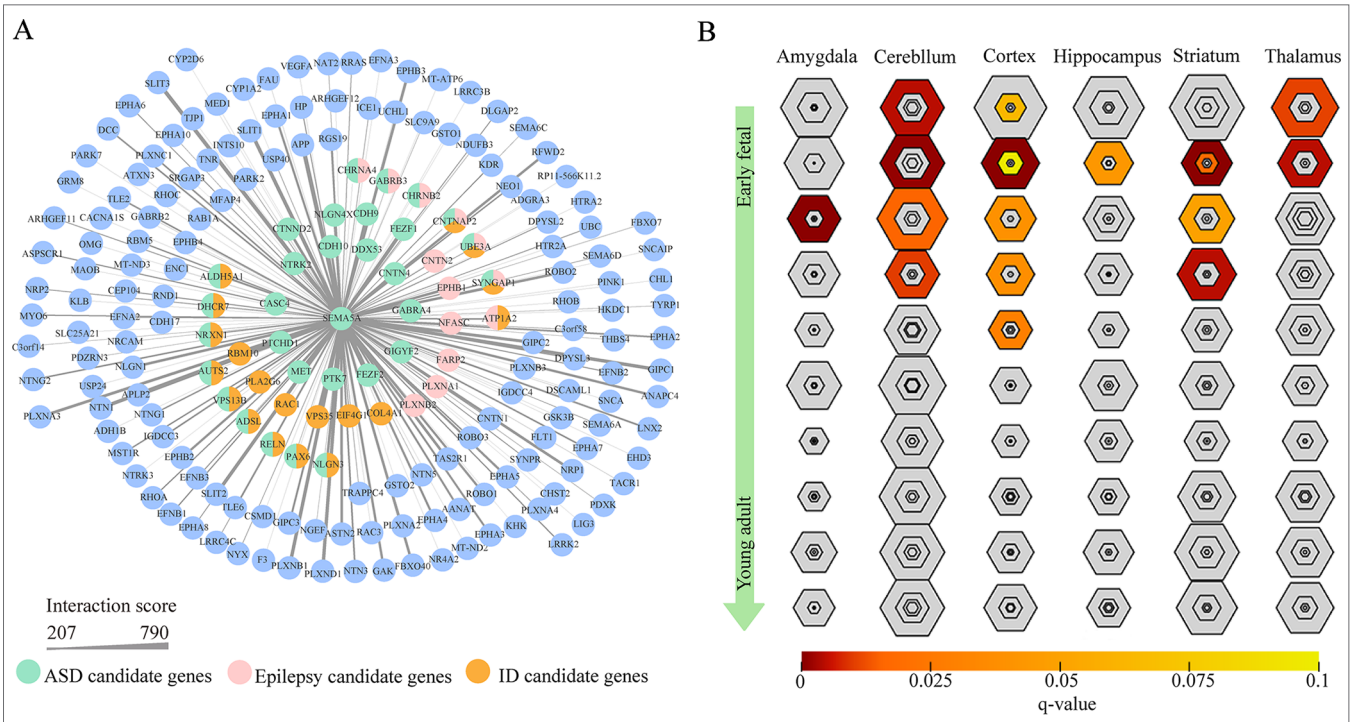

FIGURE 4 | Protein–protein interaction (PPI) network and specific expression analysis (SEA) enrichment analysis. **(A)** PPI network of genes found to interact with *SEMA5A* in the STRING database and to have interaction scores greater than 200. The nodes represent genes, and edges represent the intersection between gene pairs. The thickness of an edge denotes the interaction scores between gene pairs. **(B)** SEA of genes in the PPI network. Each hexagon shows different stringency thresholds, with colors indicating q values.

genes were significantly enriched in axon-, neuron- and synapse-associated GO terms, such as axon development ($q = 2.05 \times 10^{-49}$), axon guidance ($q = 5.22 \times 10^{-39}$), the regulation of neuron projection development ($q = 4.86 \times 10^{-31}$), and synapse organization ($q = 2.73 \times 10^{-19}$) (**Figure 5A**). Moreover, it is important to note that genes in the PPI network were relevant to ASD ($q = 1.88 \times 10^{-13}$), developmental disorders of mental health ($q = 2.48 \times 10^{-12}$), Parkinson's disease ($q = 1.61 \times 10^{-6}$) and epilepsy syndrome ($q = 8.92 \times 10^{-5}$) based on Disease Ontology (DO) analysis (**Figure 5B**).

DISCUSSION

In this study, we performed WES on three trios and presented evidence showing that a *de novo* missense mutation in *SEMA5A* was likely associated with IS. Considering the potential role of recessive inheritance in EEs, we also tried to detect homozygous inherited SNVs or indel mutations in other genes in our samples. However, there were no candidate damaging compound mutations or homozygous mutations in all patients except a compound heterozygous mutation in *CLTCL1* in patient 2. *SEMA5A* belongs to the semaphorin gene family, the members of which serve as canonical axon guidance proteins and function in pathological conditions of the nervous system (Yaron and Zheng, 2007). It is now clear that semaphorins and their receptors play many crucial roles in the development of neural circuits (Pasterkamp, 2012), including roles in neuronal migration (Hernandez-Miranda et al., 2011), axon bundling (Van Battum et al., 2015), axon pruning (Pasterkamp, 2012) and synaptic transmission (Sahay et al., 2005; Carrillo et al., 2010). Moreover, some alterations in the expression of axon guidance proteins, such as semaphorins and ephrins, have been observed in animal models of epilepsy (Barnes et al., 2003; Xia et al., 2013). It is generally accepted that semaphorins that can act as axon guidance proteins are involved in a variety of neurological diseases, including ASD (McFadden and Minshew, 2013), epilepsy (Xia et al., 2013), Parkinson's disease (Lin et al., 2009) and Alzheimer's disease (Good et al., 2004).

Moreover, semaphorins are sensitive to electrical activity and experience and form a large family consisting of eight classes, among which class 5 semaphorins (*SEMA5A* and *SEMA5B*) play essential roles in the functions of the developing nervous system (Mann et al., 2007). SEMA5A, as an integral membrane protein, acts as a bifunctional guidance cue that could be directly implicated in both attractive and inhibitory processes during axon development (Kantor et al., 2004). Furthermore, *Sema5A* was shown to be essential for the proper development of fasciculus retroflexus in a rat model (Kantor et al., 2004). A separate study showed that *Sema5A* functions as a negative regulator of synaptogenesis during early brain development and synaptic transmission (Duan et al., 2014).

Cumulative evidence obtained from various sources shows that *SEMA5A* is a transmembrane protein that has been identified as an autism susceptibility gene in humans according to a genome-wide association study (Weiss et al., 2009), cDNA microarray technology

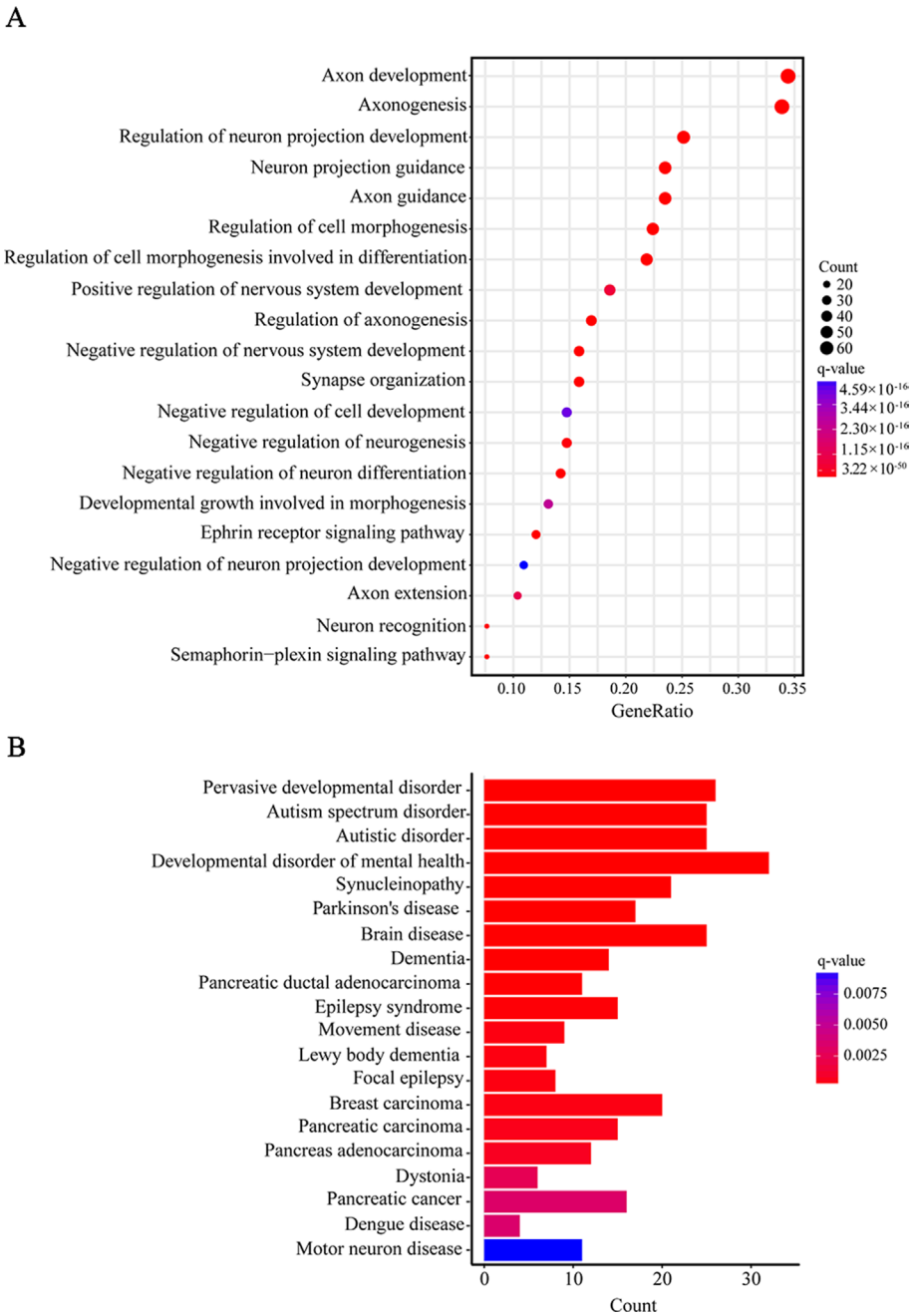

FIGURE 5 | Gene Ontology (GO) and Disease Ontology (DO) enrichment analyses. **(A)** The top 20 enriched GO terms of biological process. Circle size indicates the number of genes enriched in each term. Color saturation represents the significance level. **(B)** The top 20 terms in the DO enrichment analysis. The x-axis shows the number of genes enriched in each term. The q-value of each term is indicated by color according to the legend.

(Melin et al., 2006) and expression quantitative trait locus mapping (Cheng et al., 2013). Notably, a *de novo* microdeletion in *SEMA5A* was identified in a patient with ASD and ID (Mosca-Boidron et al., 2016). It is worth noting that ASD and EEs often occur together and share some common genetic etiologies (Srivastava and Sahin, 2017). Abnormal synaptic plasticity may represent a common pathophysiological mechanism in ASD and epilepsy that can lead to an imbalance between excitatory and inhibitory neurotransmission in the developing brain (Brooks-Kayal, 2010). *NRP2* is a receptor for both *SEMA3C* and *SEMA3F*, each of which plays an important role in axon guidance in the peripheral and central nervous systems (Chen et al., 2000) in addition to synaptic plasticity (Lee et al., 2012). Positional and functional evidence has been indicated that polymorphisms and mutations in *NRP2* may be associated with autism (Wu et al., 2007). For example, *NRP2* knockout mice were found to have a lowered seizure threshold

and were more sensitive to chemical challenges aimed at inducing epileptogenesis (Gant et al., 2009). Notably, several studies have demonstrated that *SEMA3C* and *SEMA3F*, which are homologous to *SEMA5A*, are candidate genes for EEs (Barnes et al., 2003; Mefford et al., 2011). The patient carrying damaging DNM in *SEMA5A* in this study was diagnosed with IS at the age of only 7 months, though there was no behaviors or phenotypes related to ASD based on clinical diagnosis, which may be due to the very young age of the patient. Therefore, it is necessary to continue examinations or follow-up surveys for this patient as they become older, to evaluate whether ASD develops.

CLTCL1, a clathrin heavy chain protein in humans, is associated with the neuromuscular system (Towler et al., 2004), and neuropeptide degradation and secretion during neuronal development (Nahorski et al., 2018). The homozygous R125C mutation in *CLTCL1*, inherited from heterozygous parents, was predicted to be damaging and has been identified in patients with autism (Chahrour et al., 2012). A mutation in *CLTCL1* located on chromosome 22q11.2 has also been associated with susceptibility to schizophrenia (Karayiorgou et al., 2010; Chahrour et al., 2012). DiGeorge syndrome is usually associated with the deletion of chromosome 22q11.2, which is linked with cognitive impairments, susceptibility to schizophrenia and neuroanatomical changes (Karayiorgou et al., 2010). It is worth noting that an interruption in the *CLTCL* gene observed in a patient with DiGeorge syndrome was found to contribute to the patient's phenotype, including a seizure disorder, ID and facial dysmorphia (Holmes et al., 1997). However, whether the compound mutation in *CLTCL* identified in this study was associated with IS still required further genetic evidence or functional studies to support it.

CONCLUSIONS

To our knowledge, this is a novel DNM of *SEMA5A* found in an individual with IS. Our results provide genetic and functional evidence showing that *SEMA5A* plays a significant role in the development of the brain and might be involved in several neurodevelopmental disorders, such as EE and ASD. The discovery of damaging mutations in *SEMA5A* provides further evidence supporting the role of axon guidance proteins in the pathogenesis and causes underlying EE. However, to definitively establish the role of *SEMA5A* in IS, further genetic function studies are needed.

ETHICS STATEMENT

This study was carried out with the approval provided by the Second Affiliated Hospital and Yuying Children's Hospital of Wenzhou Medical University. Written informed consent of all participants was collected from their parents or guardians at the time of recruitment for participation in the study and publication of the study.

AUTHOR CONTRIBUTIONS

QW and ZWL contributed to the drafting and revision of the manuscript, data acquisition, and analysis. ZDL, RZ, YL, and FL contributed to individual recruitment, data acquisition, and data analysis. XX, MT, and WS contributed to data acquisition and manuscript revision. XZ, JZ, and YL contributed to study concept and design, critical review, and manuscript revision.

FUNDING

This study was supported by the Zhejiang Provincial Natural Science Foundation of China (Grant No. LY18C060007).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2019.00605/full#supplementary-material

REFERENCES

Barnes, G., Puranam, R. S., Luo, Y., and McNamara, J. O. (2003). Temporal specific patterns of semaphorin gene expression in rat brain after kainic acid-induced status epilepticus. *Hippocampus* 13, 1–20. doi: 10.1002/hipo.10041

Berg, A. T., Berkovic, S. F., Brodie, M. J., Buchhalter, J., Cross, J. H., van Emde Boas, W., et al. (2010). Revised terminology and concepts for organization of seizures and epilepsies: report of the ILAE Commission on Classification and Terminology, 2005-2009. *Epilepsia* 51, 676–685. doi: 10.1111/j.1528-1167.2010.02522.x

Brooks-Kayal, A. (2010). Epilepsy and autism spectrum disorders: are there common developmental mechanisms? *Brain Dev.* 32, 731–738. doi: 10.1016/j.braindev.2010.04.010

Carrillo, R. A., Olsen, D. P., Yoon, K. S., and Keshishian, H. (2010). Presynaptic activity and CaMKII modulate retrograde semaphorin signaling and synaptic refinement. *Neuron* 68, 32–44. doi: 10.1016/j.neuron.2010.09.005

Carvill, G. L., Heavin, S. B., Yendle, S. C., McMahon, J. M., O'Roak, B. J., Cook, J., et al. (2013). Targeted resequencing in epileptic encephalopathies identifies de novo mutations in CHD2 and SYNGAP1. *Nat. Genet.* 45, 825–830. doi: 10.1038/ng.2646

Chahrour, M. H., Yu, T. W., Lim, E. T., Ataman, B., Coulter, M. E., Hill, R. S., et al. (2012). Whole-exome sequencing and homozygosity analysis implicate depolarization-regulated neuronal genes in autism. *PLoS Genet.* 8, e1002635. doi: 10.1371/journal.pgen.1002635

Chen, H., Bagri, A., Zupicich, J. A., Zou, Y., Stoeckli, E., Pleasure, S. J., et al. (2000). Neuropilin-2 regulates the development of selective cranial and sensory nerves and hippocampal mossy fiber projections. *Neuron* 25, 43–56. doi: 10.1016/S0896-6273(00)80870-3

Cheng, Y., Quinn, J. F., and Weiss, L. A. (2013). An eQTL mapping approach reveals that rare variants in the SEMA5A regulatory network impact autism risk. *Hum. Mol. Genet.* 22, 2960–2972. doi: 10.1093/hmg/ddt150

Claes, L., Del-Favero, J., Ceulemans, B., Lagae, L., Van Broeckhoven, C., and De Jonghe, P. (2001). De novo mutations in the sodium-channel gene SCN1A cause severe myoclonic epilepsy of infancy. *Am. J. Hum. Genet.* 68, 1327–1332. doi: 10.1086/320609

Collins, M. O., Husi, H., Yu, L., Brandon, J. M., Anderson, C. N., Blackstock, W. P., et al. (2006). Molecular characterization and comparison of the components and multiprotein complexes in the postsynaptic proteome. *J. Neurochem.* 97 Suppl 1, 16–23. doi: 10.1111/j.1471-4159.2005.03507.x

D'Gama, A. M., Pochareddy, S., Li, M., Jamuar, S. S., Reiff, R. E., Lam, A. N., et al. (2015). Targeted DNA sequencing from autism spectrum disorder brains implicates multiple genetic mechanisms. *Neuron* 88, 910–917. doi: 10.1016/j.neuron.2015.11.009

Dimassi, S., Labalme, A., Ville, D., Calender, A., Mignot, C., Boutry-Kryza, N., et al. (2016). Whole-exome sequencing improves the diagnosis yield in sporadic infantile spasm syndrome. *Clin. Genet.* 89, 198–204. doi: 10.1111/cge.12636

Duan, Y., Wang, S. H., Song, J., Mironova, Y., Ming, G. L., Kolodkin, A. L., et al. (2014). Semaphorin 5A inhibits synaptogenesis in early postnatal- and adult-born hippocampal dentate granule cells. *eLife* 3. doi: 10.7554/eLife.04390

Epi4K Consortium, and Epilepsy Phenome/Genome Project (2013). De novo mutations in epileptic encephalopathies. *Nature* 501, 217–221. doi: 10.1038/nature12439

Fukata, Y., and Fukata, M. (2017). Epilepsy and synaptic proteins. *Curr. Opin. Neurobiol.* 45, 1–8. doi: 10.1016/j.conb.2017.02.001

Gant, J. C., Thibault, O., Blalock, E. M., Yang, J., Bachstetter, A., Kotick, J., et al. (2009). Decreased number of interneurons and increased seizures in neuropilin 2 deficient mice: implications for autism and epilepsy. *Epilepsia* 50, 629–645. doi: 10.1111/j.1528-1167.2008.01725.x

Good, P. F., Alapat, D., Hsu, A., Chu, C., Perl, D., Wen, X., et al. (2004). A role for semaphorin 3A signaling in the degeneration of hippocampal neurons during Alzheimer's disease. *J. Neurochem.* 91, 716–736. doi: 10.1111/j.1471-4159.2004.02766.x

He, N., Lin, Z. J., Wang, J., Wei, F., Meng, H., Liu, X. R., et al. (2019). Evaluating the pathogenic potential of genes with de novo variants in epileptic encephalopathies. *Genet. Med.* 21, 17–27. doi: 10.1038/s41436-018-0011-y

Hernandez-Miranda, L. R., Cariboni, A., Faux, C., Ruhrberg, C., Cho, J. H., Cloutier, J. F., et al. (2011). Robo1 regulates semaphorin signaling to guide the migration of cortical interneurons through the ventral forebrain. *J. Neurosci.* 31, 6174–6187. doi: 10.1523/JNEUROSCI.5464-10.2011

Hilario, J. D., Rodino-Klapac, L. R., Wang, C., and Beattie, C. E. (2009). Semaphorin 5A is a bifunctional axon guidance cue for axial motoneurons in vivo. *Dev. Biol.* 326, 190–200. doi: 10.1016/j.ydbio.2008.11.007

Holmes, S. E., Riazi, M. A., Gong, W., McDermid, H. E., Sellinger, B. T., Hua, A., et al. (1997). Disruption of the clathrin heavy chain-like gene (CLTCL) associated with features of DGS/VCFS: a balanced (21;22)(p12;q11) translocation. *Hum. Mol. Genet.* 6, 357–367. doi: 10.1093/hmg/6.3.357

Iossifov, I., O'Roak, B. J., Sanders, S. J., Ronemus, M., Krumm, N., Levy, D., et al. (2014). The contribution of de novo coding mutations to autism spectrum disorder. *Nature* 515, 216–221. doi: 10.1038/nature13908

Jin, Z. B., Li, Z., Liu, Z., Jiang, Y., Cai, X. B., and Wu, J. (2018). Identification of de novo germline mutations and causal genes for sporadic diseases using trio-based whole-exome/genome sequencing. *Biol. Rev. Camb. Philos. Soc.* 93, 1014–1031. doi: 10.1111/brv.12383

Kantor, D. B., Chivatakarn, O., Peer, K. L., Oster, S. F., Inatani, M., Hansen, M. J., et al. (2004). Semaphorin 5A is a bifunctional axon guidance cue regulated by heparan and chondroitin sulfate proteoglycans. *Neuron* 44, 961–975. doi: 10.1016/j.neuron.2004.12.002

Karayiorgou, M., Simon, T. J., and Gogos, J. A. (2010). 22q11.2 microdeletions: linking DNA structural variation to brain dysfunction and schizophrenia. *Nat. Rev. Neurosci.* 11, 402–416. doi: 10.1038/nrn2841

Lau, C. G., and Zukin, R. S. (2007). NMDA receptor trafficking in synaptic plasticity and neuropsychiatric disorders. *Nat. Rev. Neurosci.* 8, 413–426. doi: 10.1038/nrn2153

Lee, K., Kim, J. H., Kwon, O. B., An, K., Ryu, J., Cho, K., et al. (2012). An activity-regulated microRNA, miR-188, controls dendritic plasticity and synaptic transmission by downregulating neuropilin-2. *J. Neurosci.* 32, 5678–5687. doi: 10.1523/JNEUROSCI.6471-11.2012

Lek, M., Karczewski, K. J., Minikel, E. V., Samocha, K. E., Banks, E., Fennell, T., et al. (2016). Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 536, 285–291. doi: 10.1038/nature19057

Lemke, J. R., Hendrickx, R., Geider, K., Laube, B., Schwake, M., Harvey, R. J., et al. (2014). GRIN2B mutations in West syndrome and intellectual disability with focal epilepsy. *Ann. Neurol.* 75, 147–154. doi: 10.1002/ana.24073

Li, J., Jiang, Y., Wang, T., Chen, H., Xie, Q., Shao, Q., et al. (2015). mirTrios: an integrated pipeline for detection of de novo and rare inherited mutations from trios-based next-generation sequencing. *J. Med. Genet.* 52, 275–281. doi: 10.1136/jmedgenet-2014-102656

Lin, L., Lesnick, T. G., Maraganore, D. M., and Isacson, O. (2009). Axon guidance and synaptic maintenance: preclinical markers for neurodegenerative disease and therapeutics. *Trends Neurosci.* 32, 142–149. doi: 10.1016/j.tins.2008.11.006

Liu, Q., Chen, C., Shen, E., Zhao, F., Sun, Z., and Wu, J. (2012). Detection, annotation and visualization of alternative splicing from RNA-Seq data with SplicingViewer. *Genomics* 99, 178–182. doi: 10.1016/j.ygeno.2011.12.003

Mann, F., Chauvet, S., and Rougon, G. (2007). Semaphorins in development and adult brain: implication for neurological diseases. *Prog. Neurobiol.* 82, 57–79. doi: 10.1016/j.pneurobio.2007.02.011

McFadden, K., and Minshew, N. J. (2013). Evidence for dysregulation of axonal growth and guidance in the etiology of ASD. *Front. Hum. Neurosci.* 7, 671. doi: 10.3389/fnhum.2013.00671

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., et al. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20, 1297–1303. doi: 10.1101/gr.107524.110

McTague, A., Howell, K. B., Cross, J. H., Kurian, M. A., and Scheffer, I. E. (2016). The genetic landscape of the epileptic encephalopathies of infancy and childhood. *Lancet Neurol.* 15, 304–316. doi: 10.1016/S1474-4422(15)00250-1

Mefford, H. C., Yendle, S. C., Hsu, C., Cook, J., Geraghty, E., McMahon, J. M., et al. (2011). Rare copy number variants are an important cause of epileptic encephalopathies. *Ann. Neurol.* 70, 974–985. doi: 10.1002/ana.22645

Melin, M., Carlsson, B., Anckarsater, H., Rastam, M., Betancur, C., Isaksson, A., et al. (2006). Constitutional downregulation of SEMA5A expression in autism. *Neuropsychobiology* 54, 64–69. doi: 10.1159/000096040

Michaelson, J. J., Shi, Y., Gujral, M., Zheng, H., Malhotra, D., Jin, X., et al. (2012). Whole-genome sequencing in autism identifies hot spots for de novo germline mutation. *Cell* 151, 1431–1442. doi: 10.1016/j.cell.2012.11.019

Mosca-Boidron, A. L., Gueneau, L., Huguet, G., Goldenberg, A., Henry, C., Gigot, N., et al. (2016). A de novo microdeletion of SEMA5A in a boy with autism spectrum disorder and intellectual disability. *Eur. J. Hum. Genet.* 24, 838–843. doi: 10.1038/ejhg.2015.211

Nabbout, R., and Dulac, O. (2003). Epileptic encephalopathies: a brief overview. *J. Clin. Neurophysiol.* 20, 393–397. doi: 10.1097/00004691-200311000-00002

Nahorski, M. S., Borner, G. H. H., Shaikh, S. S., Davies, A. K., Al-Gazali, L., Antrobus, R., et al. (2018). Clathrin heavy chain 22 contributes to the control of neuropeptide degradation and secretion during neuronal development. *Sci. Rep.* 8, 2340. doi: 10.1038/s41598-018-19980-0

Pardo, C. A., Nabbout, R., and Galanopoulou, A. S. (2014). Mechanisms of epileptogenesis in pediatric epileptic syndromes: Rasmussen encephalitis, infantile spasms, and febrile infection-related epilepsy syndrome (FIRES). *Neurotherapeutics* 11, 297–310. doi: 10.1007/s13311-014-0265-2

Pasterkamp, R. J. (2012). Getting neural circuits into shape with semaphorins. *Nat. Rev. Neurosci.* 13, 605–618. doi: 10.1038/nrn3302

Pasterkamp, R. J., and Giger, R. J. (2009). Semaphorin function in neural plasticity and disease. *Curr. Opin. Neurobiol.* 19, 263–274. doi: 10.1016/j.conb.2009.06.001

Pavone, P., Striano, P., Falsaperla, R., Pavone, L., and Ruggieri, M. (2014). Infantile spasms syndrome, West syndrome and related phenotypes: what we know in 2013. *Brain Dev.* 36, 739–751. doi: 10.1016/j.braindev.2013.10.008

Penzes, P., Cahill, M. E., Jones, K. A., VanLeeuwen, J. E., and Woolfrey, K. M. (2011). Dendritic spine pathology in neuropsychiatric disorders. *Nat. Neurosci.* 14, 285–293. doi: 10.1038/nn.2741

Ran, X., Li, J., Shao, Q., Chen, H., Lin, Z., Sun, Z. S., et al. (2015). EpilepsyGene: a genetic resource for genes and mutations related to epilepsy. *Nucleic Acids Res.* 43, D893–D899. doi: 10.1093/nar/gku943

Sahay, A., Kim, C. H., Sepkuty, J. P., Cho, E., Huganir, R. L., Ginty, D. D., et al. (2005). Secreted semaphorins modulate synaptic transmission in the adult hippocampus. *J. Neurosci.* 25, 3613–3620. doi: 10.1523/JNEUROSCI.5255-04.2005

Samocha, K. E., Robinson, E. B., Sanders, S. J., Stevens, C., Sabo, A., McGrath, L. M., et al. (2014). A framework for the interpretation of de novo mutation in human disease. *Nat. Genet.* 46, 944–950. doi: 10.1038/ng.3050

Scheffer, I. E., Berkovic, S., Capovilla, G., Connolly, M. B., French, J., Guilhoto, L., et al. (2017). ILAE classification of the epilepsies: position paper of the ILAE Commission for Classification and Terminology. *Epilepsia* 58, 512–521. doi: 10.1111/epi.13709

Shbarou, R., and Mikati, M. A. (2016). The expanding clinical spectrum of genetic pediatric epileptic encephalopathies. *Semin. Pediatr. Neurol.* 23, 134–142. doi: 10.1016/j.spen.2016.06.002

Shen, K., and Cowan, C. W. (2010). Guidance molecules in synapse formation and plasticity. *Cold Spring Harb. Perspect. Biol.* 2, a001842. doi: 10.1101/cshperspect.a001842

Song, I., and Huganir, R. L. (2002). Regulation of AMPA receptors during synaptic plasticity. *Trends Neurosci.* 25, 578–588. doi: 10.1016/S0166-2236(02)02270-1

Srivastava, S., and Sahin, M. (2017). Autism spectrum disorder and epileptic encephalopathy: common causes, many questions. *J. Neurodev. Disord.* 9, 23. doi: 10.1186/s11689-017-9202-0

Towler, M. C., Gleeson, P. A., Hoshino, S., Rahkila, P., Manalo, V., Ohkoshi, N., et al. (2004). Clathrin isoform CHC22, a component of neuromuscular and myotendinous junctions, binds sorting nexin 5 and has increased expression during myogenesis and muscle regeneration. *Mol. Biol. Cell* 15, 3181–3195. doi: 10.1091/mbc.e04-03-0249

Van Battum, E. Y., Brignani, S., and Pasterkamp, R. J. (2015). Axon guidance proteins in neurological disorders. *Lancet Neurol.* 14, 532–546. doi: 10.1016/S1474-4422(14)70257-1

Veltman, J. A., and Brunner, H. G. (2012). De novo mutations in human genetic disease. *Nat. Rev. Genet.* 13, 565–575. doi: 10.1038/nrg3241

Wang, T., Liu, Q., Li, X., Wang, X., Li, J., Zhu, X., et al. (2013). RRBS-analyser: a comprehensive web server for reduced representation bisulfite sequencing data analysis. *Hum. Mutat.* 34, 1606–1610. doi: 10.1002/humu.22444

Weiss, L. A., Arking, D. E., and The Gene Discovery Project of Johns Hopkins & the Autism Consortium. (2009). A genome-wide linkage and association scan reveals novel loci for autism. *Nature* 461, 802–808. doi: 10.1038/nature08490

Wu, S., Yue, W., Jia, M., Ruan, Y., Lu, T., Gong, X., et al. (2007). Association of the neuropilin-2 (NRP2) gene polymorphisms with autism in Chinese Han population. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* 144B, 492–495. doi: 10.1002/ajmg.b.30495

Xia, Y., Luo, C., Dai, S., and Yao, D. (2013). Increased EphA/ephrinA expression in hippocampus of pilocarpine treated mouse. *Epilepsy Res.* 105, 20–29. doi: 10.1016/j.eplepsyres.2013.01.001

Yaron, A., and Zheng, B. (2007). Navigating their way to the clinic: emerging roles for axon guidance molecules in neurological disorders and injury. *Dev. Neurobiol.* 67, 1216–1231. doi: 10.1002/dneu.20512

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Genetics

ORIGINAL RESEARCH
published: 16 August 2019
doi: 10.3389/fgene.2019.00761

Prenatal Diagnosis of Fetuses With Increased Nuchal Translucency by Genome Sequencing Analysis

Kwong Wai Choy[1,2,3†], Huilin Wang[4†], Mengmeng Shi[1†], Jingsi Chen[5†], Zhenjun Yang[1], Rui Zhang[4], Huanchen Yan[5], Yanfang Wang[4], Shaoyun Chen[4], Matthew Hoi Kin Chau[1], Ye Cao[1,6], Olivia Y.M. Chan[1], Yvonne K. Kwok[1], Yuanfang Zhu[4], Min Chen[5], Tak Yeung Leung[1,2,3] and Zirui Dong[1,2,5]*

[1] Department of Obstetrics & Gynaecology, The Chinese University of Hong Kong, Hong Kong, China, [2] Shenzhen Research Institute, The Chinese University of Hong Kong, Shenzhen, China, [3] The Chinese University of Hong Kong-Baylor College of Medicine Joint Center for Medical Genetics, Hong Kong, China, [4] Department of Central Laboratory, Bao'an Maternity and Child Healthcare Hospital Affiliated to Jinan University School of Medicine, Key Laboratory of Birth Defects Research, Birth Defects Prevention Research and Transformation Team, Shenzhen, China, [5] Department of Obstetrics and Gynecology, Key Laboratory of Reproduction and Genetics of Guangdong Higher Education Institutes, The Third Affiliated Hospital of Guangzhou Medical University, Guangzhou, China, [6] Department of Molecular and Cellular Biology, Baylor College of Medicine, Houston, TX, United States

OPEN ACCESS

Edited by:
Evica Rajcan-Separovic, University of British Columbia, Canada

Reviewed by:
Pengfei Liu, Baylor College of Medicine, United States
Neeta L. Vora, University of North Carolina at Chapel Hill, United States
Abdul Noor, University of Toronto, Canada

***Correspondence:**
Zirui Dong
elvisdong@cuhk.edu.hk

†These authors have contributed equally to this work.

Specialty section:
This article was submitted to Genetic Disorders, a section of the journal Frontiers in Genetics

Received: *16 April 2019*
Accepted: *17 July 2019*
Published: *16 August 2019*

Citation:
Choy KW, Wang H, Shi M, Chen J, Yang Z, Zhang R, Yan H, Wang Y, Chen S, Chau MHK, Cao Y, Chan OYM, Kwok YK, Zhu Y, Chen M, Leung TY and Dong Z (2019) Prenatal Diagnosis of Fetuses With Increased Nuchal Translucency by Genome Sequencing Analysis. Front. Genet. 10:761. doi: 10.3389/fgene.2019.00761

Background: Increased nuchal translucency (NT) is an important biomarker associated with increased risk of fetal structural anomalies. It is known to be contributed by a wide range of genetic etiologies from single-nucleotide variants to those affecting millions of base pairs. Currently, prenatal diagnosis is routinely performed by karyotyping and chromosomal microarray analysis (CMA); however, both of them have limited resolution. The diversity of the genetic etiologies warrants an integrated assay such as genome sequencing (GS) for comprehensive detection of genomic variants. Herein, we aim to evaluate the feasibility of applying GS in prenatal diagnosis for the fetuses with increased NT.

Methods: We retrospectively applied GS (> 30-fold) for fetuses with increased NT (≥3.5 mm) who underwent routine prenatal diagnosis. Detection of single-nucleotide variants, copy number variants, and structural rearrangements was performed simultaneously, and the results were integrated for interpretation in accordance with the guidelines of the American College of Medical Genetics and Genomics. Pathogenic or likely pathogenic (P/LP) variants were selected for validation and parental confirmation, when available.

Results: Overall, 50 fetuses were enrolled, including 34 cases with isolated increased NT and 16 cases with other fetal structural malformations. Routine CMA and karyotyping reported eight P/LP CNVs, yielding a diagnostic rate of 16.0% (8/50). In comparison, GS provided a twofold increase in diagnostic yield (32.0%, 16/50), including one mosaic turner syndrome, eight cases with microdeletions/microduplications, and seven cases with P/LP point mutations. Moreover, GS identified two cryptic insertions and two inversions. Follow-up study further demonstrated the potential pathogenicity of an apparently balanced insertion that disrupted an OMIM autosomal dominant disease-causing gene at the insertion site.

Conclusions: Our study demonstrates that applying GS in fetuses with increased NT can comprehensively detect and delineate the various genomic variants that are causative to the diseases. Importantly, prenatal diagnosis by GS doubled the diagnostic yield compared with routine protocols. Given a comparable turnaround time and less DNA required, our study provides strong evidence to facilitate GS in prenatal diagnosis, particularly in fetuses with increased NT.

Keywords: increased nuchal translucency, genome sequencing, prenatal diagnosis, genomic variants, structural rearrangement

INTRODUCTION

Detection of fetuses with increased nuchal translucency (NT) in routine first-trimester ultrasound screening has been widely used as a sensitive indication for fetal chromosomal abnormalities and/or fetal structural anomalies, such as congenital heart disorders or neurodevelopmental anomalies detected in later gestations (Leung et al., 2011; Huang et al., 2014; Socolov et al., 2017; Sinajon et al., 2019). Fetuses with increased NT and structural malformations are frequently contributed by genetic abnormalities and have poor prognoses. However, more than 80% of such cases do not obtain a causative result with the current routine prenatal diagnostic tests (Leung et al., 2011; Huang et al., 2014; Yang et al., 2017; Sinajon et al., 2019), challenging genetic counseling and clinical management. In addition, pathogenic copy number variants (CNVs) only account for 0.8% to 5.3% of these fetuses with isolated increased NT (with/without other soft markers) (Leung et al., 2011; Huang et al., 2014), and part of these cases would also have poor outcomes. Therefore, a test for the comprehensive detection of disease associated genomic variants including numerical disorders, structural rearrangements, CNVs, and point mutations in this prenatal cohort is warranted.

In prenatal diagnosis, quantitative fluorescent PCR (QF-PCR) is routinely conducted for the detection of maternal cell contamination (MCC) and common aneuploidy [such as Trisomy 21 (Choy et al., 2014; Sinajon et al., 2019)]. In addition, since 2010, chromosomal microarray analysis (CMA) has been recommended as the first-tier test for high-risk pregnancies in identification of microscopic or submicroscopic CNVs (Leung et al., 2011; Huang et al., 2014; Chau et al., 2019; Sinajon et al., 2019). However, this approach is limited by its resolution and it cannot detect single-nucleotide variants (SNVs) and small insertions/deletions (InDels). Owing to the breakthrough of molecular technologies such as next-generation sequencing and its reduction of costs over the years, whole-exome sequencing (or exome sequencing, WES) has been applied for both research purposes and clinical use (Drury et al., 2015; Fu et al., 2018; Leung et al., 2018; Normand et al., 2018; Lord et al., 2019; Petrovski et al., 2019). Emerging studies show that WES has the ability to provide genetic diagnoses ranging from 9.1% to 32% for the fetuses with a structural anomaly (Drury et al., 2015; Fu et al., 2018; Leung et al., 2018; Normand et al., 2018; Lord et al., 2019; Petrovski et al., 2019), while among these cases, WES yielded diagnoses in 3.2% to 21% of the fetuses with increased NT with/without structural malformations (Drury et al., 2015; Lord et al., 2019; Petrovski et al., 2019). However, most of these studies were conducted on prenatal cohorts after the exclusion of abnormal karyotypes and/or CMA results attributed to the cost and the limited ability of WES in CNV detection (Belkadi et al., 2015). These studies show the clinical utility of WES and CMA in prenatal diagnosis and warrants a combination of these two approaches for each case. Meanwhile, both WES and CMA are unable to detect apparently balanced structural rearrangements (or structural variants, SVs), a common limitation of the current methods, but some of these rearrangements have been demonstrated to be disease-causing (Talkowski et al., 2012a). The wide spectrum of genetic etiologies in fetuses with increased NT ranging from single-base mutations to those affecting millions of base pairs and numerical disorders warrants a holistic approach for comprehensive detection of the disease-causing genetic variants.

Our previous studies have demonstrated the feasibility and potential diagnostic utility of applying low-pass whole-genome sequencing (or genome sequencing, GS) analysis for the detection of CNVs (Dong et al., 2016; Dong et al., 2017) and chromosomal structural rearrangements (Dong et al., 2014) including balanced translocations and inversions in both clinical cohorts and presumably normal populations in the 1000 Genomes Project (Dong et al., 2018a; Dong et al., 2018b). By increasing the read depth to a minimal of 30-fold for the purpose of including SNV/InDel detection, GS is able to provide comprehensive detection of various genomic variants, thus providing a unique platform for gene discovery and potential clinical application. However, evaluation of its clinical utility is warranted.

Herein, we aimed to apply GS for the investigation of genetic contributions to fetuses with increased NT with/without structural malformations and to evaluate the possibility for its potential clinical application.

MATERIALS AND METHODS

Ethics, Consent, and Permissions

The study protocol was approved by the Ethics Committee of the Joint Chinese University of Hong Kong–New Territories East Cluster Clinical Research Ethics Committee (CREC Ref. No. 2016.713), Jinan University and Guangzhou Medical University. From year 2014 to 2018, 50 pregnant women, whose fetus was diagnosed with increased NT (≥3.5 mm) with/without structural

malformations (Leung et al., 2011; Huang et al., 2014) and had undergone prenatal diagnosis by CMA (and karyotyping if available) after a negative finding from QF-PCR (Choy et al., 2014), were recruited in this study. The recruitment criteria included fetuses with increased NT detected, and CMA results were available without the selection of 1) whether any fetal malformation was detected, 2) the results from CMA and/or karyotyping, 3) the timing for sample collection (CVS or AF), and 4) the pregnancy outcome (such as terminated pregnancies or live births) (**Figure 1**). Written informed consent was obtained from each participant for the purpose of this study, and any findings from the genome sequencing would not be disclosed to the patients. Routine CMA results were available in all cases and 90.0% (45/50) of them also had G-banded chromosome analysis results. Among them, chorionic villus sampling (CVS) was conducted for 37 cases at the time of first-trimester ultrasound screening, while amniotic fluid (AF) was obtained in the other 13 cases during the second trimester. Parental peripheral blood samples were collected at the time of prenatal sample retrieval if available.

DNA Preparation

Genomic DNA from chorionic villi or amniotic fluids (AF) were extracted using the DNeasy Blood & Tissue Kit (Cat No./ID: 69506, Qiagen, Hilden, Germany) at the time of routine CMA testing. DNA was quantified with the Quant-iT dsDNA HS Assay kit (Invitrogen, Carlsbad, CA), and DNA integrity was assessed by agarose gel electrophoresis and subsequently subjected to QF-PCR.

CMA and Karyotyping for Routine Prenatal Diagnosis

Two CMA platforms were routinely used in our three prenatal genetic diagnosis centers: CytoScan 750K (Applied Biosystems,

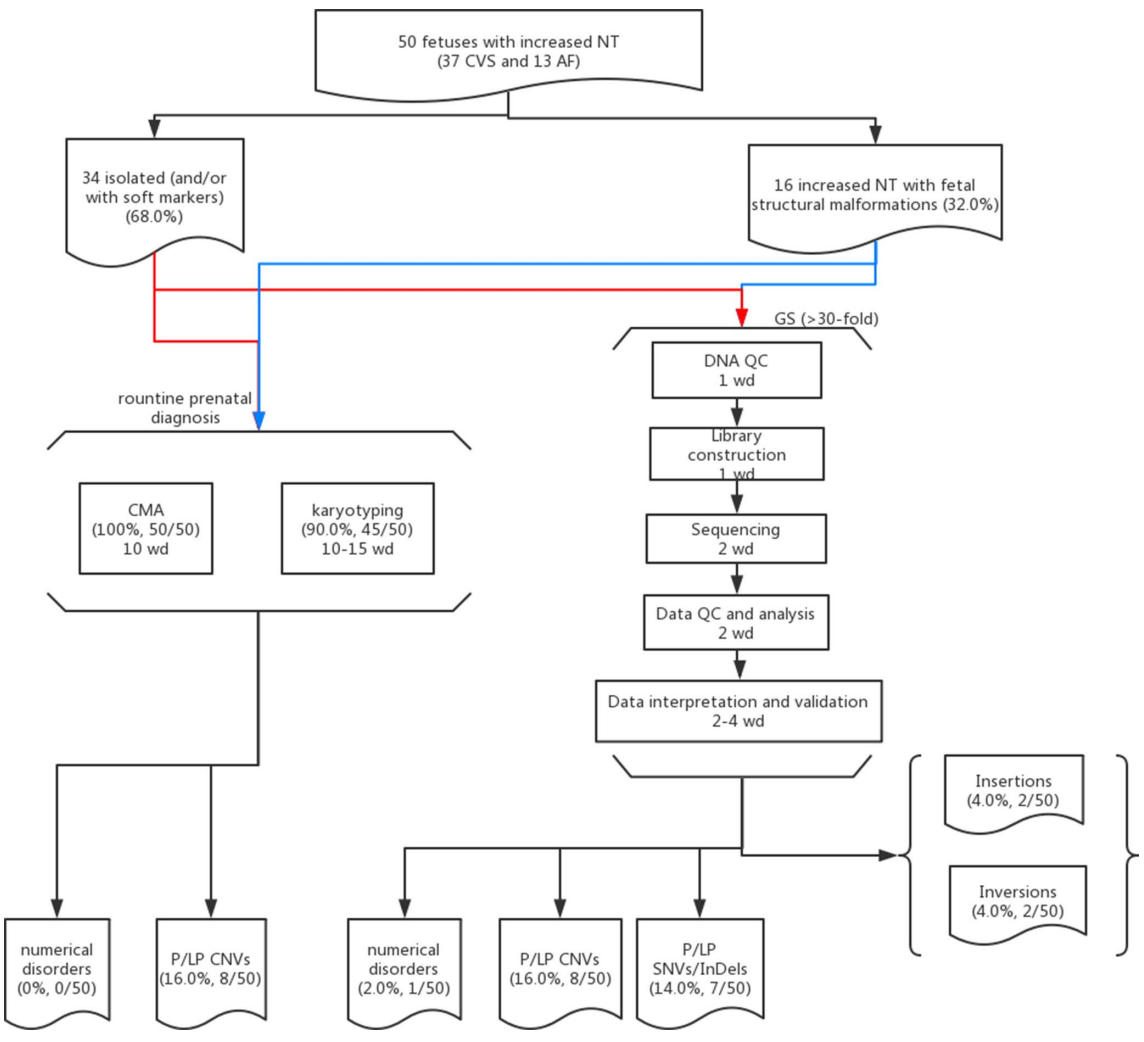

FIGURE 1 | Flowchart of this study. Detailed methods and results are described in the main text. Fifty cases with isolated (red lines) or syndromic increased NT (blue lines) and prenatal diagnosis results available (CMA and karyotyping) were selected for GS. The detailed procedures with the estimated turnaround time expressed as working day (wd) are provided in each box.

Affymetrix, Inc., Santa Clara, CA) and a well-established 8×60K fetal DNA chip (Agilent Technologies, Santa Clara, CA).

For the CytoScan 750K SNP-based platform, 250 ng of DNA samples was required and the experiment was conducted based on the manufacturer's protocol (Lin et al., 2015). Aberration analysis was performed with the ChAS 2.0 software (Affymetrix) (Lin et al., 2015). For fetal DNA chip array-based comparative genomic hybridization (aCGH) platform (Leung et al., 2011), the experiment was conducted with a total of 300 ng of DNA after treatment with RNase A (Qiagen) as input, and data analysis was performed with CytoGenomics software according to the manufacturer's protocol (Leung et al., 2011).

In 90.0% (45/50) of the cases, routine G-banded chromosome analysis of the CVS or AF was also performed by certified Medical Technologists/Cytogeneticists following international guidelines [Standards and Guidelines for Clinical Genetics Laboratories (Section 5.1) of the American College of Medical Genetics and Genomics (ACMG) and Requirements for Cytogenetic Testing of the National Pathology Accreditation Advisory Council]. A minimum of two independent cultures were set up for each specimen and 15–20 metaphases were analyzed with two metaphases karyotyped at a resolution of 400 bands. In cases suspected of mosaicism, at least 30 metaphases were analyzed. Karyograms were interpreted by at least two Medical Technologists/Cytogeneticists and diagnoses were reported according to the International System for Human Cytogenomic Nomenclature (ISCN) 2016.

GS for Fetal DNA

First, 100 ng of genomic DNA from each sample was sheared to fragment sizes ranging from 300 to 500 bp by the Covaris S2 Focused Ultrasonicator (Covaris, Inc., Woburn, MA). Library construction including end repairing, A-tailing, adapter ligation, and PCR amplification was conducted subsequently. The PCR products were then heat-denatured to form single-strand DNAs, followed by circularization with DNA ligase. After construction of the DNA nanoballs, paired-end sequencing with 100 bp at each end was carried out for each sample with a minimal read depth of 30-fold on the MGISEQ-2000 platform (BGI-Wuhan, Wuhan, China) (Huang et al., 2017).

Data Analysis and Variants Detection

QC for the paired-end reads was assessed *via* FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and subsequently aligned to the human reference genome (hg19) by Burrows–Wheeler Aligner (BWA) (Li and Durbin, 2009) and reformatted with SAMtools (Li et al., 2009). SNV and InDel detection was performed with HaplotypeCaller v3.4 from the Genome Analysis Toolkit (GATK, Broad Institute) (Mckenna et al., 2010) and annotation by ANNOVAR (Wang et al., 2010) and InterVAR (Li and Wang, 2017) with public and our in-house databases.

CNV detection and SV analyses were performed by our previously published methods (Dong et al., 2014; Dong et al., 2016; Dong et al., 2018a) with uniquely aligned reads/read pairs. For CNV analysis, all aligned reads were classified into each adjustable sliding window (50 kb with 5-kb increments) for identifying the candidate region(s) with CNVs and then they were classified again into each non-overlapping window (5 kb) for detection of the precise breakpoints with the module of Increment Rate of Coverage. The rare CNVs (*U* test, $P < 0.0001$) were then selected for further interpretation. For SV analysis, all chimeric read pairs, which were aligned to different chromosomes or to the same chromosome but with a distance larger than expected (> 10 kb), were selected for identification of translocations, inversions, insertions, or complex rearrangements. Briefly, 1) clustering: the chimeric read pairs were clustered by sorting the aligned coordinates; 2) systematic error filtering: each event was filtered against a control data set for the elimination of potential systematic errors; 3) random error filtering: each event was filtered with a cluster property matrix with the reported parameters; and 4) aligned orientations: each event was filtered based on p/q arm genetic exchange (joining type). Results of SNV, CNV, and SV analyses were integrated and reviewed for classification and interpretation of pathogenicity. The final results were also provided based on ISCN 2016.

Data Interpretation and Validation

The clinical significance of the detected SNVs, InDels, and CNVs were interpreted in accordance with the guidelines of the American College of Medical Genetics and Genomics (ACMG) and were classified into five categories: pathogenic, likely pathogenic, variant of uncertain significance, likely benign, or benign. Prioritization of SNVs/InDels in each sample was based on the following criteria: 1) whether reported by ClinVAR or HGMD (human mutation gene database); 2) with a minor allele frequency ≤5% in the databases of ExAC (http://exac.broadinstitute.org) and gnomAD (https://gnomad.broadinstitute.org); 3) located in coding region and exon–intron junctions; 4) with damaging/intolerant or splicing-change effect suggested by multiple biological algorithms (SIFT, Polyphen-2, MutationTaster, Human Splicing Finder, and MaxEntScan); and 5) located in an OMIM disease-causing gene. For known mutations, correlation of ultrasound finding(s) with the reported phenotype(s) was conducted. For novel variants, the priority of further classification was conducted as 1) located gene was reported to be in autosomal dominant or X-linked dominant manner; 2) affecting gene was in autosomal recessive or X-linked recessive manner, and a homozygous variant or more than one variant (suspected compound heterozygosity) were found. CNV interpretation was conducted based on our reported study (Dong et al., 2016). Potential disease-causing mutations were selected for validation and parental confirmation when available. No guidelines were available for SV interpretation; thus, gene disruption or potential gene dysregulation by the disruption of regulatory elements or topological associated domains were used for further interpretation. Parental confirmation was conducted if available for determination of the mode of inheritance of the variants.

SNV/InDel/SV were validated by Sanger sequencing. Genomic reference sequences (hg19) from each putative variant/breakpoint region were used for web-based primer design with

Primer3 (http://primer3.ut.ee/) and NCBI Primer-Blast (http://www.ncbi.nlm.nih.gov/tools/primer-blast/). PCR amplification was performed with each pair of primers (**Supplementary Table S1**) in cases and control (in-house DNA from a presumably normal male subject) simultaneously. PCR products were sequenced by Sanger sequencing on an ABI 3730 machine (Applied Biosystems, Thermo Fisher Scientific, Wilmington, DE) and sequencing results were aligned with BLAT (https://genome.ucsc.edu/cgi-bin/hgBlat?command=start) for confirmation of the mutations or re-arranged DNA sequences (Dong et al., 2016).

For CNV validation, quantitative PCR (qPCR) was conducted for additional disease-causing CNVs identified by GS. Genomic reference sequences (hg19) of each deleted/duplicated region were used for web-based primer design with Primer3 and NCBI Primer-Blast. Melting curve analysis was carried out for each pair of primers, and the PCR efficiency ranging 95% to 105% was determined by using the standard curve method. Each reaction was performed in 10 μl of reaction mixture simultaneously in cases and control in triplicate on a StepOnePlus Real-Time PCR System (Applied Biosystems) with SYBR Premix Ex Taq Tli RNaseH Plus (Takara Biotechnology, Dalian, China) and default setting of the reaction condition. The copy number in each sample was determined by using the ΔΔCt method, which compared the ΔCt (cycle threshold) of the target and a copy number neutral region [an ultra-conserved region (https://ccg.epfl.ch/UCNEbase/view.php?data = ucne&entry = 5530)] in the case with that of the control. Two independent primer pairs were used for each validation (Dong et al., 2018a).

RESULTS

Overall, we recruited 50 pregnancies with fetus with increased NT in the first-trimester Down syndrome screening. Thirty-seven CVS samples were collected from the first trimester, and AF samples were collected in later gestational weeks in the other 13 cases. Thirty-four cases were reported to have isolated increased NT with/without other soft markers (68.0%), and 16 cases were diagnosed with syndromic abnormalities (increased NT and fetal structural malformations, **Table 1**). All cases have undergone routine prenatal diagnosis by CMA, and 90% of these cases also have G-banded chromosome results available. CMA with/without karyotyping yielded diagnoses in eight cases, with pathogenic/likely pathogenic (P/LP) CNVs, providing a diagnostic yield of 16.0% (**Figure 1** and **Table 1**).

Numerical Disorders and CNVs

In this study, karyotyping identified four microscopic deletions/duplications, while CMA reported eight cases with P/LP CNVs and seven with variants of uncertain significant (VOUS or VUS, **Supplementary Table S2**). Compared with CMA results, GS not only detected all P/LP CNVs and VOUS detectable by CMA but also defined additional findings including a case of mosaic turner syndrome, a case of comprehensive delineation of CNV in one case with 16p11.2 recurrent deletion syndrome (**Table 2**), and two additional VOUS (**Supplementary Table S2**).

In 17NT0005, a fetus at the 12th gestational week with NT >14.3 mm presented with cystic hygroma. Normal CMA results were reported by the SNP-based platform and no karyotyping results were available (**Figures 2A, B** and **Supplementary Figure S1**). GS reported a mosaic turner syndrome with mosaic level around 40% (**Figure 2C**). Further validation by aCGH-based platform confirmed this finding and with consistent mosaic level estimated (**Figure 2D**). Mosaic turner syndrome was known to be the causative finding for the fetus with cystic hygroma, but it was missed by the original CMA in this study (Alpman et al., 2009). In the original CMA result, no decrease of copy ratio of the Y chromosome was indicated by the allele difference data (copy ratio as around 1, **Figure 2A**), although it was slightly decreased compared with the one of the X chromosome based on the coverage difference data (**Figure 2B**). It indicated that SNP-based platform might not be sensitive enough for detection of mosaicism on chromosome Y. Since the DNA source of this sample was from CVS, we could not exclude the fact that the possibility of this mosaic turner syndrome was due to confined placental mosaicism. However, further confirmation by testing with AF sample was not possible as the pregnancy was terminated due to the severe presentation in the fetus (cystic hygroma).

In total, GS identified one mosaic turner syndrome and eight cases with P/LP microdeletions/microduplications (**Table 2**) in this group.

TABLE 1 | Prenatal detection rates of the fetuses with increased NT by CMA/Karyotyping and GS.

Clinical indications	Number of cases	CMA with/without karyotyping		GS		*P* value
		Diagnostic yield	95% C.I. (%)#	Diagnostic yield	95% C.I. (%)#	
Isolated (increased NT with/without other soft markers)	34 (68%)	5/34 (14.7%)	5.0–31.1	10/34 (29.4%)	15.1–47.5	0.144$
Syndromic (increased NT with other fetal structural malformations)	16 (32%)	3/16 (18.8%)	4.0–45.6	6/16 (37.5%)	15.2–64.6	0.433*
Overall	50	8/50 (16%)	7.2–29.1	16/50 (32%)	19.5–46.7	0.061$

#95% confidence interval was calculated by binomial exact calculation.

$Pearson chi-square.

**Fisher's exact test.*

TABLE 2 | Summary of numerical disorder and pathogenic or likely pathogenic CNVs detected.

Case ID	NT (mm)	Other sonographic findings	Karyotype	CMA	CMA result	GS	Pregnancy outcome*
17NT0005	14.3	–	–	SNP typing	Normal	(X)x1.(Y)x0~1	TOP
18NT0018	3.5	–	46,*XY*	aCGH	arr[hg19] 16p11.2(29832358_30091372)x1	seq[hg19] del(16)(p11.2)dn chr16:g.29538256_30290160del	Live birth; no scoliosis identified
17BA0324	3.8	–	46,*XY*,1qh+	aCGH	arr[hg19] 20p13p12.3(4957885_8622907)x1	seq[hg19] del(20)(p13p12.3) chr20:g.4937184_8674795del	–
17BA0551	5.0	–	46,*XX*,add(8)(p23)	aCGH	arr[hg19]8p23.3p23.2(186477_5419654) x1, arr[GRCh37]8q22.1q24.3(98148502_146230986)x3	seq[hg19] del(8)(p23.3p23.2) chr8:g.10134_5523520del, seq[hg19] dup(8)(q22.1q24.3) chr8:g.98620704_146298884dup	–
18C0241	6.0	Double cleft palate small left ventricle with a large ventricular septal defect dilated sigmoid colon	46,*XY*,del(1)(q21q25)	aCGH	arr[hg19] 1q23.1q25.2(158043081_176445395)x1 dn	seq[hg19] del(1)(q23.1q25.2)dn chr1:g.157878750_176611630del	TOP
17C0070	7.4	–	46,*XY*,dup(2)(q11.2q21)	aCGH	arr[hg19] 2q11.2q21.1(97782763_131015556)x4 dn	seq[hg19] trp(2)(q11.2q21.1)dn chr2:g.98070117_131452568trp	–
17C0608	4.1	Absent nasal bone and possible congenital heart disease	46,*XY*,del(7)(q21q22)	aCGH	arr[hg19] 7q21.3q22.3(96787493_107041142)x1 dn	seq[hg19] del(7)(q21.3q22.3)dn chr7:g.96777381_107062981del	TOP
17C0963	5.4	Pleural effusion and umbilical hernia	46,*XX*	aCGH	arr[hg19] 22q11.21(18909032_21357982)x1	seq[hg19] del(22)(q11.21) chr22:g.18912476_21471029del	Miscarriage
17C1660	4.4	Right aortic arch	46,*XY*	aCGH	arr[hg19] 22q11.21(18909032_21357982) x1 dn	seq[hg19] del(22)(q11.21)dn chr22:g.18912979_21443877del	TOP

**TOP refers to termination of pregnancy.*

SNV Detection and Interpretation

To demonstrate the ability of detecting SNVs and InDels by GS, we reported seven cases with P/LP point mutations by GS following the ACMG guidelines (**Table 3**).

The first line of our detection was to report the cases with known disease-causing mutations. In case 16C1953, a male fetus presenting with increased NT (10.9 mm), cystic hygroma and bilateral large jugular lymphatic sacs, absence of ductus venosus a-wave, and bilateral bifid thumbs in the 12th gestational week was reported with normal CMA and karyotyping. GS detected a pathogenic heterozygous point mutation NM_004333:c.G1411T(p.V471F) in *BRAF* gene, which was reported to cause Cardio-Facio-Cutaneous Syndrome (Abe et al., 2012) and Noonan Syndrome (Nystrom et al., 2008; Croonen et al., 2013) in an autosomal dominant manner (**Table 3**). In 18NT0003, a female fetus with increased NT, hydrosarca, and short limbs at the 12th week of gestation was reported with normal CMA and karyotyping findings. A heterozygous point mutation NM_001844.4:c.G2950A(p.G984S) was detected in an autosomal dominant disease-causing gene *COL2A1* by GS. A different base change NM_001844.4:c.G2950C(p.G984R) in the same location was reported to cause achondrogenesis and type II (OMIM: #200610) or type II collagenopathies (Barat-Houari et al., 2016); in addition, the variant was confirmed to be a *de novo* mutation (**Figure 3A**); thus it was classified as a likely pathogenic causative variant for the fetal phenotype (**Table 3**).

In addition, we also identified five novel mutations in four samples, including variants in three autosomal dominant genes and one autosomal recessive gene (**Table 3**). For instance, 18NT018 is a female fetus with increased NT, absent nasal bone, and reverse a-wave ductus venosus observed in the first-trimester ultrasound screening with both normal karyotyping and CMA. Morphology scan in the second trimester revealed complex congenital heart disease (single atrium and ventricle) and polycystic kidneys. GS identified a novel heterozygous deletion, NM_003482:c.16474delG(p.D5492fs), which resulted in a frameshift in the *KMT2D* gene. Mutations in *KMT2D* have been known to cause Kabuki syndrome (OMIM: #147920) in an autosomal dominant manner, commonly resulting in heart defects and renal malformations consistent with the presentation of this fetus. Therefore, it was further classified as likely pathogenic (**Table 3**). Compound heterozygous point mutations were found in the fetus 14C1232: a paternally inherited frameshift deletion NM_018076:c.1614_1615del(p.P538fs) and a maternally inherited nonsynonymous mutation NM_018076:c.C2306A(p.P769H) were detected by GS and further confirmed by Sanger sequencing (**Supplementary Figure S2**). Mutations in gene *ARMC4* in an autosomal recessive manner are known to cause ciliopathies such as ciliary dyskinesia, primary, 23 (OMIM: #615451). The couple decided to keep the pregnancy and resulted in a live birth. Follow-up study is ongoing.

Furthermore, GS also detected 23 VOUS in 18 cases (**Figures 3B**, **C**, **Supplementary Table S3**).

SV Detection and the Potential Pathogenicity

Apart from detecting CNV and point mutations, GS was able to identify structural rearrangements by utilizing paired-end reads.

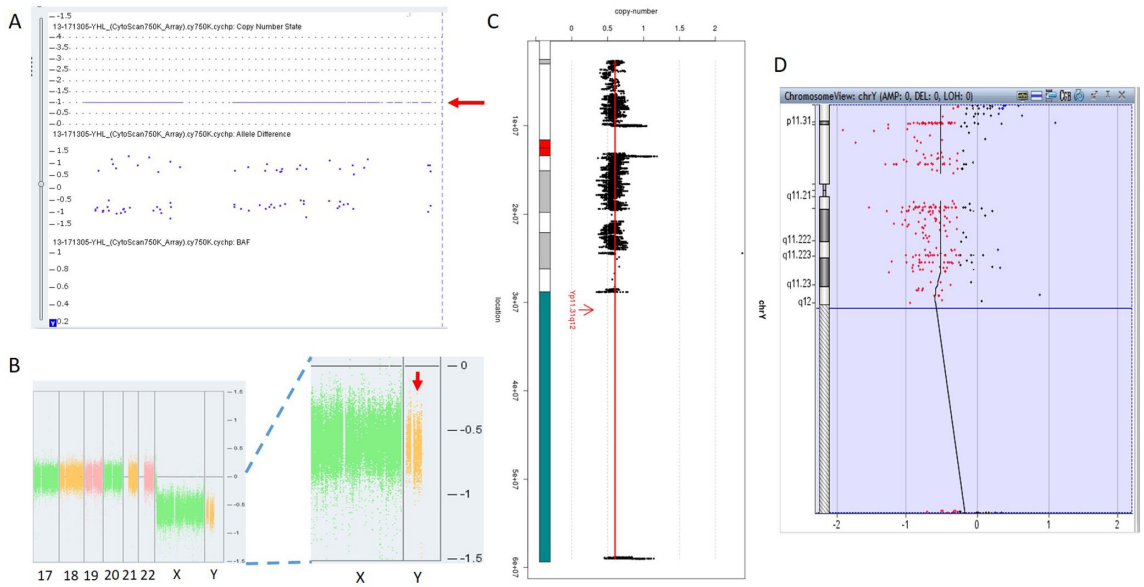

FIGURE 2 | Mosaic Turner syndrome detected by GS. In 17NT0005, CytoScan 750K CMA platform reported **(A)** a copy number as 1 for chromosome Y (indicated by a red arrow) and **(B)** apparently normal male fetus (partial figure of weighted Log2 ratio across each chromosome), although a slight decrease of Log2 ratio of chromosome Y was observed compared with the Log2 ratio of chromosome X (manifested by a red arrow). **(C)** GS reported a 40% decrease of the copy number of chromosome Y and confirmed by the aCGH CMA platform shown in **(D)**. In figure **(C)**, the copy ratio of each window is indicated by a black dot and the average copy ratio of chromosome Y is reflected by a red vertical line. Karyogram of chromosome Y is shown on the left. In figure **(D)**, dots in red, in black, and in blue indicate copy number lost, copy number neutral, and copy number gained, respectively. The average Log2 ratio is indicated by a black vertical line.

Overall, two inversions and two insertions were detected among these 50 cases (**Supplementary Table S4**).

In 18NT0003 described above, GS also reported a cryptic balanced complex insertion seq[hg19]ins(2;12)(q33.2;q24.31)g.[chr2:203384219_203384293inschr12:122757221_122907271cx]chr12:g.122757221_122907271del carrying a partial segment of gene *CLIP1* from chromosome 12 to the site of *BMPR2* in chromosome 2 (**Supplementary Table S4**, **Figure 3D**). The original fragment from chromosome 12 (seq[hg19]chr12:122757221_122907271) was 150.1 kb in size, but it was divided into 11 fragments with five segments lost when inserted to chromosome 2 (**Supplementary Figure S3**). There were 18 consecutive thymines (Ts) found in the breakpoint junction of the inserted site, not belonging to either chromosomes (2 and 12) (**Figure 3D**), indicating that the breakpoint repairing mechanism might be caused by non-homologous end joining with a cryptic insertion (Carvalho and Lupski, 2016). Parental confirmation by PCR and Sanger sequencing established this insertion to be a paternally inherited event. Regarding the potential pathogenicity, it has been known that mutations leading to loss of function in *BMPR2* would cause Primary Pulmonary Hypertension 1 and/or Pulmonary Venooclusive Disease 1 (OMIM: #600799). However, we were unable to determine whether the disruption of *BMRP2* was one of the causal factors of the fetal phenotype, since the pregnancy was terminated. We further followed up with the father (31 years old) and found sinus bradycardia by routine electrocardiograph. Although pulmonary hypertension may be later onset, sinus bradycardia might be one of the markers indicating pulmonary hypertension (Rajdev et al., 2012). Further follow-up is also ongoing.

For the other cryptic insertion and two inversions (**Supplementary Table S4**), since no OMIM genes were involved and no topological associated domains were disrupted, they were further classified as polymorphisms.

Overall Diagnostic Yield for Fetuses With Increased NT by GS

For these 50 cases, 68.0% of these cases had isolated increased NT with/without other soft markers. However, no significant difference of the diagnostic yields was found between isolated and syndromic groups (**Table 1**, chi-square test, $P = 0.5674$). In addition, GS provided a twofold diagnostic yield in all 50 cases compared with the routine test by CMA and/or karyotyping.

Comprehensive Detection and Delineation of the Variants in Individuals

In addition to the identification of individual variants (i.e., SNVs, CNVs and SVs), another advantage of using GS is to delineate the aberration regions involving various types of aberrations. For example, in 18NT0018, a male fetus with isolated increased NT, CMA reported a 259.0-kb *de novo* heterozygous deletion arr[hg19] 16p11.2(29832358_30091372)x1, which was diagnosed as 16p11.2 recurrent deletion syndrome (Lin et al., 2018). GS refined the breakpoints of the

TABLE 3 | Summary of pathogenic or likely pathogenic mutations revealed by GS.

Case	NT (mm)	Other sonographic findings	Gene	Mutation and het/hom*	Consequence	Inheritance mode	Disease association(s) [MIM #]	Pregnancy outcome	Inheritance confirmation#
14C1232	3.73	Mild ventricular disproportion	*ARMC4*	NM_018076:c.1614_1615del (p.P538fs) het	Frameshift	AR	Ciliary dyskinesia, primary, 23 (615451)	Live birth	Pat
				NM_018076: c.C2306A (p.P769H) het	Missense	AR	Ciliary dyskinesia, primary, 23 (615451)	Live birth	Mat
15C0337	4.18	–	*ANKRD11*	NM_001256182: c.2404dupC (p.L802fs) het	Frameshift	AD	KBG syndrome (148050)	Live birth; newborn exam: low set ears, increased nuchal fold; right hand single transverse crease; bilateral clinodactyly; postnatal follow-up at three years old showed bilateral hearing loss (40 db)	*De novo*
15C0802	5.03	–	*GATA4*	NM_002052:c.C1325T(p.A442V) het	Missense	AD	Atrial septal defect 2 (607941); Tetralogy of Fallot (187500)	Live birth	Mat
18NT019	3.5	–	*NSD1*	NM_022455:c.3797-2A > G het	Splicing site	AD	Sotos syndrome 1(117550)	–	–
16C1953	10.9	Large jugular lymphatic sac; cystic hygroma; absent ductus venosus a-wave; bilateral hand with bifid thumb.	*BRAF*	NM_004333: c.G1411T (p.V471F) het	Missense	AD	Noonan syndrome 7 (613706); Cardiofaciocutaneous syndrome (115150)	TOP	–
18NT0003	12	Hydrosarca; cystic hygroma; short long bones	*COL2A1*	NM_001844: c.G2950A (p.G984S) het	Missense	AD	Achondrogenesis, type II (200610)	TOP	*De novo*
18NT018	3.5	Reverse "a" wave in DV, absent nasal bone and possible congenital heart disease: single atrium, single ventricle, bilateral renal volume increased with polycystic changes	*KMT2D*	NM_003482: c.16474delG (p.D5492fs) het	Frameshift	AD	Kabuki syndrome 1(147920)	TOP	–

**het/hom refer to heterozygous mutation or homozygous.*

#Mat and Pat refer to maternally and paternally inherited, respectively.

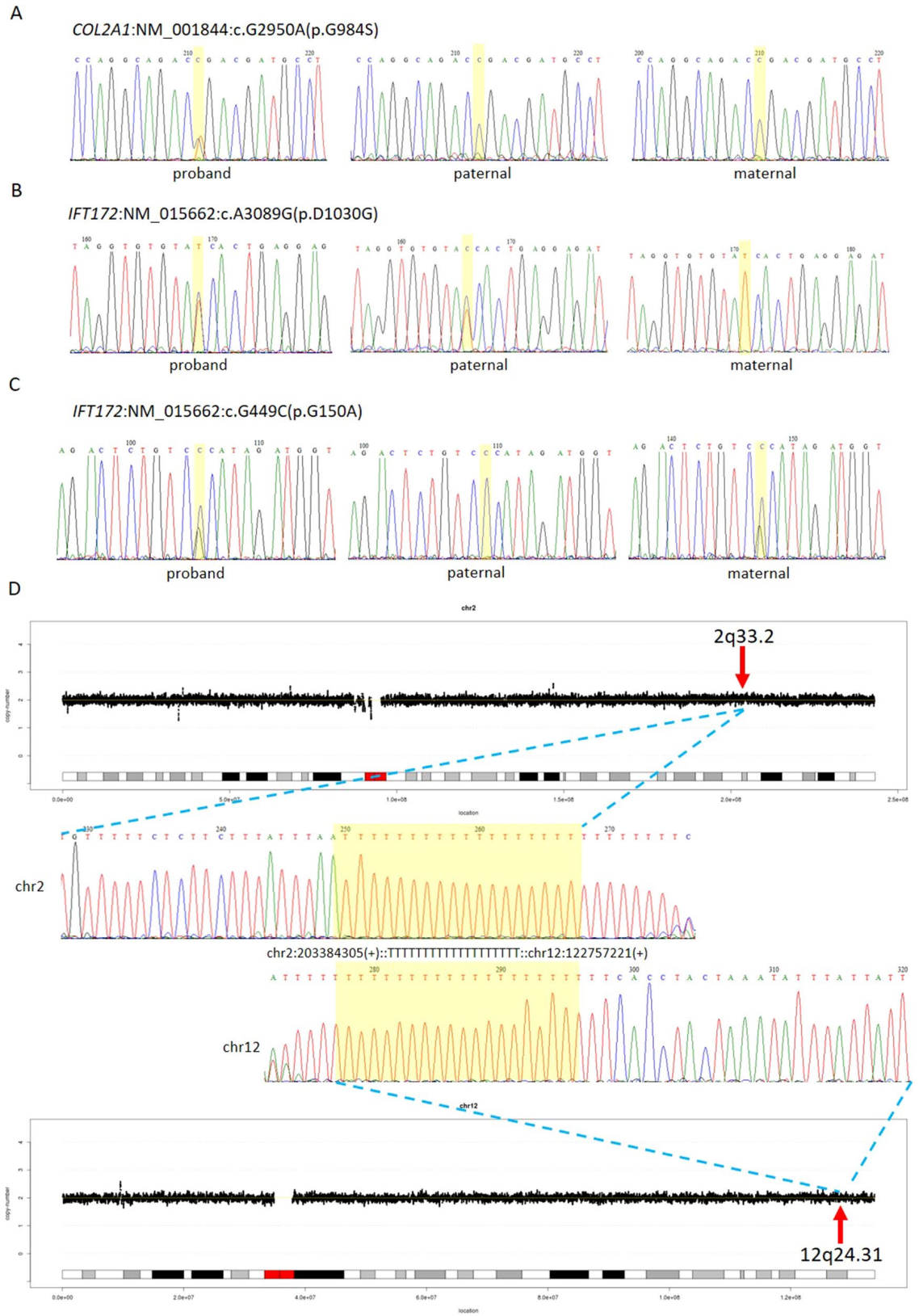
A
COL2A1:NM_001844:c.G2950A(p.G984S)
proband
paternal
maternal
B
IFT172:NM_015662:c.A3089G(p.D1030G)
proband
paternal
maternal
C
IFT172:NM_015662:c.G449C(p.G150A)
proband
paternal
maternal
D
chr2
2q33.2
chr2
chr2:203384305(+)::TTTTTTTTTTTTTTTTTT::chr12:122757221(+)
chr12
chr12
12q24.31

FIGURE 3 |

FIGURE 3 | Comprehensive definition of the genetic etiologies in 18NT0003. GS with parental confirmation reported **(A)** a *de novo* heterozygous mutation NM_001844:c.G2950A(p.G984S) in *COL2A1*, **(B)** a paternal heterozygous mutation NM_015662:c.A3089G: (p.D1030G) in *IFT172*, and **(C)** a maternal heterozygous mutation NM_015662:c.G449C: (p.G150A) in *IFT172*. All the mutated sites are highlighted in yellow. **(D)** Distributions of copy ratios in chromosome 2 and chromosome 12 are shown in the top and the bottom, respectively. The inserted site in chromosome 2 and the rearranged segments in chromosome 12 are indicated by red arrows with the band numbers; both of these two regions were copy number neutral. Dots in black indicate the copy ratio in each window across different chromosomes. *X* axis and *Y* axis represent the genomic coordinates and the copy ratios. Dotted lines in gray indicate the copy number 0, 1, 3, and 4, respectively. Sanger sequencing results of the breakpoints are shown in the middle. Forward sequencing is shown on the top, while the reverse complementary sequencing result of the reverse sequencing is shown at the bottom. The breakpoint coordinates in chromosome 2 and chromosome 12 are shown with the aligned orientation, while the inserted sequence is remarked in between and highlighted in yellow in both Sanger sequencing results.

CNV to a 751.9-kb *de novo* deletion seq[hg19] del(16p11.2) chr16:g.29538256_30290160del (**Figure 4A**) that involved the *TBX6* gene (T-Box 6, **Figure 4B**) not covered by the CNV reported by CMA. Heterozygous deletion of *TBX6* was further confirmed by qPCR experiment (**Figure 4C**) in the proband and the parents. The smaller CNV reported by CMA was contributed by the absence of probes in the region next to the reported region (**Figure 4B**). *TBX6* is currently involved in 16p11.2 recurrent deletion syndrome (Lin et al., 2018) and is highly correlated with scoliosis if there was a presence of a hemizygous T-C-A haplotype (Wu et al., 2015; Liu et al., 2019a). However, there was no mutant allele detected for each of these three common single-nucleotide polymorphisms (SNPs: rs2289292, rs3809624, and rs3809627) in the fetus. Further follow-up with the family: the pregnancy was kept and no scoliosis was found in the infant, echoing lower risk of having scoliosis in the absence of a hemizygous T-C-A haplotype in patient with 16p11.2 recurrent deletion syndrome (Wu et al., 2015; Liu et al., 2019a).

DISCUSSION

In this study, we applied GS for 50 fetuses with increased NT with/without other fetal structural malformations. In comparison, GS provided an overall diagnostic rate of 32.0%, which was a twofold increase compared to the current prenatal diagnosis tests (16.0%). Additional diagnoses include one mosaic turner syndrome, eight cases with P/LP CNVs (**Table 2**), and seven cases with P/LP point mutations (**Table 3**). In addition, GS reported two cryptic insertions and two inversions, and we further demonstrated the potential pathogenicity of disrupting OMIM disease-causing gene by the inserted site in the follow-up study. Overall, GS demonstrates the ability to detect the various disease-causing variants in human diseases such as fetuses with increased NT (**Figure 5**).

For the detection of numerical disorders and CNVs, GS detected additional findings including one case of mosaic turner syndrome and one case of comprehensive delineation of 16p11.2 recurrent deletion syndrome, compared with CMA (**Table 2**).

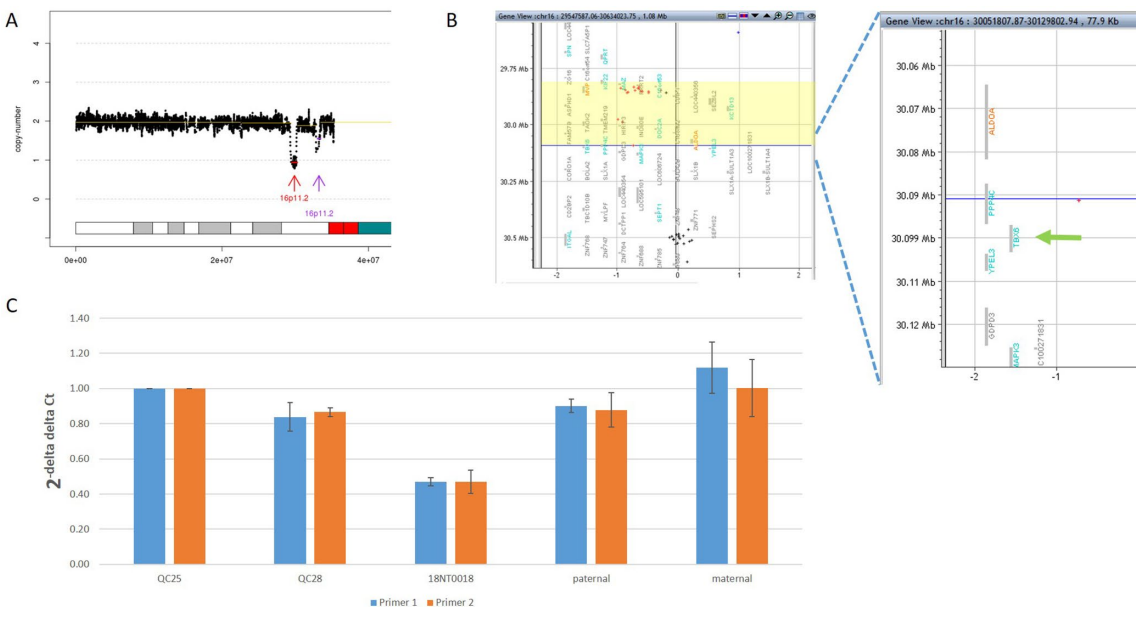

FIGURE 4 | Comprehensive delineation of 16p11.2 recurrent deletion syndrome. In 18NT0018, **(A)** GS reported a 751.9-kb *de novo* pathogenic deletion seq[hg19] del(16)(p11.2) chr16:g.29538256_30290160del (indicated by a red arrow) and a polymorphism deletion (population-based *U* test $P > 0.01$, indicated by a purple arrow) in chromosome 16p11.2; **(B)** original CMA reported a 259.0-kb heterozygous deletion arr[hg19] 16p11.2(29832358_30091372)x1 highlighted in yellow and without a probe locating in gene *TBX6*; **(C)** quantitative PCR two independent pairs of primers targeting *TBX6* shows approximately 0.5 copy ratio in 18NT0018 compared with normal control (Q25). The parental experiments confirmed that the heterozygous deletion in 18NT0018 is in *de novo* manner.

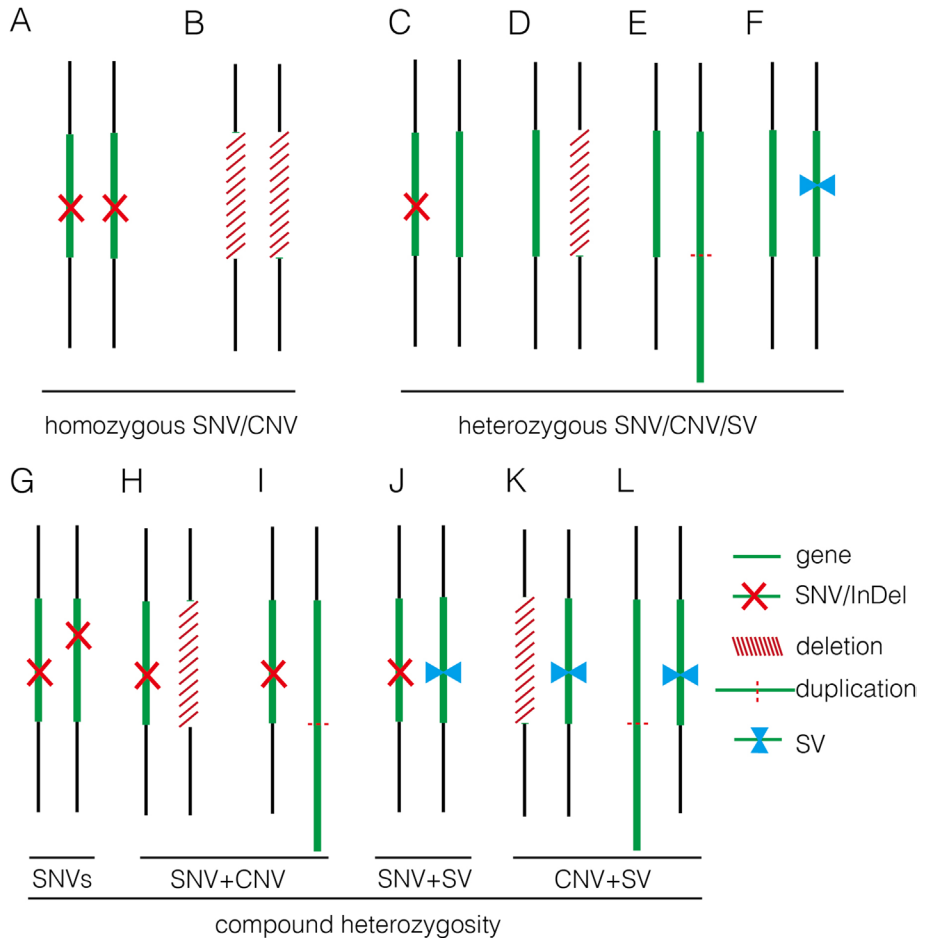

FIGURE 5 | Summary of various genomic variants proposed to contribute to human diseases. **(A)** and **(B)** show that the disease was caused by a homozygous SNV or a homozygous deletion commonly affecting a disease-causing gene in autosomal recessive manner. **(C–F)** indicate that the disease was caused by a heterozygous SNV [NM_004333: c.G1411T (p.V471F) in 16C1953], a heterozygous deletion (seq[hg19] del(16)(p11.2)dn chr16:g.29538256_30290160del in 18NT0018), a triplication (seq[hg19] trp(2)(q11.2q21.1)dn chr2:g.98070117_131452568trp in 17C0070), and a heterozygous SV (seq[hg19]ins(2;12)(q33.2;q24.31)g.[chr2:203384219_203384293inschr12:122757221_122907271cx] chr12:g.122757221_122907271del in 18NT0003), respectively, commonly affecting a disease-causing gene in autosomal dominant manner; **(G–L)** show that the disease was contributed by compound heterozygosity, a combination of two or more variants. **(G)** The typical mode of compound heterozygosity (NM_018076:c.1614_1615del (p.P538fs) and NM_018076: c.C2306A (p.P769H) from different parental origin in 14C1232). A gene content is indicated by a green bar. SNV/InDel, deletion, duplication, and SV are indicated by a red cross, a white bar (with red slashes), an elongated green bar (with red dotted line), and a pair of opposite blue arrows, respectively.

For the case with the mosaic turner syndrome (~40% mosaic level) additionally detected by GS (**Figure 2**), it was undetected by CMA possibly due to the algorithm employed for calling the copy number changes by calculating the allele differences. Further confirmation of this mosaic finding by using aCGH-based CMA platform shows the cross-platform differences of CMA testing (Haraksingh et al., 2017; Wang et al., 2019). Of note, in a routine setting, QF-PCR experiment was conducted prior to CMA for the detection of MCC and common aneuploidy. However, the missed detection of decreased copy ratio of the Y chromosome by QF-PCR in this case might be due to the low-level mosaicism presented (Shin et al., 2016). In addition, GS was able to delineate the 16p11.2 recurrent deletion, which included the disease-causing gene *TBX6*, the missed detection of which in this customized CMA platform was owing to the lack of probes located in the target region (**Figure 4B**). In this study, the detection results from only two CMA platforms were used for comparison and the missed detections and delineations described might be limited to these specific CMA platforms. A large-scale and systematic comparison of the performance of different currently available CMA platforms in prenatal diagnosis is warranted as certain platforms such as CytoScan HD and CytoSNP 850K have demonstrated to provide accurate breakpoints of CNVs detected and reliable detection of mosaicism, with a level greater than 20–30% (Haraksingh et al., 2017; Wang et al., 2019).

On top of the scope of current prenatal genetic diagnosis by karyotyping and CMA, GS also detected seven cases with P/LP SNVs and InDels, including three known pathogenic mutations/sites and five novel mutations (in four cases). The incidence of P/LP mutations in this study was 14.0%, which was higher than

the newly reported studies utilizing trio-based WES (Lord et al., 2019; Petrovski et al., 2019). A possible reason is the limited sample size or echoing the message that GS is more powerful than WES for detecting exome variants, particularly for SNVs and InDels (Belkadi et al., 2015). However, no WES data were available for comparison in our study.

Current clinical guidelines for variants interpretation are available for SNVs/InDels and CNVs, but there are no guidelines available for SV interpretation. Our study demonstrates the feasibility of detecting SVs by utilizing paired-end reads from GS. With that, we were able to identify inversions and insertions and to show the potential pathogenicity in a paternally inherited insertion disruption of gene *BMRP2*. Follow-up study conducted in the father indicated a sinus bradycardia from routine electrocardiograph. Detection of such cryptic rearrangements would help us explore a "blind spot" using the routine methods. However, CNVs and SVs are predominantly mediated by repetitive elements [commonly >1 kb (van Heesch et al., 2013)], which causes difficulty in the identification of the flanking unique sequences by standard GS with small-insert libraries employed in this study (Chen et al., 2008; Carvalho and Lupski, 2016; Dong et al., 2018a). Further study with other approaches such as mate-pair sequencing might be an alternative method for identifying additional pathogenic SVs (Talkowski et al., 2012a; Dong et al., 2019), which is enriched in the disease groups (Talkowski et al., 2012b). Overall, apart from identifying individual genomic variants, GS also shows its advantages in comprehensively defining different mutation types in different alleles. It is supported by the absence of a hemizygous T-C-A haplotype in the region of 16p11.2 recurrent deletion syndrome (Wu et al., 2015; Liu et al., 2019a) (**Figure 5**).

Overall, by taking together the various genomic variants, the diagnostic rates of GS were not significantly different between the isolated and syndromic groups (**Table 1**). For the isolated group, GS provided a diagnostic yield as 29.4%, which was significantly higher than previously reported (Lord et al., 2019). One of the reasons might be owing to the small sample size. Nonetheless, our current data truly show a higher diagnostic yield. Further study with a larger sample size is warranted. However, there were still 34 cases without a positive diagnosis by genome sequencing. Pregnancy outcomes were available for 18 cases, 7 (38.9%) of which were terminated pregnancies due to severe fetal malformations, such as pleural effusion, multicystic dysplastic kidneys, and complex congenital heart disease. It indicates that genome sequencing is still not revealing all the causative variants and the clinical decision of continuing or termination of pregnancy is largely dependent on the severity of fetal malformation(s) found in each particular case. Nonetheless, reanalysis of the GS data frequently upon updating literatures might be able to increase the genetic diagnosis (Liu et al., 2019b) and, thus, provide the causative answers for the family.

In this study, we aimed to evaluate the feasibility of applying GS for prenatal diagnosis in fetuses with increased NT. Apart from the increased diagnostic yield, the requirement of DNA amount as input for GS (100 ng) was less, compared with CMA (250 ng for CytoScan 750K or 300 ng for the 8×60K fetal chip), which would further facilitate GS in prenatal diagnosis, particularly for the AF samples from early gestational weeks. In addition, in comparison with the methods for routine prenatal diagnosis (CMA and karyotyping), GS would be able to provide a comparable turnaround time of 10 working days from DNA preparation to variant validation (**Figure 1**). Although the reagent cost of GS is less than 1000 USD per case, the demands of computational resources and labor costs for data analysis and interpretation are still higher than routine methods (**Figure 1**). Parental confirmation is important for determination of the variants' pathogenicity. In this study, we applied GS only for the proband and conducted parental confirmation by Sanger sequencing, which was laborious. Although applying trio-based GS testing would be helpful for variant filtering in initial data analysis, it would triplicate the cost, thus preventing the value of clinical application.

Regarding the interpretation of genomic variants, similarly with the prenatal studies by WES, even with parental confirmation, GS study would still yield a number of VOUS, the pathogenicity and clinical significance of which are still uncertain. One of the reasons would be lack of clinical symptom(s) in early pregnancy; follow-up with detailed diagnosis would be warranted. For instance, in 18NT0003, a fetus with increased NT, hydrosarca, and short limbs, GS reported a known LP mutation in *COL2A1* that might already explain the malformation. In addition, GS with parental confirmation also reported another two variants in a compound heterozygosity manner, including a paternally inherited heterozygous variant NM_015662:c.A3089G(p.D1030G) and a maternally inherited heterozygous variant NM_015662:c.G449C(p.G150A) in gene *IFT172* (**Figures 3B**, **C**). Mutations in *IFT172* are known to cause retinitis pigmentosa (OMIM: #616394) or short-rib thoracic dysplasia (OMIM: # 615630); the latter might be also be partially correlated with the fetal phenotype but might only present in later gestation. However, further confirmation of the pathogenicity by correlating the phenotype in later gestational weeks was not possible as the pregnancy was terminated. Nonetheless, GS still provides a comprehensive foundation for gene discovery. Moreover, one limitation of this study was the retrospective approach and the limited sample size, a prospective back to back in comparison of GS's performance with routine prenatal diagnosis approaches with a larger sample size, and other clinical indications are warranted in the near future.

In conclusion, our study demonstrates that applying GS in fetuses with increased NT can provide comprehensive detection of various disease-causing genomic variants, including SNVs, InDels, CNVs, and structural rearrangements. Compared with current routine prenatal diagnosis approaches, GS not only shows increased sensitivity for detecting mosaic numerical disorders and comprehensive delineation of P/LP CNVs but also provides the ability of identifying causative point mutations and balanced structural rearrangements that are beyond the detection scope of current prenatal diagnosis protocols. Although future prospective studies with larger sample sizes are warranted, given the comparable turnaround time and less DNA required in addition to the increased diagnostic yield, our study provides the first strong evidence to facilitate GS in prenatal diagnosis, particularly in fetuses with increased NT.

DATA AVAILABILITY

Genome sequencing data used in this study have been made available in the CNGB Nucleotide Sequence Archive (CNSA: https://db.cngb.org/cnsa/) under the accession number CNP0000437.

ETHICS STATEMENT

The study protocol was approved by the Ethics Committee of the Joint Chinese University of Hong Kong–New Territories East Cluster Clinical Research Ethics Committee (CREC Ref. No. 2016.713), Jinan University, and Guangzhou Medical University.

AUTHOR CONTRIBUTIONS

KWC, MC, TYL, and ZD designed the study. HW, JC, RZ, OC, YZ, and TYL collected the samples and followed up. KWC, MS, ZY, YC, YK, and ZD performed the analysis and data interpretation. MS, HY, YW, SC, and MHKC conducted the validation. KWC, HW, MS, JC, and ZD wrote the manuscript.

FUNDING

This project is supported by the National Natural Science Foundation of China (81741009, 81671470, and 31801042), the Health and Medical Research Fund (04152666 and 01120766), the Guangzhou Science and Technology Program (2014 Y2-00551, 201504282321393, 201604020078, and 201604020091), and the National Key Research and Development Program of China (2018YFC1004104).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2019.00761/full#supplementary-material.

REFERENCES

Abe, Y., Aoki, Y., Kuriyama, S., Kawame, H., Okamoto, N., Kurosawa, K., et al. (2012). Prevalence and clinical features of Costello syndrome and cardio-facio-cutaneous syndrome in Japan: findings from a nationwide epidemiological survey. *Am. J. Med. Genet. A.* 158A, 1083–1094. doi: 10.1002/ajmg.a.35292

Alpman, A., Cogulu, O., Akgul, M., Arikan, E. A., Durmaz, B., Karaca, E., et al. (2009). Prenatally diagnosed Turner syndrome and cystic hygroma: incidence and reasons for referrals. *Fetal. Diagn. Ther.* 25, 58–61. doi: 10.1159/000199869

Barat-Houari, M., Sarrabay, G., Gatinois, V., Fabre, A., Dumont, B., Genevieve, D., et al. (2016). Mutation update for COL2A1 gene variants associated with type II collagenopathies. *Hum. Mutat.* 37, 7–15. doi: 10.1002/humu.22915

Belkadi, A., Bolze, A., Itan, Y., Cobat, A., Vincent, Q. B., Antipenko, A., et al. (2015). *Whole-genome sequencing is more powerful than whole-exome sequencing for detecting exome variants. Proc. Natl. Acad. Sci. U. S. A.* 112, 5473–5478. doi: 10.1073/pnas.1418631112

Carvalho, C. M., and Lupski, J. R. (2016). Mechanisms underlying structural variant formation in genomic disorders. *Nat. Rev. Genet.* 17, 224–238. doi: 10.1038/nrg.2015.25

Chau, M. H. K., Cao, Y., Yvonne Kwok, K. Y., Chan, S., Chan, Y. M., Wang, H., et al. (2019). Characteristics and mode of inheritance of pathogenic copy number variants in prenatal diagnosis. *Am. J. Obstet. Gynecol.* doi: 10.1016/j.ajog.2019.06.007

Chen, W., Kalscheuer, V., Tzschach, A., Menzel, C., Ullmann, R., Schulz, M. H., et al. (2008). Mapping translocation breakpoints by next-generation sequencing. *Genome Res.* 18, 1143–1149. doi: 10.1101/gr.076166.108

Choy, K. W., Kwok, Y. K., Cheng, Y. K., Wong, K. M., Wong, H. K., Leung, K. O., et al. (2014). Diagnostic accuracy of the BACs-on-beads assay versus karyotyping for prenatal detection of chromosomal abnormalities: a retrospective consecutive case series. *BJOG* 121, 1245–1252. doi: 10.1111/1471-0528.12873

Croonen, E. A., Nillesen, W., Schrander, C., Jongmans, M., Scheffer, H., Noordam, C., et al. (2013). Noonan syndrome: comparing mutation-positive with mutation-negative dutch patients. *Mol. Syndromol.* 4, 227–234. doi: 10.1159/000350686

Dong, Z., Jiang, L., Yang, C., Hu, H., Wang, X., Chen, H., et al. (2014). A robust approach for blind detection of balanced chromosomal rearrangements with whole-genome low-coverage sequencing. *Hum. Mutat.* 35, 625–636. doi: 10.1002/humu.22541

Dong, Z., Wang, H., Chen, H., Jiang, H., Yuan, J., Yang, Z., et al. (2018a). Identification of balanced chromosomal rearrangements previously unknown among participants in the 1000 Genomes Project: implications for interpretation of structural variation in genomes and the future of clinical cytogenetics. *Genet. Med.* 20, 697–707. doi: 10.1038/gim.2017.170

Dong, Z., Xie, W., Chen, H., Xu, J., Wang, H., Li, Y., et al. (2017). Copy-number variants detection by low-pass whole-genome sequencing. *Curr. Protoc. Hum. Genet.* 94, 8 17 11–8 17 16. doi: 10.1002/cphg.43

Dong, Z., Ye, L., Yang, Z., Chen, H., Yuan, J., Wang, H., et al. (2018b). Balanced chromosomal rearrangement detection by low-pass whole-genome sequencing. *Curr. Protoc. Hum. Genet.* 96, 8 18 11–18 18 16. doi: 10.1002/cphg.51

Dong, Z., Zhang, J., Hu, P., Chen, H., Xu, J., Tian, Q., et al. (2016). Low-pass whole-genome sequencing in clinical cytogenetics: a validated approach. *Genet. Med.* 18, 940–948. doi: 10.1038/gim.2015.199

Dong, Z., Zhao, X., Li, Q., Yang, Z., Xi, Y., Alexeev, A., et al. (2019). Development of coupling controlled polymerizations by adapter-ligation in mate-pair sequencing for detection of various genomic variants in one single assay. *DNA Res.* doi: 10.1093/dnares/dsz011

Drury, S., Williams, H., Trump, N., Boustred, C., Lench, N., Scott, R. H., et al. (2015). Exome sequencing for prenatal diagnosis of fetuses with sonographic abnormalities. *Prenat. Diagn.* 35, 1010–1017. doi: 10.1002/pd.4675

Fu, F., Li, R., Li, Y., Nie, Z. Q., Lei, T., Wang, D., et al. (2018). Whole exome sequencing as a diagnostic adjunct to clinical testing in fetuses with structural abnormalities. *Ultrasound Obstet. Gynecol.* 51, 493–502. doi: 10.1002/uog.18915

Haraksingh, R. R., Abyzov, A., and Urban, A. E. (2017). Comprehensive performance comparison of high-resolution array platforms for genome-wide Copy Number Variation (CNV) analysis in humans. *BMC Genomics* 18, 321. doi: 10.1186/s12864-017-3658-x

Huang, J., Liang, X., Xuan, Y., Geng, C., Li, Y., Lu, H., et al. (2017). A reference human genome dataset of the BGISEQ-500 sequencer. *Gigascience* 6, 1–9. doi: 10.1093/gigascience/gix024

Huang, J., Poon, L. C., Akolekar, R., Choy, K. W., Leung, T. Y., and Nicolaides, K. H. (2014). Is high fetal nuchal translucency associated with submicroscopic chromosomal abnormalities on array CGH? *Ultrasound Obstet. Gynecol.* 43, 620–624. doi: 10.1002/uog.13384

Leung, G. K. C., Mak, C. C. Y., Fung, J. L. F., Wong, W. H. S., Tsang, M. H. Y., Yu, M. H. C., et al. (2018). Identifying the genetic causes for prenatally diagnosed structural congenital anomalies (SCAs) by whole-exome sequencing (WES). *BMC Med. Genomics* 11, 93. doi: 10.1186/s12920-018-0409-z

Leung, T. Y., Vogel, I., Lau, T. K., Chong, W., Hyett, J. A., Petersen, O. B., et al. (2011). Identification of submicroscopic chromosomal aberrations in fetuses with increased nuchal translucency and apparently normal karyotype. *Ultrasound Obstet. Gynecol.* 38, 314–319. doi: 10.1002/uog.8988

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25, 1754–1760. doi: 10.1093/bioinformatics/btp324

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352

Li, Q., and Wang, K. (2017). InterVar: clinical interpretation of genetic variants by the 2015 ACMG-AMP guidelines. *Am. J. Hum. Genet.* 100, 267–280. doi: 10.1016/j.ajhg.2017.01.004

Lin, S., Shi, S., Zhou, Y., Ji, Y., Huang, P., Wu, J., et al. (2018). Intrauterine phenotypic features associated with 16p11.2 recurrent microdeletions. *Prenat. Diagn.* 38, 381–389. doi: 10.1002/pd.5245

Lin, S. B., Xie, Y. J., Chen, Z., Zhou, Y., Wu, J. Z., Zhang, Z. Q., et al. (2015). Improved assay performance of single nucleotide polymorphism array over conventional karyotyping in analyzing products of conception. *J. Chin. Med. Assoc.* 78, 408–413. doi: 10.1016/j.jcma.2015.03.010

Liu, J., Wu, N., Deciphering Disorders Involving, S., Study, C.O, Yang, N., Takeda, K., Chen, W., et al. (2019a). TBX6-associated congenital scoliosis (TACS) as a clinically distinguishable subtype of congenital scoliosis: further evidence supporting the compound inheritance and TBX6 gene dosage model. *Genet. Med.* doi: 10.1038/s41436-018-0377-x

Liu, P., Meng, L., Normand, E. A., Xia, F., Song, X., Ghazi, A., et al. (2019b). Reanalysis of clinical exome sequencing data. *N. Engl. J. Med.* 380, 2478–2480. doi: 10.1056/NEJMc1812033

Lord, J., Mcmullan, D. J., Eberhardt, R. Y., Rinck, G., Hamilton, S. J., Quinlan-Jones, E., et al. (2019). Prenatal exome sequencing analysis in fetal structural anomalies detected by ultrasonography (PAGE): a cohort study. *Lancet* 393, 747–757. doi: 10.1016/S0140-6736(18)31940-8

Mckenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., et al. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20, 1297–1303. doi: 10.1101/gr.107524.110

Normand, E. A., Braxton, A., Nassef, S., Ward, P. A., Vetrini, F., He, W., et al. (2018). Clinical exome sequencing for fetuses with ultrasound abnormalities and a suspected Mendelian disorder. *Genome Med.* 10, 74. doi: 10.1186/s13073-018-0582-x

Nystrom, A. M., Ekvall, S., Berglund, E., Bjorkqvist, M., Braathen, G., Duchen, K., et al. (2008). Noonan and cardio-facio-cutaneous syndromes: two clinically and genetically overlapping disorders. *J. Med. Genet.* 45, 500–506. doi: 10.1136/jmg.2008.057653

Petrovski, S., Aggarwal, V., Giordano, J. L., Stosic, M., Wou, K., Bier, L., et al. (2019). Whole-exome sequencing in the evaluation of fetal structural anomalies: a prospective cohort study. *Lancet* 393, 758–767. doi: 10.1016/S0140-6736(18)32042-7

Rajdev, A., Garan, H., and Biviano, A. (2012). Arrhythmias in pulmonary arterial hypertension. *Prog. Cardiovasc. Dis.* 55, 180–186. doi: 10.1016/j.pcad.2012.06.002

Shin, Y. J., Chung, J. H., Kim, D. J., Ryu, H. M., Kim, M. Y., Han, J. Y., et al. (2016). Quantitative fluorescent polymerase chain reaction for rapid prenatal diagnosis of fetal aneuploidies in chorionic villus sampling in a single institution. *Obstet. Gynecol. Sci.* 59, 444–453. doi: 10.5468/ogs.2016.59.6.444

Sinajon, P., Chitayat, D., Roifman, M., Wasim, S., Carmona, S., Ryan, G., et al. (2019). Increased nuchal translucency: results from microarray and RASopathy disorders testing. *Ultrasound Obstet. Gynecol.* doi: 10.1002/uog.20352

Socolov, D., Socolov, R., Gorduza, V. E., Butureanu, T., Stanculescu, R., Carauleanu, A., et al. (2017). Increased nuchal translucency in fetuses with a normal karyotype-diagnosis and management: an observational study. *Medicine (Baltimore)* 96, e7521. doi: 10.1097/MD.0000000000007521

Talkowski, M. E., Ordulu, Z., Pillalamarri, V., Benson, C. B., Blumenthal, I., Connolly, S., et al. (2012a). Clinical diagnosis by whole-genome sequencing of a prenatal sample. *N. Engl. J. Med.* 367, 2226–2232. doi: 10.1056/NEJMoa1208594

Talkowski, M. E., Rosenfeld, J. A., Blumenthal, I., Pillalamarri, V., Chiang, C., Heilbut, A., et al. (2012b). Sequencing chromosomal abnormalities reveals neurodevelopmental loci that confer risk across diagnostic boundaries. *Cell* 149, 525–537. doi: 10.1016/j.cell.2012.03.028

Van Heesch, S., Kloosterman, W. P., Lansu, N., Ruzius, F. P., Levandowsky, E., Lee, C. C., et al. (2013). Improving mammalian genome scaffolding using large insert mate-pair next-generation sequencing. *BMC Genomics* 14, 257. doi: 10.1186/1471-2164-14-257

Wang, J. C., Radcliff, J., Coe, S. J., and Mahon, L. W. (2019). Effects of platforms, size filter cutoffs, and targeted regions of cytogenomic microarray on detection of copy number variants and uniparental disomy in prenatal diagnosis: results from 5026 pregnancies. *Prenat. Diagn.* 39, 137–156. doi: 10.1002/pd.5375

Wang, K., Li, M., and Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* 38, e164. doi: 10.1093/nar/gkq603

Wu, N., Ming, X., Xiao, J., Wu, Z., Chen, X., Shinawi, M., et al. (2015). TBX6 null variants and a common hypomorphic allele in congenital scoliosis. *N. Engl. J. Med.* 372, 341–350. doi: 10.1056/NEJMoa1406829

Yang, X., Li, R., Fu, F., Zhang, Y., Li, D., and Liao, C. (2017). Submicroscopic chromosomal abnormalities in fetuses with increased nuchal translucency and normal karyotype. *J. Matern. Fetal. Neonatal. Med.* 30, 194–198. doi: 10.3109/14767058.2016.1168394

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Genetics

ORIGINAL RESEARCH
published: 30 August 2019
doi: 10.3389/fgene.2019.00813

Prenatal Diagnosis of Microdeletions or Microduplications in the Proximal, Central, and Distal Regions of Chromosome 22q11.2: Ultrasound Findings and Pregnancy Outcome

*Shuyuan Li[1,2,3], Xu Han[1], Mujin Ye[1,2,3], Songchang Chen[1,2,3], Yinghua Shen[1], Jianmei Niu[1], Yanlin Wang[1] and Chenming Xu[1,2,3]**

[1] International Peace Maternity and Child Health Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai, China, [2] Shanghai Key Laboratory of Embryo Original Diseases, Shanghai, China, [3] Institute of Embryo-Fetal Original Adult Disease, Shanghai Jiao Tong University School of Medicine, Shanghai, China

OPEN ACCESS

Edited by:
Fan Jin,
Zhejiang University, China

Reviewed by:
Muhammad Jawad Hassan,
National University of Medical Sciences (NUMS), Pakistan
Richard Choy,
The Chinese University of Hong Kong, China

***Correspondence:**
Chenming Xu
chenming_xu2006@163.com

Specialty section:
This article was submitted to Genetic Disorders, a section of the journal Frontiers in Genetics

Received: 03 April 2019
Accepted: 06 August 2019
Published: 30 August 2019

Citation:
Li S, Han X, Ye M, Chen S, Shen Y, Niu J, Wang Y and Xu C (2019) Prenatal Diagnosis of Microdeletions or Microduplications in the Proximal, Central, and Distal Regions of Chromosome 22q11.2: Ultrasound Findings and Pregnancy Outcome. Front. Genet. 10:813. doi: 10.3389/fgene.2019.00813

Several recurrent microdeletions and microduplications in the proximal, central, and distal regions of chromosomal 22q11.2 have been identified. However, due to a limited number of patients reported in the literature, highly variable clinical phenotypes, and incomplete penetrance, the pathogenicity of some microdeletions/microduplications in 22q11.2 central and distal regions is unclear. Hence, the genetic counseling and subsequent pregnancy decision are extremely challenging, especially when they are found in structurally normal fetuses. Here, we reported 27 consecutive cases diagnosed prenatally with 22q11.2 microdeletions or microduplications by chromosomal microarray analysis in our center. The prenatal ultrasound features, inheritance of the microdeletions/microduplications, and their effects on the pregnancy outcome were studied. We found that fetuses with 22q11.2 microdeletions were more likely to present with structure defects in the ultrasound, as compared with fetuses with 22q11.2 microduplications. Both the prenatal ultrasound findings and the inheritance of the microdeletions/microduplications affected the parent's decision of pregnancy. Those with structure defects in prenatal ultrasound or occurred *de novo* often resulted in termination of the pregnancy, whereas those with normal ultrasound and inherited from healthy parent were likely to continue the pregnancy and led to normal birth. Our study emphasized that proximal, central, and distal 22q11.2 deletions or duplications were different from each other, although some common features were shared among them. More studies are warranted to demonstrate the underlying mechanisms of different clinical features of these recurrent copy-number variations, thereby to provide more information for genetic counseling of 22q11.2 microdeletions and microduplications when they are detected prenatally.

Keywords: 22q11.2, microdeletions, microduplications, prenatal diagnosis, genetic counseling

INTRODUCTION

Low copy repeats (LCRs), also known as segmental duplications, are highly homologous sequence (greater than 95% sequence identity) and comprise approximately 4–5% of the human genome (Bailey et al., 2001). Misalignment of LCRs during meiosis through the well-established mechanism of nonallelic homologous recombination (NAHR) can lead to recurrent copy-number variations (CNVs), including microdeletions and microduplication. When dosage-sensitive gene(s) involved, the microdeletions or microduplications may result in abnormal phenotypes (Dittwald et al., 2013). Eight LCRs, naming LCR22A-H, have been identified in chromosome 22q11.2, one region showing high frequency of genomic rearrangement (Shaikh et al., 2007). Several recurrent microdeletions and microduplications in chromosome 22q11.2 have been identified, including chromosome 22q11.2 deletion syndrome [also known as DiGeorge syndrome (#188400) or velocardiofacial syndrome (#192430), hereafter "22q11.2DS"], chromosome 22q11.2 deletion syndrome, distal (#611867), chromosome 22q11.2 duplication syndrome (#608363), and some others not recorded in the Online Mendelian Inheritance in Man database (22q11.2 central deletion or duplication, 22q11.2 distal duplication, etc.) (**Table 1** and **Figure 1**) (Burnside, 2015).

The 22q11.2DS is the most common recurrent microdeletions in humans with a frequency estimated at 1:4,000 to 1:8,000 live births (McDonald-McGinn et al., 2015). The clinical features of patients with 22q11.2DS are variable and include cardiac defects, palatal abnormalities, characteristic facial features, learning difficulties, and immune deficiencies (Bassett et al., 2011). Approximately 85–90% of patients with 22q11.2DS result from a 3-Mb deletion extending from LCR22A to LCR22D, while 10–15% of the patients have a smaller "nested" ~1.5-Mb deletions involving LCR22A to LCR22B (McDonald-McGinn et al., 2015). The *TBX1* (*602054) gene, located between the LCR22A and LCR22B, is the main candidate gene responsible for most of the features of 22q11.2DS, and the phenotypes of patients with LCR22A-B and LCR22A-D deletions are clinically indistinguishable (Carlson et al., 1997). Microduplications of the same region as 22q11.2DS have also been reported and are defined as chromosome 22q11.2 duplication syndrome (#608363) (Ensenauer et al., 2003; Hassed et al., 2004). The phenotypes of individuals with chromosome 22q11.2 duplication syndrome are highly variable, which range from apparently normal to severe malformations with developmental delay (Hassed et al., 2004; Wentzel et al., 2008). Some rare atypical deletions/duplications of shorter size, mainly involved LCR22B-D or LCR22C-D and not encompassed the *TBX1* gene, have been reported and now are proposed as "central" 22q11.2 deletions/duplications (Burnside, 2015; Rump et al., 2014). The "distal" 22q11.2 deletions/duplications, mediated by NAHR of the five distal LCR22s, LCR22D-H, have also been demonstrated (Ben-Shachar et al., 2008). Although the patients with central or distal 22q11.2 deletions/duplications share some characteristic features with 22q11.2DS, they have unique clinical characterizes with high phenotypic variability (Burnside, 2015; Rump et al., 2014).

TABLE 1 | Reported microdeletions or microduplications in chromosomal 22q11.2.

	LCR	Chromosome physical location (hg19)	OMIM syndrome	Haploinsufficiency/ Triplosensitivity score [a]	Classification of pathogenicity
22q 11.2 microdeletion					
Proximal	A-B/D	18,912,231-20,287,208/21,465,672	Chromosome 22q11.2 deletion syndrome: DGS (#188400) or VCFS (# 192430)	3	Pathogenic
Central	B/C-D	20,731,986-21,465,672	/	2	VOUS-LP
Distal Type I	D-E/F	21,917,117-23,649,111	Chromosome 22q11.2 deletion syndrome, distal (#611867)	3	Pathogenic
Type II	E-F	23,119,414-23,649,111	/	2	VOUS-LP
Type III	F-G	23,831,202-24,632,821	/	NA	VOUS
	E-H	23,119,414-24,994,433	/	NA	VOUS
	D-H	21,917,117-24,994,433	Chromosome 22q11.2 deletion syndrome, distal (#611867)	NA	Pathogenic
22q11.2 microduplication					
Proximal	A-B/D	18,912,231-20,287,208/21,465,672	Chromosome 22q11.2 duplication syndrome (#608363)	3	Pathogenic
Central	B/C-D	20,731,986-21,465,672	/	1	VOUS
Distal Type I	D-E/F	21,917,117-23,649,111	/	3	Pathogenic
Type II	E-F	23,119,414-23,649,111	/	1	VOUS
Type III	F-G	23,831,202-24,632,821	/	NA	VOUS
	E-H	23,119,414-24,994,433	/	NA	VOUS
	D-H	21,917,117-24,994,433	/	NA	Pathogenic

[a]*Haploinsufficiency score (for deletions) and triplosensitivity score (for duplications) of the region curated in the ClinGen Dosage Sensitivity Map (https://www.ncbi.nlm.nih.gov/projects/dbvar/clingen).*

LCR, low copy repeats; DGS, DiGeorge syndrome; VCFS, velocardiofacial syndrome; NA, not available; VOUS, variant of unknown significance; LP, likely pathogenic; OMIM, Online Mendelian Inheritance in Man (www.omim.org).

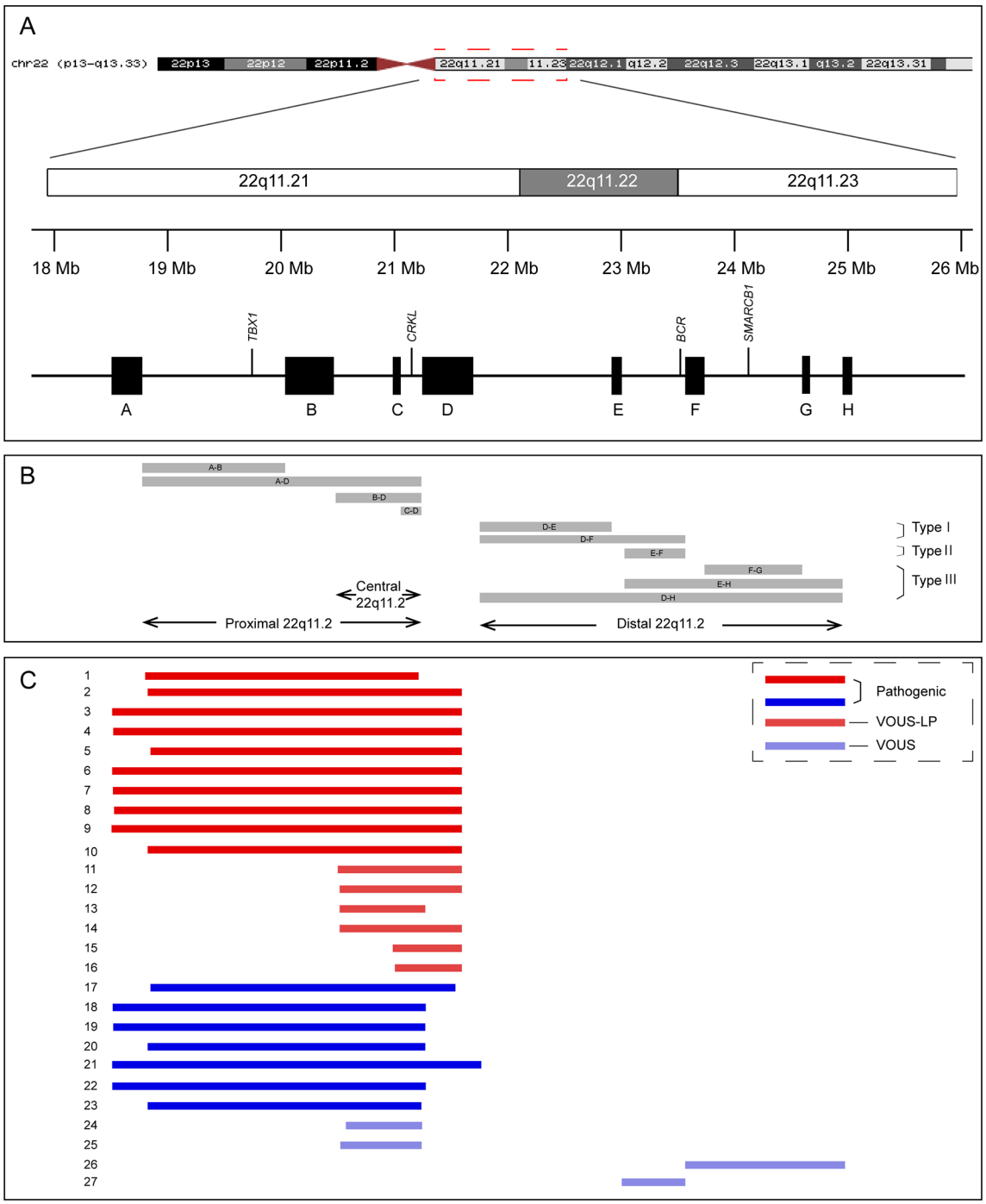

FIGURE 1 | Schematic representation of chromosome 22q11.2 region **(A)**, the recurrent copy number variations reported in this region **(B)**, and the 27 cases included in this study **(C)**. VOUS, variant of unknown significance; LP, likely pathogenic.

Chromosomal microarray analysis (CMA) is a high-resolution technology capable of detecting aneuploidy, as well as microduplications and microdeletions, throughout the genome. The use of CMA in prenatal diagnosis has been recommended by the American College of Obstetricians and Gynecologist in 2013 (American College of Obstetricians and Gynecologists Commitee on Genetics, 2013). With the wide use of CMA in prenatal diagnosis, more and more fetuses with 22q11.2 microdeletions or microduplications with variable sizes have been identified. It has been reported that the prevalence of 22q11.2DS in fetuses with congenital heart defects is as high as 7% (Mademont-Soler et al., 2013). However, due to a limited number of patients reported in

the literature, highly variable clinical phenotypes, and incomplete penetrance, the pathogenicity of microdeletions/microduplications in 22q11.2 central and distal regions (types II and III) is unclear (variants of unknown significance) (**Figure 1** and **Table 1**). Hence, the genetic counseling and subsequent pregnancy decision are extremely difficult, especially when they are found in structurally normal fetuses.

Here, we reported 27 consecutive cases diagnosed prenatally with 22q11.2 microdeletions or microduplications by CMA from December 2015 to September 2018 in our center. The prenatal ultrasound features, inheritance of the CNVs, and their effects on the pregnancy outcome were studied. Our study will provide more information for genetic counseling of 22q11.2 microdeletions and microduplications in prenatal diagnosis.

METHODS

Study Population

This study was conducted in the Reproductive Genetic Center of International Peace Maternal and Child Health Hospital (IPMCH) of Shanghai Jiao Tong University School of Medicine. From December 2015 to September 2018, 5,464 pregnant women received an invasive prenatal diagnostic test for CMA analysis in our center. Among them, 16 fetuses of 22q11.2 microdeletion (0.29%) and 11 fetuses of 22q11.2 microduplication (0.20%) were detected by CMA. According to the gestational age (range: 12–28 weeks, median: 22 weeks), fetal samples were obtained using chorionic villus sampling (n = 1), amniocentesis (n = 21), or cord blood sampling (n = 5). All cases diagnosed with 22q11.2 microdeletions or microduplications were further consulted regarding the prognosis and additionally followed up for the clinical outcome. Written informed consent was obtained from the parents in accordance with the Declaration of Helsinki, and the study was approved by the Ethics Committee of the IPMCH (number of Institutional Review Board approval: GJEC-A-2015-11-1).

Chromosomal Microarray Analysis and Quantitative Real-Time PCR

Genomic DNA was isolated according to standard procedures (Li et al., 2019). CMA was performed using Agilent 4X180K SurePrint Prenatal Research Array (Agilent Technologies, Santa Clara, CA,USA) from December 2015 to August 2016 (n = 2) and using Affymetrix CytoScan 750K Array (Affymetrix, Inc., Santa Clara, CA, USA) from September 2016 to September 2018 (n = 25). CNVs were determined using Agilent CytoGenomics (Agilent Technologies, Santa Clara, CA, USA) or Affymetrix Chromosome Analysis Suite software 3.2 (Affymetrix, Inc., Santa Clara, CA, USA), depending on the platform that was used. All results were evaluated using the University of California Santa Cruz human Genome Browser release of February 2009 (GRCh37/hg19). When parental blood samples were available, the inheritance of the detected microdeletions or microduplications of 22q11.2 was determined using CytoScan® 750K Array or real-time quantitative PCR (qPCR). The qPCR was performed in a LightCycler 480 II (Roche Applied Science, Mannheim, Germany) qPCR machine according to the manufacturer's instructions. The *HBB* gene was used as housekeeping gene, and the qPCR primers were shown in **Table S1**. Each sample was analyzed in triplicate. The temperature condition for qPCR was 95°C for 5 min; followed by 40 cycles consisting of 95°C for 15 s and 60°C for 1 min.

Cytogenetic Analysis

A G-banding karyotype analysis was also performed in all 27 fetal samples. Twenty metaphase cells were checked for numerical abnormalities of chromosomes, and five metaphase cells were carefully examined to detect structural chromosomal abnormalities.

Statistical Analysis

Categorical variables were summarized as number (percent) and were compared using chi-square test or Fisher's exact test. All analyses were performed using SAS statistical software (release 9.3, SAS Institute Inc., Cary, NC, USA), and P value < 0.05 was considered to be statistically significant.

RESULTS

Chromosomal Microarray Analysis and Karyotyping Result

Among the 16 cases diagnosed with 22q11.2 microdeletions, 10 of them were diagnosed with 22q11.2DS (proximal 22q11.2 region, LCR22-A to -D deletion, patient 1-10), and 6 others had deletions in the central 22q11.2 region 4 of LCR22-B to -D deletions (patients 11–14) and 2 of LCR22-C to -D deletions (patients 15–16). With respect to the 11 cases of 22q11.2 microduplications, 7 cases were detected with duplications in the proximal 22q11.2 region (LCR22-A to -D duplications, patient 17-23), 2 cases had duplications in the central 22q11.2 region (LCR22-B to -D duplications, patient 24-25), and 2 cases had duplications in the distal 22q11.2 region 1 of LCR22-F to -G duplications (patient 26), 1 of LCR22-E to -F duplications (patient 27). All the prenatal samples revealed normal karyotypes (**Table 2**, **Table S2,** and **Figure 1**).

Prenatal Ultrasound Findings

Compared with fetuses with 22q11.2 microduplications, fetuses with 22q11.2 microdeletions were more likely to present with structure defects in the ultrasound. As shown in **Table 2** and **Table S2**, all the 10 fetuses (100%) with 22q11.2DS had abnormal prenatal ultrasound findings, including congenital heart defect (n = 7), multiple congenital abnormalities (n = 2), and congenital renal agenesis (n = 1). Abnormal ultrasound findings were also observed in four of the six fetuses (66.7%) with central 22q11.2 deletion, including congenital heart defect (n = 2), congenital anomaly of nervous system (n = 1), and multiple congenital abnormalities (n = 1). In contrast, only one of the seven fetuses (14.3%) with proximal 22q11.2 microduplications was detected with defects by prenatal ultrasound scan. Moreover, three of the seven fetuses with proximal 22q11.2 microduplications showed

TABLE 2 | Characteristics of cases with prenatally diagnosed microduplications or microdeletions of 22q11.2 in this study.

	LCR	N	abnormal US[a], n (%)	*De novo*[b], n (%)	TOP, n (%)
22q11.2 microdeletion					
Proximal	A-B/D	10	10 (100)	8 (80)	10 (100)
Central	B/C-D	6	4 (66.7)	0 (0)[b]	4 (66.7)
22q11.2 microduplication					
Proximal	A-B/D	7	1 (14.3)	4 (57.1)[b]	4 (57.1)
Central	B/C-D	2	0 (0)	0 (0)	0 (0)
Distal type II+ III	E-F/F-G	2	0 (0)	0 (0)[b]	0 (0)

[a]*defined as structure defects.*
[b]*the inheritance was unknown in 6 fetuses, including 3 fetuses with central deletions, 2 fetuses with proximal duplications, and 1 fetus with distal duplications.*
US, ultrasound; TOP, termination of pregnancy.

increased nuchal translucency in the first trimester ultrasound (≥3.0 mm), and one of the two fetuses with distal 22q11.2 microduplications presented with echogenic bowel in the second trimester ultrasound. All the other cases were presented with normal ultrasound.

Inheritance

The parental samples were available in 21 cases, including 13 cases with microdeletions and 8 cases with microduplications (**Table 2** and **Table S2**). Eight of 10 cases (80%) with proximal deletions and 4 of 7 cases (57.1%) with proximal duplications occurred *de novo*. While, all other cases, including two fetuses with proximal deletions (one paternal and one maternal), one fetus with proximal duplications (paternal), three fetuses with central deletions (maternal), two fetuses with central duplications (one paternal and one maternal), and one fetus with proximal duplications (paternal) were inherited from parents. All the parents of the fetuses with inherited 22q11.2 microdeletions/microduplications were phenotypic normal, except the father of fetus 10 who presented with congenital atrial septal defect and mild developmental delay. In the remaining six cases, the inheritance could not be established, as the parents were not available for (or did not agree to) testing.

Pregnancy Outcome

Eighteen cases resulted in the induced termination of pregnancy (TOP), including 10 cases (100%) with proximal 22q11.2 deletions, 4 of 6 cases (66.7%) with 22q11.2 central deletions, and 4 of 7 cases (57.1%) with proximal 22q11.2 duplications (**Table 2**). The effect of abnormal ultrasound findings and the inheritance of the deletions or duplications on the rate of TOP were examined. As shown in **Table 3**, the rate of TOP in the cases with abnormal ultrasound was significantly higher than that in the cases presented with normal ultrasound (100 vs 25%, $P = 4 \times 10^{-5}$). Before the decision was made, the inheritance of the CNVs was available in only nine fetuses (**Table S2**). In these fetuses, the rate of TOP was associated with the inheritance of the deletions or duplications (100% in the *de novo* group vs 25% in the inherited group, $P = 0.02$) (**Table 3**).

To clarify the independent effect of the inheritance on the pregnancy outcome, the TOP rate was further analyzed in fetuses with normal ultrasound. We found that all the four fetuses with normal ultrasound and inherited from healthy parent were born normally, whereas all the two fetuses with normal ultrasound and occurred *de novo* resulted in induced TOP (**Table S2**).

TABLE 3 | Effect of ultrasound features, inheritance of the microdeletions, or microduplications on pregnancy outcome.

		N	TOP, n (%)	*P**
Ultrasound	Abnormal[$]	15	15 (100)	4×10^{-5}
	Normal	12	3 (25)	
Inheritance	*De novo*	4	4 (100)	0.02
	Inherited	5	1 (25)	

**chi-square test or Fisher's exact test.*
[$]*defined as structure defect.*
TOP, termination of pregnancy.

DISCUSSIONS

The application of CMA in prenatal diagnosis has greatly improved the detection rate of recurrent microdeletions and microduplications, which are common causes of congenital anomalies and neuropsychiatric disorders (Cooper et al., 2011; Grati et al., 2015). However, it is also accompanied by the detection of some CNVs with uncertain clinical significance, which may lead to great challenges in genetic counseling and parental anxiety. In this study, we reported on 27 new prenatally diagnosed cases of microdeletions or microduplications in the proximal (10 deletions and 7 duplications), central (6 deletions and 2 duplications), and distal (2 duplications) of chromosome 22q11.2, with particular attention being paid to the prenatal ultrasound findings and the pregnancy outcome of these fetuses.

Numerous studies about 22q11.2DS have been reported. Although the phenotypes of patients with 22q11.2DS are variable, the penetrance is nearly complete (McDonald-McGinn et al., 2015). In our study, all fetuses with 22q11.2DS presented with abnormal ultrasound findings. In contrast, only one of the seven fetuses (14.3%) with proximal 22q11.2 duplications was detected with structure defects, demonstrating the milder and highly variable phenotypes of proximal 22q11.2 duplications. The milder clinical phenotypes may contribute to the less

number of duplication cases reported in literature compared with 22q11.2DS, although they are complementary to each other and predicted to occur at the same frequency (Portnoi, 2009). Recently, a case-cohort study in Danish population found that the prevalence of 22q11.2 microduplications was 1 in 1,606, about twice of 22q11.2 microdeletions, demonstrating distinct selective pressures on these rearrangements (Olsen et al., 2018). Of note, increased nuchal translucency was detected in three fetuses with proximal 22q11.2 duplications, which was in accordance with the study conducted by Celine and coworkers (about 37% fetuses presented with increased nuchal translucency) (Dupont et al., 2015). Previous studies had demonstrated that more than 90% of 22q11.2DS occurred *de novo*, and this was confirmed in our study (McDonald-McGinn et al., 2001). Our study also found 57.1% of proximal 22q11.2 microduplications were *de novo*, which was different from the previous reports (Van Campenhout et al., 2012). One reason for the inconsistency may be the small sample size in our study.

Compared with the well-defined 22q11.2DS, the reports of fetuses with central or distal 22q11.2 microdeletions/microduplications are limited. Previously, the central 22q11.2 deletions were recognized as "atypical/nested deletions" of 22q11.2DS (Rump et al., 2014; Verhagen et al., 2012; Garcia-Minaur et al., 2002). However, the *TBX1* gene, which was considered to be the major candidate gene for the main features of 22q11.2DS, was not included in the central 22q11.2 region, and recent studies proposed that the central 22q11.2 deletions were distinct form the 22q11.2DS (Burnside, 2015; Rump et al., 2014). Compared with patients with 22q11.2DS, patients with central 22q11.2 deletions had a lower prevalence of congenital heart defects while nearly equal prevalences of renal and urogenital anomalies, developmental delays, cognitive impairments, and behavioral problems (Rump et al., 2014; Verhagen et al., 2012). The *CRKL* (*602007) gene was thought to be the candidate gene in the pathogenesis of the 22q11.2 central deletion (Lopez-Rivera et al., 2017; Haller et al., 2017; Breckpot et al., 2012). Six 22q11.2 central deletions were detected in our study, indicating that central 22q11.2 deletions were common recurrent CNVs (Burnside, 2015). However, the prevalence of 22q11.2 central deletions in the general population has not been reported. To date, very few cases with central 22q11.2 duplications have been reported (Pebrel-Richard et al., 2012; Fernandez et al., 2009). The phenotypes of individuals with central 22q11.2 duplications were variable, ranging from clinically normal to severe developmental delay with profound intellectual disability. One recent study supposed the gene *PI4K* as a candidate gene responsible for the neurodevelopmental phenotypes in individuals with central 22q11.2 duplications (Woodward et al., 2019). Compared with the 22q11.2DS, most central deletions/duplications were familial in the reported cases, and our study was consistent with this (Rump et al., 2014; Clements et al., 2017). However, as the reported cases are limited, the proportion of *de novo* occurrence of this CNV needs to be studied further.

The four distal LCRs, LCR22E-H, were smaller than LCR22A-D, and the distal 22q11.2 deletions and duplications resulted from NAHR of them were less common compared with the 22q11.2DS (Coppinger et al., 2009). It was confirmed in our study that no fetus with distal 22q11.2 deletions and only two fetuses with 22q11.2 distal duplications were detected. The distal 22q11.2 microduplications or microdeletions were enriched in clinical population (Coe et al., 2014), and the phenotypes of individuals with distal 22q11.2 microdeletions or microduplications were variable with incomplete penetrance (Wincent et al., 2010; Tan et al., 2011). Mikhail et al. suggested the recurrent distal 22q11.2 microdeletions do not represent a single clinical entity and proposed to categorize them into three types with unique clinical features and risks according to their genomic position (Mikhail et al., 2014). Most of cases reported as distal 22q11.2 CNVs were in the region D–E/F and can be classified into type I microdeletions or microduplications (Mikhail et al., 2014). In contrast, few cases with type II/III microdeletions or microduplications have been reported, and the pathogenicity of distal type II/III 22q11.2 microdeletions or microduplications is unclear, calling for more case reports. In addition, the causal genes responsible for the phenotypes of the 22q11.2 distal microdeletions and microduplications are still unknown.

In our hospital, counseling on the 22q11.2 microdeletions or microduplications was provided by a geneticist in prenatal diagnosis center. We observed that 66.7% of the parents decided to terminate the pregnancy. Those with structure defects in prenatal ultrasound or occurred *de novo* often led to TOP, whereas those with normal ultrasound and inherited from healthy parent were likely to continue the pregnancy and led to normal birth. However, as 22q11.2 deletions/duplications were associated with many neuropsychiatric disorders including developmental delay, long-term monitoring and follow-up of these carriers were necessary.

In conclusion, our results exhibited the extreme variability of the 22q11.2 recurrent microdeletions and microduplications. Compared with the fetuses with 22q11.2 microduplications, fetuses with 22q11.2 microdeletions were more likely to present with structure defects in the ultrasound. Both the prenatal ultrasound findings and the inheritance of the CNVs affected the parent's decision of pregnancy. Our study emphasized that proximal, central, and distal 22q11.2 deletions or duplications were different from each other, although some common features were shared among them. More studies are warranted to demonstrate the underlying mechanisms of different clinical features of these recurrent CNVs, thereby to provide more information for genetic counseling of 22q11.2 microdeletions and microduplications in prenatal diagnosis.

DATA AVAILABILITY

The CNV data for this study can be found in the ClinVar database (https://www.ncbi.nlm.nih.gov/clinvar/, RCV000788056-RCV000788073).

ETHICS STATEMENT

This study was carried out in accordance with the Declaration of Helsinki. The protocol was approved by the Ethics Committee

of the IPMCH, and written informed consent was obtained from the parents.

AUTHOR CONTRIBUTIONS

SL drafted and revised the manuscript, with further contributions from XC. SL, XH, SC and YS performed the experiment. SL, XH, and XC carried out data analysis. YW and JN collected clinical sample and interpretation of clinical data. All authors were involved in the design and conception of the study, in addition to reading and approving the final manuscript.

FUNDING

This work was supported by the National Natural Science Foundation of China (81871136, 81771638, 81501231), the Shanghai Municipal Commission of Science and Technology Program (16411963300), the Shanghai Municipal Health and Family Planning Committee (20164Y0212), the Interdisciplinary Program of Shanghai Jiao Tong University (YG2016MS40), the National Key Research and Development Program of China (2016YFC0905103), the International Peace Maternity and Child Health Hospital Clinical Research Project (GFY5817, GFY5818), and the Shanghai Municipal Key Clinical Specialty.

ACKNOWLEDGMENTS

We thank the families who participated in this study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2019.00813/full#supplementary-material

REFERENCES

Bailey, J. A., Yavor, A. M., Massa, H. F., Trask, B. J., and Eichler, E. E. (2001). Segmental duplications: organization and impact within the current human genome project assembly. *Genome Res.* 11, 1005–1017. doi: 10.1101/gr.GR-1871R

Bassett, A. S., McDonald-McGinn, D. M., Devriendt, K., Digilio, M. C., Goldenberg, P., Habel, A., et al. (2011). Practical guidelines for managing patients with 22q11.2 deletion syndrome. *J. Pediatr.* 159, 332–339 e1. doi: 10.1016/j.jpeds.2011.02.039

Ben-Shachar, S., Ou, Z., Shaw, C. A., Belmont, J. W., Patel, M. S., Hummel, M., et al. (2008). 22q11.2 distal deletion: a recurrent genomic disorder distinct from DiGeorge syndrome and velocardiofacial syndrome. *Am. J. Hum. Genet.* 82, 214–221. doi: 10.1016/j.ajhg.2007.09.014

Breckpot, J., Thienpont, B., Bauters, M., Tranchevent, L. C., Gewillig, M., Allegaert, K., et al. (2012). Congenital heart defects in a novel recurrent 22q11.2 deletion harboring the genes CRKL and MAPK1. *Am. J. Med. Genet. A* 158A, 574–580. doi: 10.1002/ajmg.a.35217

Burnside, R. D. (2015). 22q11.21 Deletion syndromes: a review of proximalcentral, and distal deletions and their associated features. *Cytogenet. Genome Res.* 146, 89–99. doi: 10.1159/000438708

Carlson, C., Sirotkin, H., Pandita, R., Goldberg, R., McKie, J., Wadey, R., et al. (1997). Molecular definition of 22q11 deletions in 151 velo-cardio-facial syndrome patients. *Am. J. Hum. Genet.* 61, 620–629. doi: 10.1086/515508

Clements, C. C., Wenger, T. L., Zoltowski, A. R., Bertollo, J. R., Miller, J. S., de Marchena, A. B., et al. (2017). Critical region within 22q11.2 linked to higher rate of autism spectrum disorder. *Mol. Autism* 8, 58. doi: 10.1186/s13229-017-0171-7

Coe, B. P., Witherspoon, K., Rosenfeld, J. A., van Bon, B. W., Vulto-van Silfhout, A. T., Bosco, P., et al. (2014). Refining analyses of copy number variation identifies specific genes associated with developmental delay. *Nat. Genet.* 46, 1063–1071. doi: 10.1038/ng.3092

Cooper, G. M., Coe, B. P., Girirajan, S., Rosenfeld, J. A., Vu, T. H., Baker, C., et al. (2011). A copy number variation morbidity map of developmental delay. *Nat. Genet.* 43, 838–846. doi: 10.1038/ng.909

Coppinger, J., McDonald-McGinn, D., Zackai, E., Shane, K., Atkin, J. F., Asamoah, A., et al. (2009). Identification of familial and de novo microduplications of 22q11.21-q11.23 distal to the 22q11.21 microdeletion syndrome region. *Hum. Mol. Genet.* 18, 1377–1383. doi: 10.1093/hmg/ddp042

Dittwald, P., Gambin, T., Gonzaga-Jauregui, C., Carvalho, C. M., Lupski, J. R., Stankiewicz, P., et al. (2013). Inverted low-copy repeats and genome instability–a genome-wide analysis. *Hum. Mutat.* 34, 210–220. doi: 10.1002/humu.22217

Dupont, C., Grati, F. R., Choy, K. W., Jaillard, S., Toutain, J., Maurin, M. L., et al. (2015). Prenatal diagnosis of 24 cases of microduplication 22q11.2: an investigation of phenotype-genotype correlations. *Prenat. Diagn.* 35, 35–43. doi: 10.1002/pd.4478

Ensenauer, R. E., Adeyinka, A., Flynn, H. C., Michels, V. V., Lindor, N. M., Dawson, D. B., et al. (2003). Microduplication 22q11.2, an emerging syndrome: clinical, cytogenetic, and molecular analysis of thirteen patients. *Am. J. Hum. Genet.* 73, 1027–1040. doi: 10.1086/378818

Fernandez, L., Nevado, J., Santos, F., Heine-Suner, D., Martinez-Glez, V., Garcia-Minaur, S., et al. (2009). A deletion and a duplication in distal 22q11.2 deletion syndrome region. Clinical implications and review. *BMC Med. Genet.* 10, 48. doi: 10.1186/1471-2350-10-48

Garcia-Minaur, S., Fantes, J., Murray, R. S., Porteous, M. E., Strain, L., Burns, J. E., et al. (2002). A novel atypical 22q11.2 distal deletion in father and son. *J. Med. Genet.* 39, E62. doi: 10.1136/jmg.39.10.e62

Grati, F. R., Molina Gomes, D., Ferreira, J. C., Dupont, C., Alesi, V., Gouas, L., et al. (2015). Prevalence of recurrent pathogenic microdeletions and microduplications in over 9500 pregnancies. *Prenat. Diagn.* 35, 801–809. doi: 10.1002/pd.4613

Haller, M., Mo, Q. X., Imamoto, A., and Lamb, D. J. (2017). Murine model indicates 22q11.2 signaling adaptor CRKL is a dosage-sensitive regulator of genitourinary development. *P. Natl. Acad. Sci. U.S.A.* 114, 4981–4986. doi: 10.1073/pnas.1619523114

Hassed, S. J., Hopcus-Niccum, D., Zhang, L., Li, S., and Mulvihill, J. J. (2004). A new genomic duplication syndrome complementary to the velocardiofacial (22q11 deletion) syndrome. *Clin. Genet.* 65, 400–404. doi: 10.1111/j.0009-9163.2004.0212.x

Li, S., Han, X., Wang, Y., Chen, S., Niu, J., Qian, Z., et al. (2019). Chromosomal microarray analysis in fetuses with congenital anomalies of the kidney and urinary tract: a prospective cohort study and meta-analysis. *Prenat. Diagn.* 39, 165–174. doi: 10.1002/pd.5420

Lopez-Rivera, E., Liu, Y. P., Verbitsky, M., Anderson, B. R., Capone, V. P., Otto, E. A., et al. (2017). Genetic drivers of kidney defects in the digeorge syndrome. *N. Engl. J. Med.* 376, 742–754. doi: 10.1056/NEJMoa1609009

Mademont-Soler, I., Morales, C., Soler, A., Martinez-Crespo, J. M., Shen, Y., Margarit, E., et al. (2013). Prenatal diagnosis of chromosomal abnormalities in fetuses with abnormal cardiac ultrasound findings: evaluation of chromosomal microarray-based analysis. *Ultrasound Obstet. Gynecol.* 41, 375–382. doi: 10.1002/uog.12372

McDonald-McGinn, D. M., Sullivan, K. E., Marino, B., Philip, N., Swillen, A., Vorstman, J. A., et al. (2015). 22q11.2 deletion syndrome. *Nat. Rev. Dis. Primers* 1, 15072.

McDonald-McGinn, D. M., Tonnesen, M. K., Laufer-Cahana, A., Finucane, B., Driscoll, D. A., Emanuel, B. S., et al. (2001). Phenotype of the 22q11.2 deletion

in individuals identified through an affected relative: Cast a wide FISHing net! *Genet. Med.* 3, 23–29. doi: 10.1097/00125817-200101000-00006

Mikhail, F. M., Burnside, R. D., Rush, B., Ibrahim, J., Godshalk, R., Rutledge, S. L., et al. (2014). The recurrent distal 22q11.2 microdeletions are often de novo and do not represent a single clinical entity: a proposed categorization system. *Genet. Med.* 16, 92–100. doi: 10.1038/gim.2013.79

American College of Obstetricians and Gynecologists Committee on Genetics (2013). Committee Opinion No. 581: the use of chromosomal microarray analysis in prenatal diagnosis. *Obstet. Gynecol.* 122, 1374–1377. doi: 10.1097/01.AOG.0000438962.16108.d1

Olsen, L., Sparso, T., Weinsheimer, S. M., Dos Santos, M. B. Q., Mazin, W., Rosengren, A., et al. (2018). Prevalence of rearrangements in the 22q11.2 region and population-based risk of neuropsychiatric and developmental disorders in a Danish population: a case-cohort study. *Lancet Psychiatry* 5, 573–580. doi: 10.1016/S2215-0366(18)30168-8

Pebrel-Richard, C., Kemeny, S., Gouas, L., Eymard-Pierre, E., Blanc, N., Francannet, C., et al. (2012). An atypical 0.8 Mb inherited duplication of 22q11.2 associated with psychomotor impairment. *Eur. J. Med. Genet.* 55, 650–655. doi: 10.1016/j.ejmg.2012.06.014

Portnoi, M. F. (2009). Microduplication 22q11.2: a new chromosomal syndrome. *Eur. J. Med. Genet.* 52, 88–93. doi: 10.1016/j.ejmg.2009.02.008

Rump, P., de Leeuw, N., van Essen, A. J., Verschuuren-Bemelmans, C. C., Veenstra-Knol, H. E., Swinkels, M. E., et al. (2014). Central 22q11.2 deletions. *Am. J. Med. Genet. A* 164A, 2707–2723. doi: 10.1002/ajmg.a.36711

Shaikh, T. H., O'Connor, R. J., Pierpont, M. E., McGrath, J., Hacker, A. M., Nimmakayalu, M., et al. (2007). Low copy repeats mediate distal chromosome 22q11.2 deletions: sequence analysis predicts breakpoint mechanisms. *Genome Res.* 17, 482–491. doi: 10.1101/gr.5986507

Tan, T. Y., Collins, A., James, P. A., McGillivray, G., Stark, Z., Gordon, C. T., et al. (2011). Phenotypic variability of distal 22q11.2 copy number abnormalities. *Am. J. Med. Genet. A* 155A, 1623–1633. doi: 10.1002/ajmg.a.34051

Van Campenhout, S., Devriendt, K., Breckpot, J., Frijns, J. P., Peeters, H., Van Buggenhout, G., et al. (2012). Microduplication 22q11.2: a description of the clinical, developmental and behavioral characteristics during childhood. *Genet. Counsel* 23, 135–148.

Verhagen, J. M., Diderich, K. E., Oudesluijs, G., Mancini, G. M., Eggink, A. J., Verkleij-Hagoort, A. C., et al. (2012). Phenotypic variability of atypical 22q11.2 deletions not including TBX1. *Am. J. Med. Genet. A* 158A, 2412–2420. doi: 10.1002/ajmg.a.35517

Wentzel, C., Fernstrom, M., Ohrner, Y., Anneren, G., and Thuresson, A. C. (2008). Clinical variability of the 22q11.2 duplication syndrome. *Eur. J. Med. Genet.* 51, 501–510. doi: 10.1016/j.ejmg.2008.07.005

Wincent, J., Bruno, D. L., van Bon, B. W., Bremer, A., Stewart, H., Bongers, E. M., et al. (2010). Sixteen new cases contributing to the characterization of patients with distal 22q11.2 microduplications. *Mol. Syndromol.* 1, 246–254. doi: 10.1159/000327982

Woodward, K. J., Stampalia, J., Vanyai, H., Rijhumal, H., Potts, K., Taylor, F., et al. (2019). Atypical nested 22q11.2 duplications between LCR22B and LCR22D are associated with neurodevelopmental phenotypes including autism spectrum disorder with incomplete penetrance. *Mol. Genet. Genomic Med.* 7, e00507 doi: 10.1002/mgg3.507.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Genetics

CASE REPORT
published: 11 September 2019
doi: 10.3389/fgene.2019.00817

A Novel Silent Mutation in the *L1CAM* Gene Causing Fetal Hydrocephalus Detected by Whole-Exome Sequencing

*Yixi Sun[1,2,3], Yanfeng Li[1,2,3], Min Chen[1,2,3] Yuqin Luo[1,2,3], Yeqing Qian[1,2,3], Yanmei Yang[1,2,3], Hong Lu[4], Fenlan Lou[5] and Minyue Dong[1,2,3]**

[1] Department of Reproductive Genetics, Women's Hospital, School of Medicine, Zhejiang University, Hangzhou, China, [2] Key Laboratory of Reproductive Genetics, Ministry of Education (Zhejiang University), Hangzhou, China, [3] Key Laboratory of Women's Reproductive Health of Zhejiang Province, Hangzhou, China, [4] Department of Ultrasound, Women's Hospital, School of Medicine, Zhejiang University, Hangzhou, China, [5] Department of Diagnostic Radiology, Women's Hospital, School of Medicine, Zhejiang University, Hangzhou, China

OPEN ACCESS

Edited by:
Yueqiu Tan, Central South University, China

Reviewed by:
Liang-Liang Fan, Central South University, China
Na Zhu, Columbia University, United States

***Correspondence:**
Minyue Dong
dongmy@zju.edu.cn

Specialty section:
This article was submitted to Genetic Disorders, a section of the journal Frontiers in Genetics

Received: *08 June 2019*
Accepted: *07 August 2019*
Published: *11 September 2019*

Citation:
Sun Y, Li Y, Chen M, Luo Y, Qian Y, Yang Y, Lu H, Lou F and Dong M (2019) A Novel Silent Mutation in the L1CAM Gene Causing Fetal Hydrocephalus Detected by Whole-Exome Sequencing. Front. Genet. 10:817. doi: 10.3389/fgene.2019.00817

X-linked hydrocephalus (XLH), a genetic disorder, has an incidence of 1/30,000 male births. The great proportion of XLH is ascribed to loss-of-function mutations of L1 cell adhesion molecule gene (*L1CAM*), but silent mutations in *L1CAM* with pathogenic potential were rare and were usually ignored especially in whole-exome sequencing (WES) detection. In the present study, we describe a novel silent *L1CAM* mutation in a Chinese pregnant woman reporting continuous five times pregnancies with fetal hydrocephalus. After fetal blood sampling, we found c.453G > T (p.Gly151 =) in the *L1CAM* gene of the fetus by WES; RT-PCR of the messenger RNA (mRNA) from cord blood mononuclear cells and subsequent sequence analysis identified the mutation created a potential 5′ splice site consensus sequence, which would result in an in-frame deletion of 72 bp from exon 5 and 24 amino acids of the *L1CAM* protein. Heterozygous mutations were confirmed in analyzing DNA and mRNA from peripheral blood mononuclear cells of the woman, and a severe L1 syndrome was confirmed by fetal ultrasound scan and MRI. Our study first indicated c.453G > T (p.Gly151 =) in *L1CAM* could be disease causing for hydrocephalus, which would aid in genetic counseling for the prenatal diagnosis of hydrocephalus. Meanwhile, it suggested some silent mutations detected in WES should not be ignored; splicing predictions of these mutations were necessary.

Keywords: hydrocephalus, *L1CAM*, whole-exome sequencing, silent mutation, splicing mutation

INTRODUCTION

The L1 cell adhesion molecule gene (*L1CAM*) is a neuronal cell adhesion molecule belonging to the immunoglobulin superfamily; it possesses key functions in the development of the nervous system (Itoh and Fushiki, 2015). Mutations in *L1CAM* have been related to X-linked neurological syndromes, which are summarized as L1 diseases. They are classified as follows: X-linked hydrocephalus (XLH) due to stenosis of the aqueduct of Sylvius (HSAS), MASA syndrome (intellectual disability, aphasia, shuffling gait, adducted thumbs), spastic paraparesis type 1 (SP1), and X-linked agenesis of corpus callosum (ACC) (Weller and Gartner, 2001; Itoh and Fushiki, 2015).

About 282 disease-causing mutations (DMs) in the *L1CAM* gene have been reported in HGMD® Professional 2019.2 (https://portal.biobase-international.com/hgmd/pro/all.php). Alterations in the *L1CAM* gene are varied; mutation data analyses from 282 patients disclose 51% missense and nonsense mutations, 25% deletions, 5% insertions, and 19% splice site changes, but silent mutations in *L1CAM* with pathogenic potential were rare, and silent mutations were often ignored especially in whole-exome sequencing (WES) detection.

In this study, using WES, we screened the fetal DNA of a Chinese pregnant woman who has reported five continuous pregnancies with fetal hydrocephalus; we only found a novel silent mutation c.453G > T (p.Gly151 =) in the *L1CAM* gene. Interestingly, through further analysis, we indicated the silent mutation created a potential 5′ splice site consensus sequence, which would result in an in-frame deletion of 72 bp from exon 5 and 24 amino acids of the *L1CAM* protein.

CASE PRESENTATION

A 28-year-old healthy woman was referred to our clinic after four voluntary terminations of pregnancy due to fetal hydrocephalus at other hospitals. All fetuses were male. When arriving at our hospital (Women's Hospital, School of Medicine, Zhejiang University, Zhejiang, China), she was already on her fifth pregnancy at 24 weeks of gestation, with a fetal hydrocephalus by image examinations. To explore the genetic cause, fetal blood sampling was conducted at 26 weeks of gestational age. Conventional cytogenetic studies were performed for both fetal and parental samples, and the fetal sample was further analyzed by single-nucleotide polymorphism (SNP) array and WES.

This study was carried out in accordance with the recommendations of the Ethics Committee of Women's Hospital, School of Medicine Zhejiang University, and informed consent was acquired from all the participants of this study in accordance with the Declaration of Helsinki. The study protocol was approved by the Review Board of the Women's Hospital, School of Medicine, Zhejiang University in China.

MATERIALS AND METHODS

Karyotype and SNP Array

The karyotypes of fetal cord blood and peripheral cord blood were determined by conventional karyotyping of at least 30 blood lymphocytes, which were arrested at metaphase by colchicines. G-banding karyotypes of cultured cells were performed at the 320–400-band level with a resolution of around 10 Mb. SNP array was performed by the CytoScan™ HD array (Affymetrix, USA) according to the manufacturer's instruction, with around 2,600,000 markers including 750,000 SNP probes and 1,900,000 non-polymorphism probes for comprehensive whole-genome coverage. Data were analyzed by the Chromosome Analysis Suite (ChAS) software (Affymetrix, Santa Clara, CA) based on the GRCh37/hg19 assembly. The reporting threshold of the copy number result was set at 500 kb with a marker count of ≥50 for gains and at 200 kb with a marker count of ≥50 for losses.

Whole-Exome Sequencing

The main part of WES was provided by the Beijing Genomics Institute. Genomic DNA was extracted by a DNeasy Blood Kit (Qiagen, CA) and then was fragmented by Covaris LE220 (Massachusetts, USA) to generate a paired-end library (200–250 bp). All amplified libraries were performed on the BGISEQ-500 platform, the single-strand DNA was mixed with MGIEasy™ DNA Library Prep Kit V1 (BGI, Shenzhen, China) and then sequenced using 100SR chemistry with BGISEQ-500RS high-throughput sequencing Kit (BGI, Shenzhen, China).

Clean reads (with a length of 90 bp) derived from targeted sequencing and filtering were then aligned to the human genome reference (hg19) using the Burrows-Wheeler Aligner (BWA) Multi-Vision software package (Li and Durbin, 2009). After alignment, the output files were used to perform sequencing coverage and depth analysis of the target region, single-nucleotide variants (SNVs), and indel calling, we used the GATK software to detect SNVs and indels (McKenna et al., 2010), all SNVs and indels were filtered and estimated *via* multiple databases, including the National Center for Biotechnology Information (NCBI) Single-Nucleotide Polymorphism Database (dbSNP), HapMap, 1000 Genomes Project dataset, and database of 100 Chinese healthy adults. We used Condel, SIFT, PolyPhen-2, LRT, Mutation Taster, and PhyloP to predict the effect of variants. Pathogenic variants are assessed under the protocol issued by the American College of Medical Genetics and Genomics (ACMG) (Richards et al., 2015). The Human Gene Mutation Database (HGMD) was used to screen mutations. All potential pathogenic variants were validated using Sanger sequencing methods.

RNA Extraction, PCR, and Sequencing

Peripheral blood mononuclear cells (PMBCs) and cord blood mononuclear cells (CBMCs) were isolated by Ficoll density gradient separation. Total RNA was extracted from PMBCs and the CBMCs using TRIzol (Takara, Japan). Extracted total RNAs were reverse-transcribed using RT Kit (Takara, Japan). PCR was performed using GoldStar Best Master Mix (CWBIO, Beijing). Primer sequences are listed: L1CAM-DNA-5F, CCCACCCGTC CTTTCCTA; L1CAM-DNA-5R, CGCTCGTCCTGCTTGATGT; L1CAM-mRNA-4-6-F, GGTGTCCACTTCAAACCCAA; and L1CAM-mRNA-4-6-R, GCGGCTTCCTGTCAATCA. Sanger sequencing was performed by an ABI 3130 DNA analyzer.

RESULTS

A 28-year-old healthy woman was referred to our clinic after four voluntary terminations of pregnancy due to fetal hydrocephalus. All fetuses were male (**Figure 1A**). The familial pedigree seemed to show XLH. She was already on here fifth pregnancy at 26 weeks of gestation. Fetal ventriculomegaly was detected by fetal ultrasound

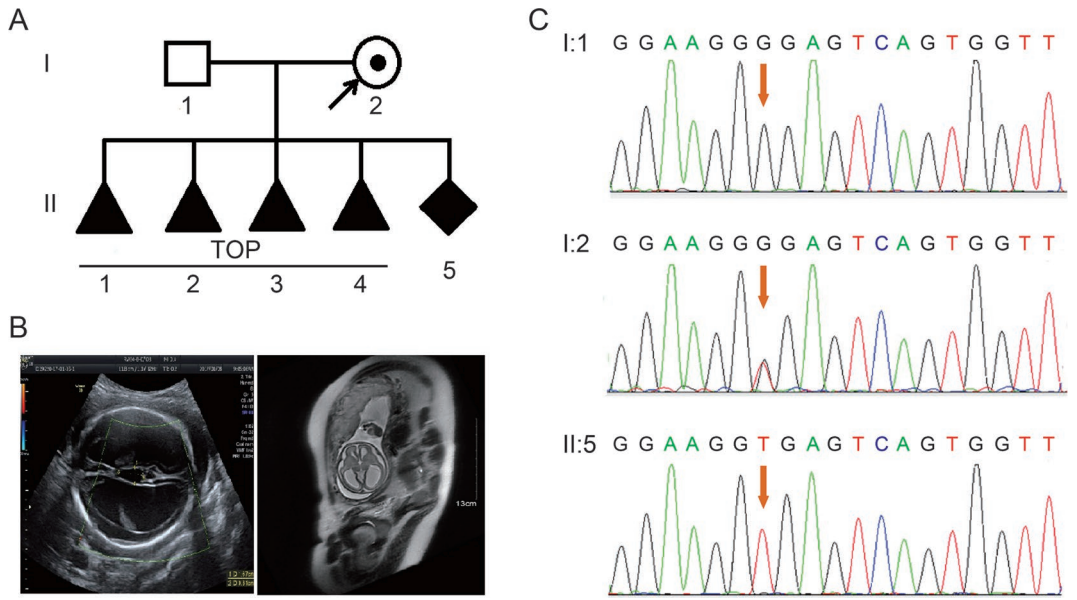

FIGURE 1 | (A) Pedigree of the family. TOP, termination of pregnancy. **(B)** Imaging examinations of the fetus. Fetal ultrasound scan and fetal MRI showed there was severe hydrocephalus in the fetus. **(C)** Sequence analysis of genomic DNA from family members. The genotypes of *L1CAM* were wild type, c.453G > T Het, and c.453G > T Hom, in I:1 (husband), I:2 (pregnant woman), and II:5 (fetus). The mutation is indicated by the red arrows.

scan and MRI, which consistently demonstrated the presence of hydrocephalus. They showed that the bilateral cerebral ventricle and the third ventricle were obviously dilated, and there was severe hydrocephalus in the intracerebral and agenesis of the corpus callosum (**Figure 1B**).

In order to explore the possible genetic cause, we performed karyotype analysis and SNP array to analyze the fetal blood sampling and found no positive findings. Choroidal neovascularization (CNV) result has been deposited in the Gene Expression Omnibus (GEO); the accession number is GSE133063, as appended below (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE133063).

Then, we detected the fetus by WES. The analytic strategy for finding likely pathogenic variant identification was shown in **Figure S1**. A list of variants (**Table S1**) were obtained through the screening of variant frequencies, mutation status, and inheritance mode. Taking the hydrocephalus-associated genes (HP:0000238, http://compbio.charite.de/hpoweb/showterm?id=HP:0000238#id=HP_0000238) (**Table S2**) into consideration, there was no additional notable mutation except for the silent mutation of c.453G > T in exon 5 of the *L1CAM* gene (NM_000425.3). c.453G > T was not reported in HGMD and ClinVar and was not found in dbSNP, gnomAD, and other datasets. According to the standards and guidelines of the ACMG (Richards et al., 2015), it had not yet reached the criterion of "pathogenic" or "likely pathogenic," but there was no other potential mutations; we had no choice but to make a further analysis of the silent mutation found.

According to traditional thinking, this base substitution occurred in the third base in codon 151, which encodes a glycine, thereby creating a neutral mutation (p.Gly151 =). This variant was confirmed in DNA extracted from fetal cord blood and peripheral blood in the couple by Sanger sequencing (**Figure 1C**). The woman carried the heterozygous mutation, and her husband was a wild-type genotype.

With Mutation Taster (http://www.mutationtaster.org/), c.453G > T was scored as "disease causing." It showed that protein features might be affected and the splice site might be changed; we were curious about the potential splicing effects of the *L1CAM* function in this silent mutation. The silent mutation was tested using the following online software products: NetGene2 (http://www.cbs.dtu.dk/services/NetGene2/) and NNSplice (http://www.fruitfly.org/seq_tools/splice.html); the 5′ potential splice site was also predicted to be created in the *L1CAM* c.453G > T mutation using the software products (**Figure S2**). The results showed that this silent mutation created a potential 5′ splice site 72 bp upstream from the normal exon 6/intron 6 splice site (**Figure 2A**). If this is the case, we can find the length change of *L1CAM* messenger RNA (mRNA) between I:2 and II:5 (**Figure 2A**). RT-PCR was performed using primers designed to amplify exons 4–6 in *L1CAM* mRNA. Indeed, the results showed a short band of truncation in fetal cDNA PCR (II:5), while the band amplified from *L1CAM* mRNA contained the expected long band in husband cDNA PCR (I:1) and long/short bands in the pregnant woman cDNA PCR (I:2) (**Figure 2B**). Direct sequencing of the amplified fragment showed that the deletion involved the last 72 bp of exon 5 in male fetal cDNA (the woman was a carrier) (**Figure 2C**). We got the crucial pathogenic evidence.

This silent mutation resulted in 24 amino acids of L1CAM protein (residues 151–174); Lys (K) was substituted by Glu (E) at codon 175 (**Figure 2A**). There were alignment of multiple L1CAM protein sequences across several species and conservation of the

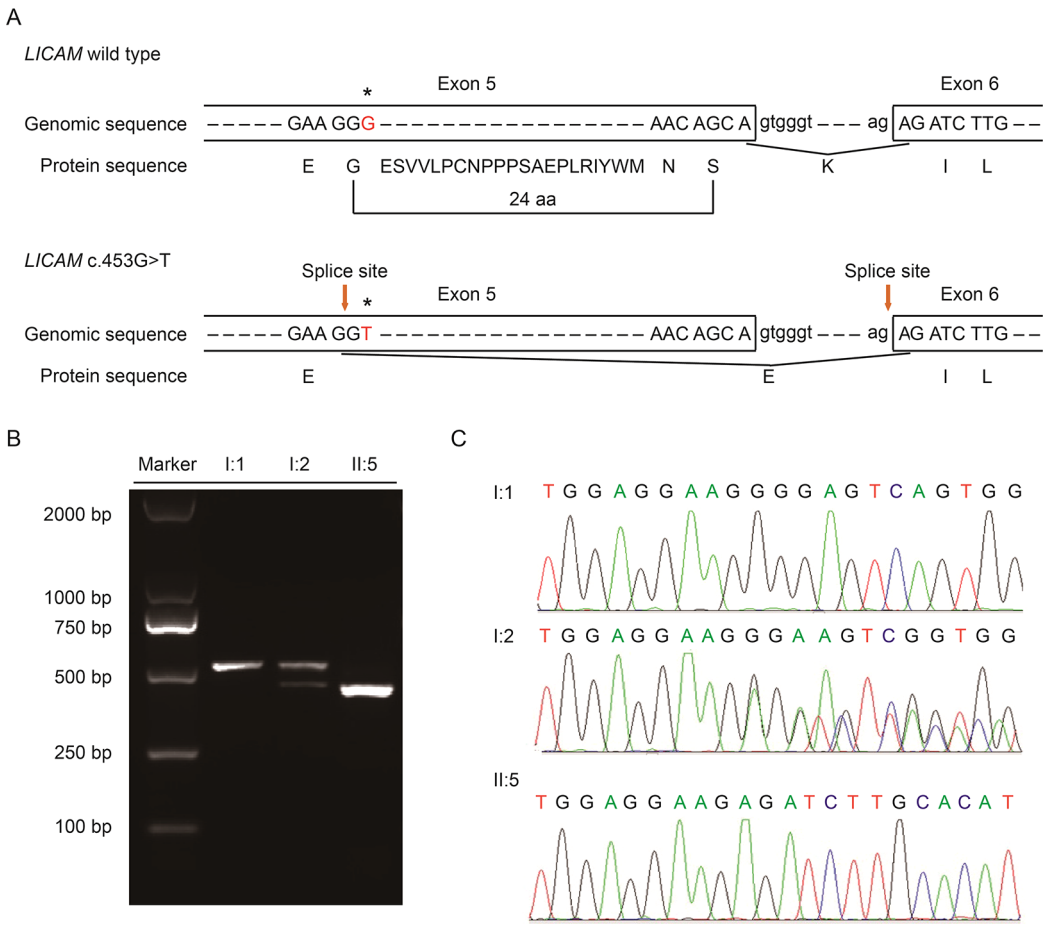

FIGURE 2 | (A) Schematic representation of exon 5, intron 6, and exon 6 organization in *L1CAM*. **(B)** RT-PCR analysis of exons 5 and 6 of the *L1CAM* cDNA from peripheral blood mononuclear cells (PMBCs) and cord blood mononuclear cells (CBMCs). Agarose gel electrophoresis of RT-PCR products generated from I:1 (husband), I:2 (pregnant woman), and II:5 (fetus). **(C)** Sequence analysis of the RT-PCR product from PMBCs of the couple and CBMCs of the fetus.

missing amino acids in L1CAM across mammals: *Homo sapiens*, *Pan troglodytes*, *Bos taurus*, *Mus musculus*, and *Rattus norvegicus* (**Figure 3A**). Wild-type and c.453G > T splicing mutation L1CAM proteins were predicted by the software CPHmodels-3.2 Server (http://www.cbs.dtu.dk/services/CPHmodels/) (**Figure 3B**). Immunoglobulin-like (Ig-like) domain 2 (residues 134–230) of wild-type and splicing mutation L1CAM proteins is shown in **Figure 3C**. *L1CAM* c.453G > T splicing mutation altered the protein structure, especially the Ig-like domain 2.

DISCUSSION

Silent mutations were often detected by WES, but insufficient attention has been paid, leading to the omission of DMs. In this study, we employed WES to explore the genetic cause of a Chinese family with hydrocephalus but only found a novel silent mutation in *L1CAM*, which forced us to make a further analysis. Fortunately, we proved that the silent mutation created a new 5' splice site and was a DM.

Mutations in *L1CAM* can cause an X-linked L1 disease, but clinical symptoms are variable; mutations produce unexpected phenotypes. In the study, the five suffering fetuses are all males, which is consistent with an inheritance pattern. The fetal ultrasound scan and MRI show a typical L1 disease, including XLH and agenesis of the corpus callosum. It improves our understanding on the genotype–phenotype correlation of L1CAM.

L1CAM c.453G > T (p.Gly151 =) was initially thought to have no effect on the protein sequence. But other silent mutations, c.924C > T (p.Gly308 =) and c.645C > T (p.Gly215 =), in the *L1CAM* gene have been reported to be DMs (Du et al., 1998; Vos et al., 2010). The c.924C > T mutation resulted in the activation of a new splice site 69 bp 5′ to the normal exon 8/intron 8 donor splice site, and it has been declared as a "disease-causing" site for hydrocephalus (Du et al., 1998). For c.645C > T in *L1CAM*, 51 bp was deleted with the activation of a new exon 6/intron 6 donor splice (Vos et al., 2010). Our present study was similar; the mutation of c.453G > T created a potential 5′ splice site upstream from the normal exon 5/intron 5 splice site. All these silent mutations created new donor splice sites, resulting in the exon being skipped. It reminded us to pay close attention to these silent mutations, which may affect splicing of proteins.

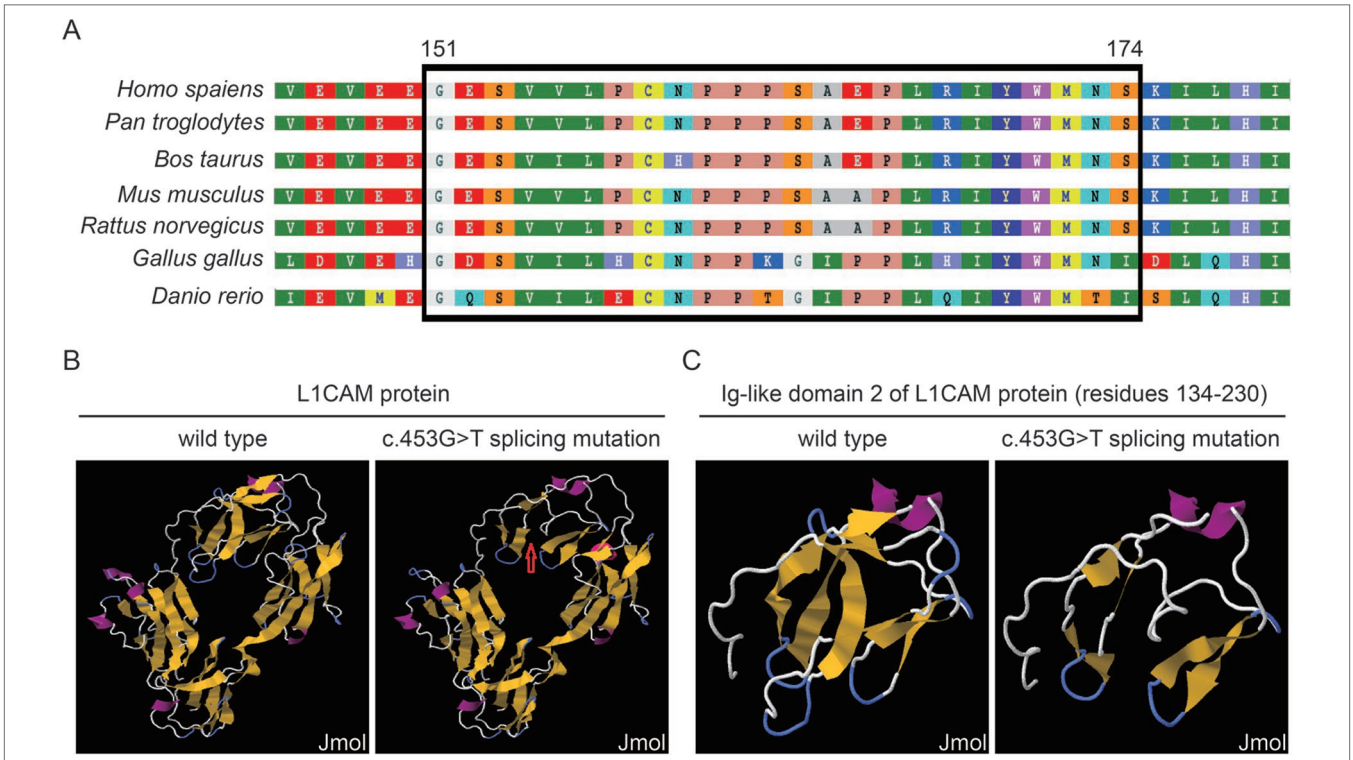

FIGURE 3 | (A) Alignment of multiple L1CAM protein sequences across species. The *L1CAM* c.453G > T resulted in 24 amino acids of L1CAM protein (residues 151–174) missing in the conserved amino acid region in different species. The black column shows the missing amino acids. **(B)** The structures of wild-type and c.453G > T splicing mutation L1CAM protein as predicted by the software CPHmodels-3.2 Server. **(C)** The structures of immunoglobulin-like (Ig-like) domain 2 (residues 134–230) of wild-type and splicing mutation L1CAM protein.

As a transmembrane glycoprotein and a member of the immunoglobulin superfamily of cell adhesion molecules, the L1CAM protein can interact at the cell surface with a number of different glycoproteins, and homophilic binding is probably its principal mode of interaction (Wei and Ryu, 2012). The studies on the crystal structure of Ig-like domains 1–4 in neurofascin suggested that many pathological L1 mutations affect conserved amino acid residues within these domains and interfere with homophilic interactions (Liu et al., 2011), especially as verified by the function research of Ig-like domain 2 (Zhao et al., 1998). In our study, we speculated that *L1CAM* c.453G > T altered Ig-like domain 2 in the extracellular part of the L1CAM protein, leading to the abnormal extracellular interaction, failing to start initiating downstream the signaling pathway. A further study dedicated to the mass spectrometry of this L1CAM variant would clarify specifically what molecular ensemble is produced in the cell.

In summary, through WES, we reported a novel silent mutation c.453G > T in *L1CAM* which produces a 5′ splice site responsible for hydrocephalus. This abnormal protein variant was predicted to alter Ig-like domain 2, which might affect L1CAM protein homophilic binding. In addition, we performed prenatal genetic diagnosis for the pregnant woman reporting five continuous pregnancies with hydrocephalus. Meanwhile, it suggested some silent mutations detected in WES should not be ignored; splicing predictions of these mutations were necessary. It provided a new genetic basis for prenatal diagnosis and pre-implantation prenatal diagnosis of hydrocephalus.

DATA AVAILABILITY

Publicly available datasets were analyzed in this study. This data can be found here: GSE133063(https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE133063).

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Review Board of the Women's Hospital, School of Medicine, Zhejiang University in China. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin. Written informed consent was obtained from the individual(s), and minor(s)' legal guardian/next of kin, for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

YS, YLi, MC, YLu, YQ, and YY conducted experiments. YS prepared the figures. MC and YLi analyzed the WES data. YLu and YQ performed karyotype analysis and SNP array. YY recruited samples. HL and FL provided imaging examinations. YS and MD wrote the manuscript. All authors read and approved the final manuscript.

FUNDING

This study was supported by the National Natural Science Foundation of China (Grant No. 81801441), the Key Research and Development Program of the Zhejiang province (Grant No. 2019C03025), the National Key Research and Development Program of China (Grant No. 2016YFC1000703), and the Medical Scientific Research Foundation of Zhejiang Province (Grant No. 2014KYA246).

ACKNOWLEDGMENTS

We thank the patients enrolled in this research. We thank Dr. Jiong Gao (BGI Genomics, BGI-Shenzhen, Shenzhen 518083, China) for his assistance during the preparation of this manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2019.00817/full#supplementary-material

FIGURE S1 | Analytic strategy for finding likely pathogenic variant identification by WES.

FIGURE S2 | Donor splice sites predicted by NetGene2 and NNSplice.

TABLE S1 | List of variants through the screening of variants frequencies, mutation status, and inheritance mode.

TABLE S2 | List of hydrocephalus-associated genes (Export for HP:0000238).

REFERENCES

Du, Y. Z., Dickerson, C., Aylsworth, A. S., and Schwartz, C. E. (1998). A silent mutation, C924T (G308G), in the L1CAM gene results in X linked hydrocephalus (HSAS). *J. Med. Genet.* 35 (6), 456–462. doi: 10.1136/jmg.35.6.456

Itoh, K., and Fushiki, S. (2015). The role of L1cam in murine corticogenesis, and the pathogenesis of hydrocephalus. *Pathol. Int.* 65 (2), 58–66. doi: 10.1111/pin.12245

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25 (14), 1754–1760. doi: 10.1093/bioinformatics/btp324

Liu, H., Focia, P. J., and He, X. (2011). Homophilic adhesion mechanism of neurofascin, a member of the L1 family of neural cell adhesion molecules. *J. Biol. Chem.* 286 (1), 797–805. doi: 10.1074/jbc.M110.180281

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., et al. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20 (9), 1297–1303. doi: 10.1101/gr.107524.110

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., et al. (2015). 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.* 17 (5), 405–424. doi: 10.1038/gim.2015.30

Vos, Y. J., de Walle, H. E. K., Bos, K. K., Stegeman, J. A., ten Berge, A. M., Bruining, M., et al. (2010). Genotype-phenotype correlations in L1 syndrome: a guide for genetic counselling and mutation analysis. *J.Med. Genet.* 47 (3), 169–175. doi: 10.1136/jmg.2009.071688

Wei, C. H., and Ryu, S. E. (2012). Homophilic interaction of the L1 family of cell adhesion molecules. *Exp. Mol. Med.* 44 (7), 413–423. doi: 10.3858/emm.2012.44.7.050

Weller, S., and Gartner, J. (2001). Genetic and clinical aspects of X-linked hydrocephalus (L1 disease): mutations in the L1CAM gene. *Hum. Mutat.* 18 (1), 1–12. doi: 10.1002/humu.1144

Zhao, X., Yip, P. M., and Siu, C. H. (1998). Identification of a homophilic binding site in immunoglobulin-like domain 2 of the cell adhesion molecule L1. *J. Neurochem.* 71 (3), 960–971 doi: 10.1046/j.1471-4159.1998.71030960.x.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Genetics

CASE REPORT
published: 13 September 2019
doi: 10.3389/fgene.2019.00835

Compound Heterozygosity for Novel Truncating Variants in the *LMOD3* Gene as the Cause of Polyhydramnios in Two Successive Fetuses

Ye Wang[1], Caixia Zhu[1], Liu Du[2], Qiaoer Li[3], Mei-Fang Lin[2], Claude Férec[4,5], David N. Cooper[6], Jian-Min Chen[4†] and Yi Zhou[1†]*

[1] Fetal Medicine Center, Department of Obstetrics and Gynecology, The First Affiliated Hospital of Sun Yat-Sen University, Guangzhou, China, [2] Department of Ultrasonic Medicine, The First Affiliated Hospital of Sun Yat-Sen University, Guangzhou, China, [3] Jiangmen Central Hospital, Affiliated Jiangmen Hospital of Sun Yat-Sen University, Jiangmen, China, [4] EFS, Univ Brest, Inserm, UMR 1078, GGB, Brest, France, [5] CHU Brest, Service de Génétique, Brest, France, [6] Institute of Medical Genetics, School of Medicine, Cardiff University, Cardiff, United Kingdom

OPEN ACCESS

Edited by:
Fan Jin,
Zhejiang University, China

Reviewed by:
Liang-Liang Fan,
Central South University, China
Wenbin Zou,
Changhai Hospital, China

***Correspondence:**
Yi Zhou
zhouyi6@mail.sysu.edu.cn

†These authors share senior authorship

Specialty section:
This article was submitted to Genetic Disorders, a section of the journal Frontiers in Genetics

Received: *24 May 2019*
Accepted: *13 August 2019*
Published: *13 September 2019*

Citation:
Wang Y, Zhu C, Du L, Li Q, Lin M-F, Férec C, Cooper DN, Chen J-M and Zhou Y (2019) Compound Heterozygosity for Novel Truncating Variants in the LMOD3 Gene as the Cause of Polyhydramnios in Two Successive Fetuses. Front. Genet. 10:835. doi: 10.3389/fgene.2019.00835

Polyhydramnios is sometimes associated with genetic defects. However, establishing an accurate diagnosis and pinpointing the precise genetic cause of polyhydramnios in any given case represents a major challenge because it is known to occur in association with over 200 different conditions. Whole exome sequencing (WES) is now a routine part of the clinical workup, particularly with diseases characterized by atypical manifestations and significant genetic heterogeneity. Here we describe the identification, by means of WES, of novel compound heterozygous truncating variants in the *LMOD3* gene [i.e., c.1412delA (p.Lys471Serfs*18) and c.1283dupC (p.Gly429Trpfs*35)] in a Chinese family with two successive fetuses affected with polyhydramnios, thereby potentiating the prenatal diagnosis of nemaline myopathy (NM) in the proband. *LMOD3* encodes leiomodin-3, which is localized to the pointed ends of thin filaments and acts as a catalyst of actin nucleation in skeletal and cardiac muscle. This is the first study to describe the prenatal and postnatal manifestations of *LMOD3*-related NM in the Chinese population. Of all the currently reported NM-causing *LMOD3* nonsense and frameshifting variants, c.1412delA generates the shortest truncation at the C-terminal end of the protein, underscoring the critical role of the WH2 domain in LMOD3 structure and function. Survey of the prenatal phenotypes of all known *LMOD3*-related severe NM cases served to identify fetal edema as a novel presenting feature that may provide an early clue to facilitate prenatal diagnosis of the disease.

Keywords: *LMOD3*, nemaline myopathy, polyhydramnios, prenatal diagnosis, truncating variant, whole exome sequencing

INTRODUCTION

Polyhydramnios, an excess of amniotic fluid in the amniotic sac, is diagnosed if the single deepest pocket (SDP) is of ≥8 cm or the amniotic fluid index (AFI) is of ≥25 cm (Moore and Cayle, 1990; Magann et al., 2007). It is present in 1–3% of pregnancies (Maymon et al., 1998; Biggio et al., 1999; Magann et al., 2007) and may occur as a consequence of both environmental

(e.g., perinatal exposure to TORCH infections, maternal gestational and pregestational diabetes) and genetic factors (Bartha et al., 2003; Keshavarz et al., 2005; Touboul et al., 2007; Kishore et al., 2011; Kollmann et al., 2014). Indeed, a search for polyhydramnios in the Human Phenotype Ontology database (https://hpo.jax.org/app/browse/term/HP:0001561) yielded 204 different conditions and 143 different genes. Nearly one half of all pregnancies with polyhydramnios are associated with varying degrees of fetal abnormality, including fetal death (Kollmann et al., 2014). Therefore, it is extremely important to identify the genetic causes of polyhydramnios in affected families with a view to providing genetic counselling and prenatal diagnosis for subsequent pregnancies.

With the decreasing cost of next generation sequencing, whole exome sequencing (WES) has been increasingly employed as an important diagnostic tool for clinical purposes (Boycott et al., 2013; Lee et al., 2014; Posey et al., 2017), especially in the case of diseases characterized by atypical manifestations and significant levels of genetic heterogeneity (Ku et al., 2012). Here we describe the use of WES to identify the genetic cause of polyhydramnios in two successive fetuses in a Chinese family.

CASE PRESENTATION

A 35-year-old woman was referred to our centre at the First Affiliated Hospital of Sun Yat-Sen University after her third fetus (II:3; **Figure 1A**) had been found to have polyhydramnios (SDP of 9.5 cm and an AFI of 30.7 cm) and hydrocele of testis at 29 gestational weeks (GW). Prior to this, her second fetus (II:2) had presented with polyhydramnios and generalized edema (pleural effusion, ascites, skin edema of the chest, abdomen, and scalp) at 34 GW (**Figures 2A–D**) and was then terminated at 36 GW. Her first child (II:1) is a healthy girl, now aged 6 years. Both parents were of south Chinese origin, healthy and non-consanguineous. Exposure to known mutagenic or teratogenic agents during pregnancy was not reported.

In the case of II:2, only standard G-banding karyotyping (using cord blood cells taken at 35 GW) followed by chromosomal microarray analysis were performed at the time, yielding negative findings. In the case of fetus II:3, molecular genetic analysis including WES was performed at 31 GW. The WES results were returned to us 6 weeks later, confirming a diagnosis of nemaline myopathy (NM) (see below). Genetic counselling was provided to the couple, who opted to continue the pregnancy. The boy was born by full-term vaginal delivery at 37 GW; his Apgar scores were 5 and 6 (normal, 10) at 1 and 5 min of life, respectively. He showed stiffness of limbs, little movement and scrotal swelling, and died of respiratory failure 2 days after birth.

Clinical findings in the two affected fetuses (II:2 and II:3) are illustrated in **Figure 2** and summarized in **Table 1**.

Abbreviations: AFI, amniotic fluid index; GW, gestational weeks; NM, nemaline myopathy; SDP, single deepest pocket; WES, whole exome sequencing.

MATERIALS AND METHODS

Subjects

This study was carried out in accordance with the recommendations of "ethical regulations of biomedical research involving humans, Ethics Committee of the First Hospital affiliated to Sun Yat-sen University" with written informed consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the "Ethics Committee of the First Hospital affiliated to Sun Yat-sen University". Prenatal diagnosis was undertaken through ultrasound-guided umbilical cord blood puncture in accordance with standard practice.

Karyotype and Chromosomal Microarray Analysis

Standard G-banding karyotyping was performed according to standard laboratory procedures. The chromosomal microarray experiments were conducted using the high-resolution Affymetrix CytoScan HD arrays (Affymetrix Inc., Santa Clara, CA) according to the manufacturer's protocols. Detected copy number variants were evaluated for clinical significance by comparing them with data in the scientific literature and publicly available databases: UCSC (Ku et al., 2012), DECIPHER database (https://decipher.sanger.ac.uk/), Database of Genomic Variants (DGV, http://dgv.tcag.ca/dgv/app/home), the International Standards for Cytogenomic Arrays (ISCA, https://www.iscaconsortium.org/), and Online Mendelian Inheritance in Man (OMIM, https://www.omim.org/). The results were then analyzed by the Chromosome Analysis Suite software version 3.3.0; the reporting threshold of copy number variants was set at 10 kb, with marker count at over 50, as previously reported (Wang et al., 2015).

Whole Exome Sequencing (WES)

WES was performed in fetus II:3 and its parents. Genomic DNA was randomly fragmented and then purified by means of the magnetic particle method. Sequences were captured by Agilent SureSelect version 4 (Agilent Technologies, Santa Clara, CA) according to the manufacturer's protocols. After enrichment and purification, the DNA libraries were sequenced on a NextSeq500 sequencer following the manufacturer's instructions (Illumina, San Diego). The reads were aligned to hg19/GRCh37.p10 using the Burrows-Wheeler Aligner (version 0.59) (Li and Durbin, 2009). Base quality recalibration and local realignment of the Burrows-Wheeler aligned reads were then processed by means of GATK IndelRealigner (https://software.broadinstitute.org/gatk/documentation/tooldocs/current/org_broadinstitute_gatk_tools_walkers_indels_IndelRealigner.php) and GATK BaseRecalibrator (https://software.broadinstitute.org/gatk/documentation/tooldocs/current/org_broadinstitute_gatk_tools_walkers_bqsr_BaseRecalibrator.php), respectively. Single nucleotide variations and small indels were revealed by the GATK UnifiedGenotyper (https://software.broadinstitute.org/gatk/

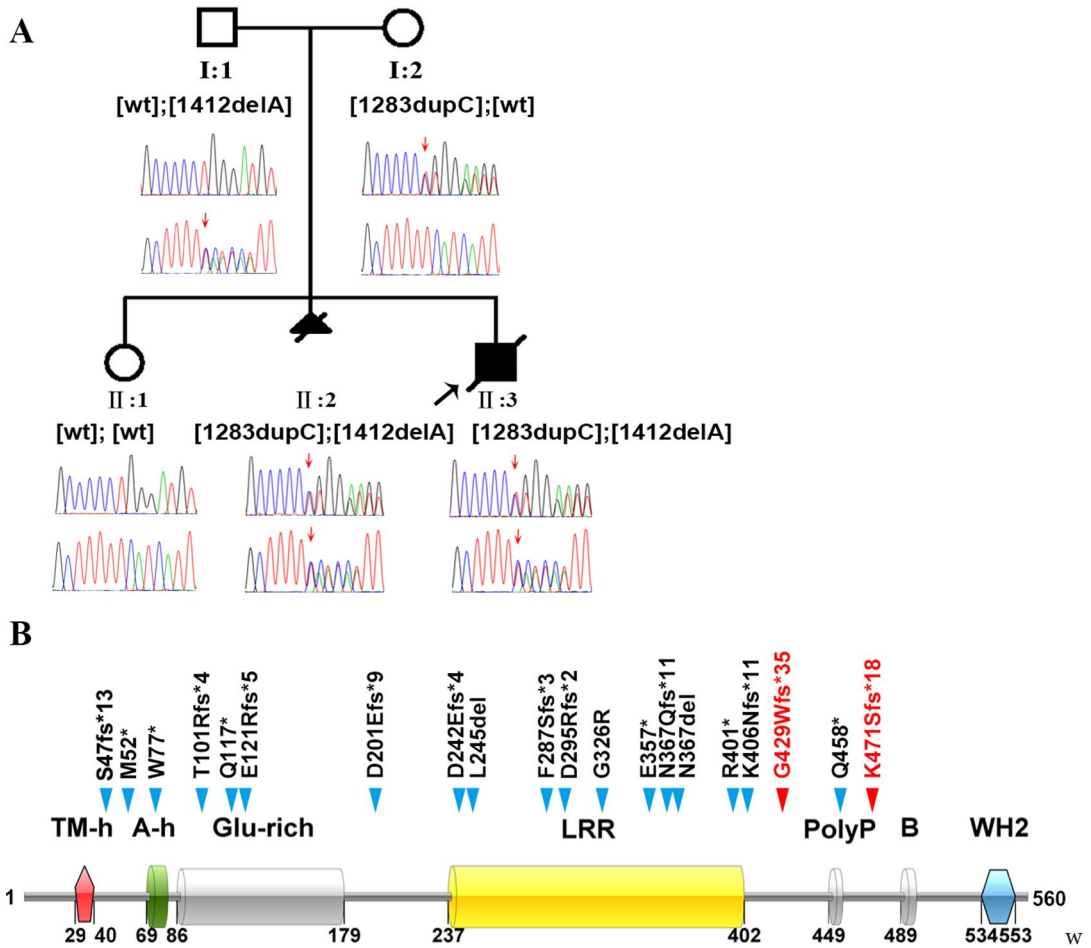

FIGURE 1 | Identification of the genetic basis of polyhydramnios in a Chinese family. **(A)** Family pedigree and DNA sequencing results. Filled triangle with oblique line indicates the fetus with polyhydramnios terminated by therapeutic abortion. Filled square with oblique line indicates the affected boy who died after birth. Arrow indicates the proband. Open symbols indicate clinically unaffected family members. *LMOD3* genotypes are provided for all subjects. Red arrows indicate the mutation positions. wt, wild-type. **(B)** Illustration of the LMOD3 structure and all currently reported severe NM-causing *LMOD3* variants. The two novel variants identified in the current study are highlighted in red. Previous reported variants (in black) were from (Yuen et al., 2014; Berkenstadt et al., 2018; Michael et al., 2019). See **Supplementary Table S1** for variant descriptions at the nucleotide sequence level.

documentation/tooldocs/current/org_broadinstitute_gatk_tools_walkers_genotyper_UnifiedGenotyper.php). Variants were finally annotated employing the Consensus Coding Sequences Database at the National Centre for Biotechnology Information (https://www.ncbi.nlm.nih.gov/CCDS/).

Putative causal variants were sought in an unbiased and hypothesis-free manner. Trio sample strategy was applied to remove the bias that may arise by proband-only sequencing. Literature and population databases were used for variant annotation, including 1000 Genomes, dbSNP, GenomAD, Clinvar, HGMD, and OMIM. The synonymous and common SNPs (MAF > 5%) were filtered out, and rare variants with high confidence were considered as disease-causing candidates for further genetic evaluation. Multiple computational algorithms were applied to assist the genetic evaluation of pathogenicity, including SIFT, Polyphen-2, and MutationTaster. Variants occurring in known phenotype-causing or -associated genes as well as in candidate genes selected on the basis of known biological, physiological or functional relevance to phenotype were considered with priority. The interpretation of variants was managed according to the American College of Medical Genetics (ACMG) guidelines.

Sanger Sequencing

All family members were subjected to Sanger sequencing of exon 2 of the *LMOD3* gene. PCR primers were designed by Oligo 6.0 (http://www.oligo.net/downloads.html). Primer sequences and PCR conditions are available upon request.

Reference Sequence and Variant Nomenclature

NM_198271.4 was employed as the *LMOD3* mRNA reference sequence. Nomenclature for the description of *LMOD3* variants

TABLE 1 | Summary of prenatal feature in 24 *LMOD3* mutation-positive cases of nemaline myopathy.

Case	Reference	Gender	Ethnicity	Polyhydramnios	Decreased fetal movements	Arthrogryposis/ contracture	Fractures	Fetal edema	Other anomalies	Preterm delivery	Age at death (or current age if alive)	*LMOD3* genotype†
1	Yuen et al., 2014	F	Algerian	+	–	+	+	–		+	neonatal	p.S47fs*13 homozygote
2		M	Belgian	–	+	+	–	–		–	10 months	p.S47fs*13 homozygote
3a		F	Portuguese	+	+	+	–	–		–	neonatal	p.M52* homozygote
3b		F	Portuguese	+	–	+	–	–		–	1 months, alive	
4		F	Japanese	+	+	–	–	–		–	2 months, alive	p.[T101Rfs*4]; [D201Efs*9]
5		F	Japanese	+	+	+	–	+	Subdural hematoma	+	10 months, alive	p.Q117* homozygote
6		F	Japanese	+	+	–	–	–	microcephaly	–	1 year, alive	p.[Q117*]; [K406Nfs*11]
7		F	Italian	+	+	+	–	–		+	4 months	p.F287Sfs*3 homozygote
8		M	Ecuadorianmn	+	+	+	+	–		+	6 weeks	p.[G326R]; [Q458*]
9		M	Swedish	+	–	+	–	–		–	5 months	p.E357* homozygote
10		M	Afghan	–	+	+	–	–		+	neonatal	p.N367Qfs*11 homozygote
11		M	Afghan	–	–	+	–	–		+	2 months	p.N367Qfs*11 homozygote
12		F	Pakistani	+	–	–	–	–		–	3 months	p.N367Qfs*11 homozygote
13		F	Pakistani	+	+	–	–	–		–	neonatal	p.N367Qfs*11 homozygote
14		F	Australian	+	+	–	–	–		–	10 years, alive	p.[N367del]; [R401*]
15		F	Australian	+	–	–	–	–		–	4 years, alive	p.[N367del]; [R401*]
16a	Abbott et al., 2017	F	Turkish	–	+	–	+	+		+	neonatal	p.S47fs*13 homozygote
16b		F	Turkish	–	+	–	+	–		+	2 months	
16c		F	Turkish	+	–	–	+	+		+	neonatal	
17	Berkenstadt et al., 2018	F	Moroccan	+	+	+	–	–		NE	induced abortion	p.[E121Rfs*5]; [L245del]
18	Michael et al., 2019	F	Turkish	+	+	+	–	+	atrial septum defect	–	8 years, alive	p.D295Rfs*2 homozygote
19		M	NE	+	+	+	–	–	atrial septum defect	–	5 months	p.E357* homozygote
20a	This study	Unknown	Chinese	+	+	–	–	+		NE	induced abortion	p.[G429Wfs*35]; [K471Sfs*18]
20b		M	Chinese	+	+	–	–	+		–	neonatal	
Total				**19/24 (79.2%)**	**17/24 (70.8%)**	**13/24 (54.1%)**	**5/24 (20.8%)**	**6/24 (25.0%)**		**9/22 (40.9%)**		

+, present; –, absent; F, female; M, male; NE, not evaluated.

*†See **Supplementary Table S1** for variant descriptions at the nucleotide sequence level.*

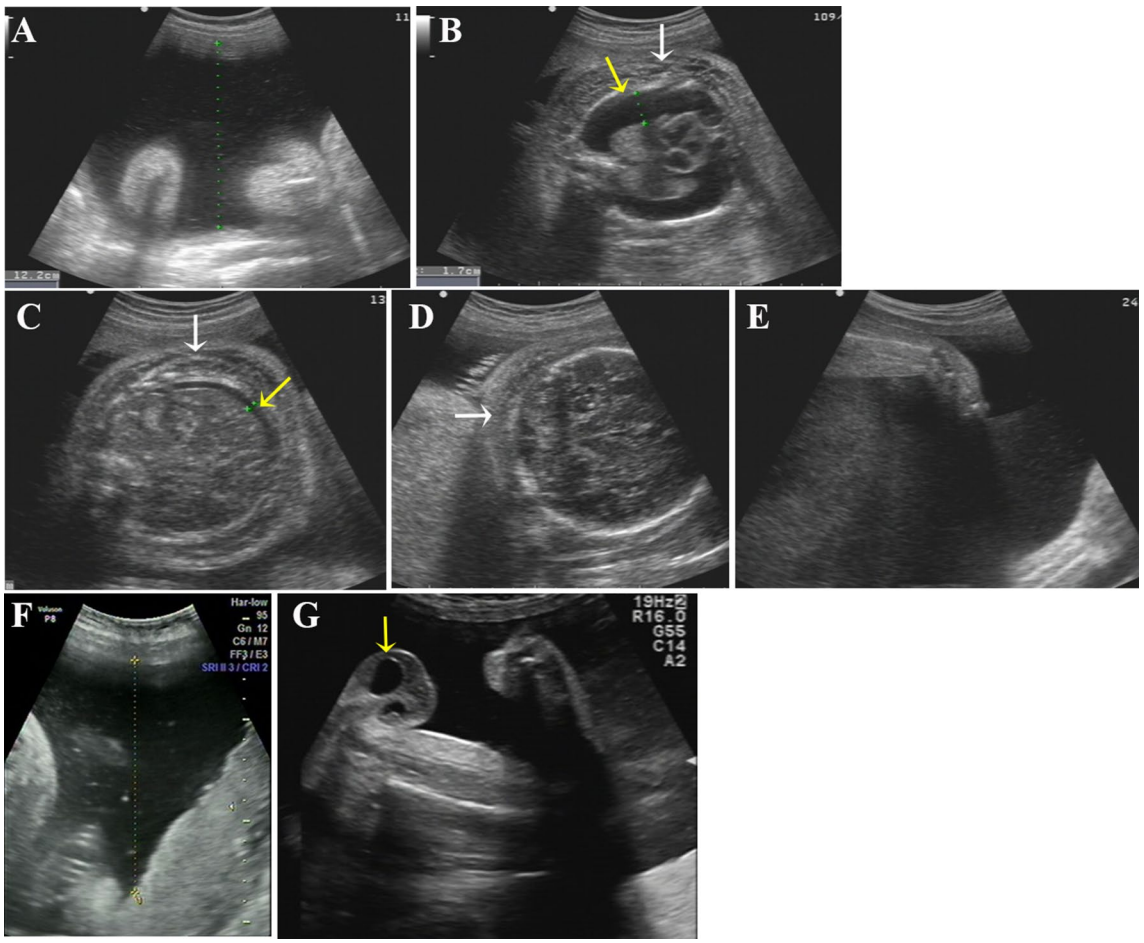

FIGURE 2 | Ultrasound photographs of the 2 fetuses with polyhydramnios. **(A–E)** Ultrasound images of the first affected fetus (II-2) at 34 GW. Two-dimensional ultrasound showing polyhydramnios (amniotic fluid deep, AFD, 12.2 cm) **(A)** and fixed limbs **(E)**. Bilateral pleural effusion (**B**, yellow arrow), ascites (**C**, yellow arrow), skin edema of chest (**B**, white arrow), abdomen (**C**, white arrow), and scalp (**D**, white arrow) were detected by ultrasound. **(F, G)** Ultrasound pictures of the second affected nemaline myopathy fetus (II-3) at 33 GW. Two-dimensional ultrasound revealed polyhydramnios (amniotic fluid deep, AFD, 11.5cm) **(F)** and hydrocele of testis **(G)** (yellow arrow).

and genotypes followed the Human Genome Variation Society (HGVS) recommendations (den Dunnen et al., 2016).

Currently Reported Variants in the *LMOD3* Gene

Currently reported variants in the *LMOD3* gene were obtained from the Professional version of the Human Gene Mutation Database (HGMD; https://www.qiagenbioinformatics.com/products/human-gene-mutation-database/) (Stenson et al., 2017) and augmented by a manual literature search.

RESULTS

Karyotyping of fetus II:3 indicated the presence of a normal set of chromosomes. Chromosomal microarray analysis also failed to identify any pathogenic copy number variants. We then screened II:3 and his parents by WES to search for putative causal variants in an unbiased and hypothesis-free manner. Variants detected were then prioritized based on sequencing quality, allele frequency in the normal population, gene product damage potential, zygosity and mode of inheritance. The mean depth of coverage for the coding regions targeted with the WES was 267×. A mean of 99.2% of bases in the targeted coding regions were covered by at least 10 reads and 99.0% of bases in the targeted coding regions were covered by more than 20 reads. Compound heterozygous truncating variants, c.1283dupC (p.Gly429Trpfs*35) and c.1412delA (p.Lys471Serfs*18), were identified in exon 2 of the *LMOD3* gene (NM_0198272) in II:3; c.1283dupC was inherited from the mother whereas c.1412delA was inherited from the father (**Figure 1A**). The genotypes in II:3 and his parents were confirmed by Sanger sequencing of exon 2 of the *LMOD3* gene. We also sequenced exon 2 of the *LMOD3* gene in II:1 and II:2. Neither variant was found in the former while both variants were found in the latter (**Figure 1A**). Neither c.1283dupC nor c.1412delA have been previously described in the literature and are also absent from

the Genome Aggregation Database (gnomAD, http://gnomad.broadinstitute.org/).

Homozygous or compound heterozygous variants (predominantly truncating and nonsense) in the *LMOD3* gene have recently been identified to be a cause of autosomal recessive NM (Yuen et al., 2014). Our identification of novel compound heterozygous truncating variants in the *LMOD3* gene in the two fetuses with polyhydramnios therefore firmly established the diagnosis of NM.

All *LMOD3* variants so far reported to be causative for severe NM are illustrated in the context of the LMOD3 protein structure (**Figure 1B**). The structure of LMOD3 is in accordance with Chereau et al. (2008), UniProt (http://www.uniprot.org/) and InterProScan 4 (Zdobnov and Apweiler, 2001). Additionally, we collated prenatal features in all 22 known *LMOD3* mutation-positive severe NM cases (**Table 1**).

DISCUSSION

Employing WES, we have successfully identified two novel compound heterozygous truncating variants, c.1283dupC and c.1412delA, in the *LMOD3* gene in two successive fetuses from the same family with polyhydramnios. *LMOD3* encodes a member of the leiomodin family of proteins, leiomodin-3, which is localized to the pointed ends of thin filaments and acts as a catalyst of actin nucleation in skeletal and cardiac muscle where it is expressed (Conley et al., 2001; Chereau et al., 2008; Qualmann and Kessels, 2009; Campellone and Welch, 2010; Tsukada et al., 2010; Cenik et al., 2015). In 2014, Yuen et al. (2014) identified *LMOD3* as a new causative gene for autosomal recessive NM. NM is a disorder that is characterized by muscle dysfunction and electron-dense protein accumulations (nemaline bodies) in myofibers (Sandaradura and North, 2015). Patients with mutations in *LMOD3* often present with a severe congenital form of early-onset generalized muscle weakness and hypotonia with respiratory insufficiency and feeding difficulties; they usually die in early infancy (Yuen et al., 2014). NM is clinically and genetically heterogeneous, making it difficult to establish a correct diagnosis from clinical features alone. Fetuses with NM usually display decreased fetal movements and polyhydramnios without any particular disease-defining features. The identification of compound heterozygous *LMOD3* variants in our family therefore provided a definite diagnosis of the disease that could not otherwise have been made merely on the basis of clinical findings in the fetus.

To date, a total of 25 *LMOD3* variants have been reported in the literature (**Supplementary Table S1**). Of these, 4 missense variants, p.Arg83His, p.Glu142Asp, p.Lys282Glu, p.Pro552His, have been reported to cause Kleine-Levin syndrome (Al Shareef et al., 2019); a further two missense variants, p.Leu550Phe and p.Gln335Arg, have been associated with a mild form of congenital NM (Schatz et al., 2018) whereas a truncating variant, c.112delG (p.Glu38Lysfs*15), has been detected in a case without a definite clinical diagnosis (Theunissen et al., 2018). By contrast, the remaining 18 variants (i.e., variants highlighted in black in **Figure 1B**) were reported to be causative for severe NM (Yuen et al., 2014; Abbott et al., 2017; Berkenstadt et al., 2018; Michael et al., 2019). Notably, 83% (n = 15) of these latter 18 variants represent truncating variants (nonsense or frameshift). The addition of our two newly identified truncating variants increased this figure to 85% (17/20). Moreover, most previous studies analyzed patients with a diagnosis or suggestive diagnosis of NM (Yuen et al., 2014). Only very recently have *LMOD3* variants been described in any detail in fetuses (Berkenstadt et al., 2018). Our study is the first to describe the prenatal and postnatal manifestations of *LMOD3*-related NM in the Chinese population.

LMOD3 contains three actin-binding domains, which are actin-binding helix [A-h], residues 69–79; leucine-rich repeat domain [LLR], residues 237–402; and Wiskott-Aldrich-syndrome protein homology 2 domain [WH2], residues 534–553 (Yuen et al., 2014). It should be noted that 1 of our 2 novel *LMOD3* mutations, c.1412delA, led to the shortest C-terminal truncation among the *LMOD3* nonsense and frameshifting variants reported to date (**Figure 1B**). This, together with the fact that c.1412delA is associated with a severe form of NM, serves to highlight the indispensable role of the WH2 domain in LMOD3 protein and function. Current models and *in vitro* assays support the view that the WH2 domain, an actin nucleator, stabilizes actin monomers in the trimer conformation, thereby promoting nucleation (Qualmann and Kessels, 2009; Cenik et al., 2015).

Finally, we evaluated the prenatal features of *LMOD3*-related severe NM reported elsewhere (Yuen et al., 2014; Abbott et al., 2017; Berkenstadt et al., 2018; Michael et al., 2019) as well as in this study. As shown in **Table 1**, the most frequently observed anomaly is polyhydramnios (79.2%), followed by decreased fetal movements (70.8%), arthrogryposis/contracture (54.1%), fetal edema (25.0%), and fractures (20.8%). This survey identified fetal edema, which has an even higher frequency than fracture (Abbott et al., 2017), to be a new noteworthy feature of *LMOD3*-related severe NM in the fetal period. Fetal edema could be explained as follows: decreased fetal movement allows accumulation of subcutaneous fluid and causes subsequent atony in the diaphragm, leading to a change in thoracic pressure and a disturbance of lymphatic circulation (Vardon et al., 1998). Taken together, we suggest that *LMOD3*-related severe NM should be considered in cases of prenatal diagnosis of fetal edema and polyhydramnios.

This is the first report of NM-causative *LMOD3* variants, and the first example of prenatal diagnosis of NM, in the Chinese population. The identification of a novel *LMOD3* variant (i.e., c.1412delA) generating the shortest LMOD3 truncation reported to date underscores the critical role of the WH2 domain in LMOD3 structure and function. Further, survey of the prenatal phenotypes of *LMOD3*-related severe NM identified fetal edema as a new presenting feature that may provide an early clue to facilitate prenatal diagnosis of the disease. Finally, we demonstrate once again the power of WES as an effective means to deliver a timely molecular diagnosis in a genetically and clinically highly heterogeneous disorder. We anticipate that the availability of genetic testing for NM will also allow clinicians to offer more reliable prognostic and recurrence risk information to families at risk.

DATA AVAILABILITY

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

ETHICS STATEMENT

Written informed consent was obtained from the parents. The study was approved by the Ethics Committee of the First Affiliated Hospital of Sun Yat-Sen University, Guangzhou, China and performed in accordance with the principles enshrined in the Declaration of Helsinki.

AUTHOR CONTRIBUTIONS

YW and YZ designed the study with the assistance of J-MC. YW performed the genetic analysis and bioinformatics evaluations. CZ, LD, QL, and M-FL conducted the clinical and ultrasound evaluations. YW and J-MC drafted the manuscript. DC critically revised the manuscript. All authors analyzed the data, contributed to revision, and approved the final manuscript.

FUNDING

This work was supported in part by the National Natural Science Foundation of China (no. 31801045) and the PhD Start-up Fund of Natural Science Foundation of Guangdong Province, China (2018A030310244).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2019.00835/full#supplementary-material

REFERENCES

Abbott, M., Jain, M., Pferdehirt, R., Chen, Y., Tran, A., Duz, M. B., et al. (2017). Neonatal fractures as a presenting feature of *LMOD3*-associated congenital myopathy. *Am. J. Med. Genet. A* 173, 2789–2794. doi: 10.1002/ajmg.a.38383

Al Shareef, S., Basit, S., Li, S., Pfister, C., Pradervand, S., Lecendreux, M., et al. (2019). Kleine-Levin syndrome is associated with *LMOD3* variants. *J. Sleep Res.* 28, e12718. doi: 10.1111/jsr.12718

Bartha, J. L., Martinez-Del-Fresno, P., and Comino-Delgado, R. (2003). Early diagnosis of gestational diabetes mellitus and prevention of diabetes-related complications. *Eur. J. Obstet. Gynecol. Reprod. Biol.* 109, 41–44. doi: 10.1016/S0301-2115(02)00480-3

Berkenstadt, M., Pode-Shakked, B., Barel, O., Barash, H., Achiron, R., Gilboa, Y., et al. (2018). *LMOD3*-associated nemaline myopathy: prenatal ultrasonographic, pathologic, and molecular findings. *J. Ultrasound Med.* 37, 1827–1833. doi: 10.1002/jum.14520

Biggio, J. J., Wenstrom, K. D., Dubard, M. B., and Cliver, S. P. (1999). Hydramnios prediction of adverse perinatal outcome. *Obstet. Gynecol.* 94, 773–777. doi: 10.1016/S0029-7844(99)00370-1

Boycott, K. M., Vanstone, M. R., Bulman, D. E., and MacKenzie, A. E. (2013). Rare-disease genetics in the era of next-generation sequencing: discovery to translation. *Nat. Rev. Genet.* 14, 681–691. doi: 10.1038/nrg3555

Campellone, K. G., and Welch, M. D. (2010). A nucleator arms race: cellular control of actin assembly. *Nat. Rev. Mol. Cell Biol.* 11, 237–251. doi: 10.1038/nrm2867

Cenik, B. K., Garg, A., McAnally, J. R., Shelton, J. M., Richardson, J. A., Bassel-Duby, R., et al. (2015). Severe myopathy in mice lacking the MEF2/SRF-dependent gene leiomodin-3. *J. Clin. Invest.* 125, 1569–1578. doi: 10.1172/JCI80115

Chereau, D., Boczkowska, M., Skwarek-Maruszewska, A., Fujiwara, I., Hayes, D. B., Rebowski, G., et al. (2008). Leiomodin is an actin filament nucleator in muscle cells. *Science* 320, 239–243. doi: 10.1126/science.1155313

Conley, C. A., Fritz-Six, K. L., Almenar-Queralt, A., and Fowler, V. M. (2001). Leiomodins: larger members of the tropomodulin (Tmod) gene family. *Genomics* 73, 127–139. doi: 10.1006/geno.2000.6501

den Dunnen, J. T., Dalgleish, R., Maglott, D. R., Hart, R. K., Greenblatt, M. S., McGowan-Jordan, J., et al. (2016). HGVS recommendations for the description of sequence variants: 2016 Update. *Hum. Mutat.* 37, 564–569. doi: 10.1002/humu.22981

Keshavarz, M., Cheung, N. W., Babaee, G. R., Moghadam, H. K., Ajami, M. E., and Shariati, M. (2005). Gestational diabetes in Iran: incidence, risk factors and pregnancy outcomes. *Diabetes Res. Clin. Pract.* 69, 279–286. doi: 10.1016/j.diabres.2005.01.011

Kishore, J., Misra, R., Paisal, A., and Pradeep, Y. (2011). Adverse reproductive outcome induced by Parvovirus B19 and TORCH infections in women with high-risk pregnancy. *J. Infect. Dev. Ctries* 5, 868–873. doi: 10.3855/jidc.1533

Kollmann, M., Voetsch, J., Koidl, C., Schest, E., Haeusler, M., Lang, U., et al. (2014). Etiology and perinatal outcome of polyhydramnios. *Ultraschall Med.* 35, 350–356. doi: 10.1055/s-0034-1366115

Ku, C. S., Cooper, D. N., Polychronakos, C., Naidoo, N., Wu, M., and Soong, R. (2012). Exome sequencing: dual role as a discovery and diagnostic tool. *Ann. Neurol.* 71, 5–14. doi: 10.1002/ana.22647

Lee, H., Deignan, J. L., Dorrani, N., Strom, S. P., Kantarci, S., Quintero-Rivera, F., et al. (2014). Clinical exome sequencing for genetic identification of rare Mendelian disorders. *JAMA* 312, 1880–1887. doi: 10.1001/jama.2014.14604

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754–1760. doi: 10.1093/bioinformatics/btp324

Magann, E. F., Chauhan, S. P., Doherty, D. A., Lutgendorf, M. A., Magann, M. I., and Morrison, J. C. (2007). A review of idiopathic hydramnios and pregnancy outcomes. *Obstet. Gynecol. Surv.* 62, 795–802. doi: 10.1097/01.ogx.0000290349.58707.e0

Maymon, E., Ghezzi, F., Shoham-Vardi, I., Franchi, M., Silberstein, T., Wiznitzer, A., et al. (1998). Isolated hydramnios at term gestation and the occurrence of peripartum complications. *Eur. J. Obstet. Gynecol. Reprod. Biol.* 77, 157–161. doi: 10.1016/S0301-2115(97)00250-9

Michael, E., Hedberg-Oldfors, C., Wilmar, P., Visuttijai, K., Oldfors, A., and Darin, N. (2019). Long-term follow-up and characteristic pathological findings in severe nemaline myopathy due to *LMOD3* mutations. *Neuromuscul. Disord.* 29, 108–113. doi: 10.1016/j.nmd.2018.12.009

Moore, T. R., and Cayle, J. E. (1990). The amniotic fluid index in normal human pregnancy. *Am. J. Obstet. Gynecol.* 162, 1168–1173. doi: 10.1016/0002-9378(90)90009-V

Posey, J. E., Harel, T., Liu, P., Rosenfeld, J. A., James, R. A., Coban, A. Z., et al. (2017). Resolution of disease phenotypes resulting from multilocus genomic variation. *N. Engl. J. Med.* 376, 21–31. doi: 10.1056/NEJMoa1516767

Qualmann, B., and Kessels, M. M. (2009). New players in actin polymerization — WH2-domain-containing actin nucleators. *Trends Cell Biol.* 19, 276–285. doi: 10.1016/j.tcb.2009.03.004

Sandaradura, S., and North, K. N. (2015). LMOD3: the "missing link" in nemaline myopathy? *Oncotarget* 6, 26548–26549. doi: 10.18632/oncotarget.5267

Schatz, U. A., Weiss, S., Wenninger, S., Schoser, B., Muss, W. H., Bittner, R. E., et al. (2018). Evidence of mild founder *LMOD3* mutations causing nemaline myopathy 10 in Germany and Austria. *Neurology* 91, e1690–e1694. doi: 10.1212/WNL.0000000000006428

Stenson, P. D., Mort, M., Ball, E. V., Evans, K., Hayden, M., Heywood, S., et al. (2017). The human gene mutation database: towards a comprehensive repository of inherited mutation data for medical research, genetic diagnosis and next-generation sequencing studies. *Hum. Genet.* 136, 665–677. doi: 10.1007/s00439-017-1779-6

Theunissen, T., Nguyen, M., Kamps, R., Hendrickx, A. T., Sallevelt, S., Gottschalk, R., et al. (2018). Whole exome sequencing is the preferred strategy to identify the genetic defect in patients with a probable or possible mitochondrial cause. *Front. Genet.* 9, 400. doi: 10.3389/fgene.2018.00400

Touboul, C., Boileau, P., Picone, O., Foix-l'Helias, L., Frydman, R., and Senat, M. (2007). Outcome of children born out of pregnancies complicated by unexplained polyhydramnios. *BJOG* 114, 489–492. doi: 10.1111/j.1471-0528.2006.01258.x

Tsukada, T., Pappas, C. T., Moroz, N., Antin, P. B., Kostyukova, A. S., and Gregorio, C. C. (2010). Leiomodin-2 is an antagonist of tropomodulin-1 at the pointed end of the thin filaments in cardiac muscle. *J. Cell Sci.* 123, 3136–3145. doi: 10.1242/jcs.071837

Vardon, D., Chau, C., Sigodi, S., Figarella-Branger, D., and Boubli, L. (1998). Congenital rapidly fatal form of nemaline myopathy with fetal hydrops and arthrogryposis. A case report and review. *Fetal Diagn. Ther.* 13, 244–249. doi: 10.1159/000020847

Wang, Y., Su, P., Hu, B., Zhu, W., Li, Q., Yuan, P., et al. (2015). Characterization of 26 deletion CNVs reveals the frequent occurrence of micro-mutations within the breakpoint-flanking regions and frequent repair of double-strand breaks by templated insertions derived from remote genomic regions. *Hum. Genet.* 134, 589–603. doi: 10.1007/s00439-015-1539-4

Yuen, M., Sandaradura, S. A., Dowling, J. J., Kostyukova, A. S., Moroz, N., Quinlan, K. G., et al. (2014). Leiomodin-3 dysfunction results in thin filament disorganization and nemaline myopathy. *J. Clin. Invest.* 124, 4693–4708. doi: 10.1172/JCI75199

Zdobnov, E. M., and Apweiler, R. (2001). InterProScan—an integration platform for the signature-recognition methods in InterPro. *Bioinformatics* 17, 847–848 doi: 10.1093/bioinformatics/17.9.847.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Genetics

ORIGINAL RESEARCH
published: 18 September 2019
doi: 10.3389/fgene.2019.00862

Improved Molecular Diagnosis of McCune–Albright Syndrome and Bone Fibrous Dysplasia by Digital PCR

Francesca Marta Elli[1], Luisa de Sanctis[2], Massimiliano Bergallo[2], Maria Antonia Maffini[1], Arianna Pirelli[1], Ilaria Galliano[2], Paolo Bordogna[3], Maura Arosio[1,3] and Giovanna Mantovani[1,3]*

[1] Department of Clinical Sciences and Community Health, University of Milan, Milan, Italy, [2] Department of Public Health and Pediatric Sciences, University of Torino, Regina Margherita Children's Hospital-AOU Cittàdella Salute e dellaScienza, Torino, Italy, [3] Endocrinology Unit, Fondazione IRCCS Ca' GrandaOspedale Maggiore Policlinico, Milan, Italy

McCune–Albright syndrome (MAS) is a rare congenital disorder characterized by the association of endocrine and nonendocrine anomalies caused by somatic activating variants of *GNAS*. The mosaic state of variants makes the clinical presentation extremely heterogeneous depending on involved tissues. Biological samples bearing a low level of mosaicism frequently lead to false-negative results with an underestimation of causative molecular alterations, and the analysis of biopsies is often needed to obtain a molecular diagnosis. To date, no reliable analytical method for the noninvasive testing of blood is available. This study was aimed at validating a novel and highly sensitive technique, the digital PCR (dPCR), to increase the detection rate of *GNAS* alterations in patients with a clinical suspicion of MAS and, in particular, in blood. We screened different tissues (blood, bone, cutis, ovary, and ovarian cyst) collected from 54 MAS patients by different technical approaches. Considering blood, Sanger was unable to detect mutations, the allele-specific PCR and the co-amplification at lower denaturation temperature had a 9.1% and 18.1% detection rate, respectively, whereas the dPCR reached a 37.8% detection rate. In conclusion, the dPCR resulted in a cost-effective, reliable, and rapid method allowing the selective amplification of low-frequency variants and able to improve *GNAS* mutant allele detection, especially in the blood.

Keywords: McCune–Albright's syndrome, bone fibrous dysplasia, precocious puberty, GNAS, mosaicism, digital PCR

OPEN ACCESS

Edited by:
Yueqiu Tan,
Central South University, China

Reviewed by:
Malgorzata Wasniewska,
University of Messina, Italy
Eduardo Fernandez-Rebollo,
University of Southern Denmark, Denmark

****Correspondence:***
Francesca Marta Elli
francesca.elli@unimi.it

Specialty section:
This article was submitted to Genetic Disorders, a section of the journal Frontiers in Genetics

Received: *04 May 2019*
Accepted: *19 August 2019*
Published: *18 September 2019*

Citation:
Elli FM, de Sanctis L, Bergallo M, Maffini MA, Pirelli A, Galliano I, Bordogna P, Arosio M and Mantovani G (2019) Improved Molecular Diagnosis of McCune–Albright Syndrome and Bone Fibrous Dysplasia by Digital PCR. Front. Genet. 10:862. doi: 10.3389/fgene.2019.00862

INTRODUCTION

The McCune–Albright syndrome (MAS, MIM #174800) is a rare congenital disorder that comprises the clinical triad of endocrinopathies, cafè-au-lait pigmented skin lesions (SP), and fibrous dysplasia of bone (FD). The endocrine dysregulation may include various autonomous hormonal hyperfunctions, such as precocious puberty, hyperthyroidism, growth hormone excess, adrenal hyperplasia, hypophosphatemic osteomalacia, and GH- and/or PRL- secreting pituitary tumors (McCune, 1936; Albright et al., 1937; Robinson et al., 2016; Boyce et al., 2017). In particular, the more frequent and studied alterations of the endocrine glands are the gonadal hyperfunction (characterized by episodes of hyperestrogenism with a consequent reduction in gonadotropin secretion), thyroid

abnormalities (nodular goiter with nodules >1cm), and the GH-IGF1 axis hyperactivity (Matarazzo et al., 2006; Tessaris et al., 2012; Tessaris et al., 2006).

The disease demonstrated sporadic occurrence, phenotypic heterogeneity, variable involvement of different endocrine glands, and a pattern of skin and bone lesions following the lines of the embryologic development, thus it was proposed that MAS derived from a postzygotic genetic defect, leading to a mosaic distribution of mutated cells. Such hypothesis was confirmed by the identification of somatic activating variants affecting the alpha subunit of the stimulatory G protein (Gsα, encoded by GNAS, MIM*139320) (Landis et al., 1989; Weinstein et al., 1991; Schwindinger et al., 1992; Candeliere et al., 1997). Most gain-of-function *GNAS* genetic variants occurred at the Arg202 residue (historically and still frequently reported as Arg201) and determined the constitutional activation of the stimulatory G protein and, consequently, of adenylyl cyclase (Landis et al., 1989). The tissue-by-tissue study by droplet digital PCR (ddPCR) of autopsy samples from a patient confirmed that MAS can affect a wider range of tissues, leading to extraskeletal manifestations, such as gastrointestinal reflux and/or polyps, pancreatitis, hepatobiliary disease, cardiac disease (sudden death, tachycardia, high output heart failure, and aortic root dilatation), platelet dysfunction, and cancer (bone, breast, testes, and thyroid) (Salpea and Stratakis, 2014; Vasilev et al., 2014).

The mosaic state of MAS/FD-associated variants determines an extremely heterogeneous presentation of clinical manifestations in patients, depending upon the extent to which affected tissues are involved, and makes establishing the correct molecular diagnosis a real challenge. As a matter of fact, the investigation of biological samples, such as peripheral blood that typically bears a low level of mosaicism, leads to high rates of false-negative results and the underestimation of causative molecular alterations in MAS/FD patients (Hannon et al., 2003; Karadag et al., 2004; Lumbroso et al., 2004; Lietman et al., 2005; Kalfa et al., 2006; Kuznetsov et al., 2008; Liang et al., 2011; Narumi et al., 2013).

A real help to overcome the underestimation of mosaic variants will possibly arrive by the development of single-cell sequencing technologies that, actually, are under development as a liquid biopsy tool to investigate cancer patients noninvasively. In particular, limits, such as elevated costs and obtaining enough genetic material for library preparation, make these methods still far from the adoption by the health care system, but, in the next future, they could represent a promising tool for somatic variants analysis (Wei et al., 2017; Zhu et al., 2018).

Different technical approaches, each with its pros and cons, had been developed in the past years with the aim to detect MAS/FD-associated variants. The first attempt dates back to 1997 when Candeliere and colleagues, (1997) combined consecutive and repeated cycles of PCR amplification with site-directed mutagenesis to generate a PCR product from the normal allele susceptible to restriction nuclease digestion. Other researchers tried to overcome the problem by using a peptide nucleic acid (PNA) primer to inhibit the amplification of the normal allele, combining PNA probes with the fluorescence resonance energy transfer (FRET) technique, pyrosequencing, next-generation sequencing (NGS), or ddPCR (Hannon et al., 2003; Karadag et al., 2004; Liang et al., 2011; Narumi et al., 2013; Vasilev et al., 2014). Selective enrichment methods reached good levels in low-abundance variants detection, but they presented several limits as being expensive, time-consuming, and leading to an elevated risk of generating PCR artifacts and cross-contamination (Qin et al., 2016). In particular, the NGS approach with deep sequence coverage enhances sensitivity and allows for accurate quantification of the level of mosaicism, but the use of a customized DNA probe library followed by deep NGS analysis with a mean coverage depth per base of approximately 800× translates into a cost-per-patient analysis too high for the health care system.

The improvement in the sensitivity of genetic tests for defects in mosaic state associated with MAS/FD is an important step to improve the diagnosis of patients, thus the aim of the present study was to set up and validate a novel technique, digital PCR (dPCR), to increase the detection rate of activating *GNAS* alterations in different tissue samples. The dPCR is a method used for absolute quantification of nucleic acids based on the amplification of single molecules of template with target-specific fluorescent-labeled assays. The subsequent analysis yields the relative or absolute quantification results from the raw imaging data. It was conceived in 1992 and applied to the quantification of KRAS mutations in DNA from colorectal cancer patients. This approach has many potential applications, including the detection and quantification of low-level pathogens, rare genetic sequences, copy number variations (CNVs), gene expression in single cells, and quantification of circulating miRNA expression. It works by partitioning a sample of DNA or cDNA into many individual parallel PCR reactions and allows the detection of sequence variants present at a very low frequency in a pool of wild-type background (Sykes et al., 1992; Vogelstein and Kinzler, 1999). According to the manufacturer's specification, optimized dPCR assays should detect and quantify rare mutant prevalence to ≤0.1%.

In this study, we evaluated the usefulness of a novel technique, the dPCR, to obtain a molecular confirmation in patients with a clinical suspicion of MAS/FD by the screening of different tissues collected from a case series of 54 MAS/FD patients using and comparing different technical approaches. In particular, thanks to the expected ability of dPCR to detect low-abundance somatic single-nucleotide alterations, our main aim was to find *GNAS* mutant alleles in blood samples from MAS/FD patients.

MATERIALS AND METHODS

Patients and Biological Samples

The present study included 54 patients with a clinical diagnosis of MAS/FD. For 41 patients, a single tissue only was available for the analysis [24 blood samples (BL) and 17 biopsies: 10 bone (BO), 4 cutis (CB), 3 ovarian tissue (OT), and/or ovarian cyst (OC)], whereas for 13 patients, we investigated and compared the DNA extracted from different tissues (2 BL+BO, 6 BL+CB/fibroblasts from cutaneous biopsy, and 4 BL+OT/OC). Overall, 79 different samples were tested for the presence of MAS/FD-related activating variants within *GNAS* exon 8. The MAS

clinical suspicion was based on the presence of at least persistent and/or recurrent ovarian cysts associated with peripheral precocious puberty or another typical sign of MAS between cafè-au-lait skin spots or fibrous bone dysplasia. Clinical details, including some patient-specific additional features, and the molecular diagnosis are resumed in the **Supplementary Table 1**. Wild-type controls were healthy people, whereas mutant controls were MAS patient-derived *GNAS* mutated tissues. Informed consent for genetic studies was obtained from all subjects involved in the study.

DNA samples were extracted from peripheral blood in EDTA collection tubes (Flexigene DNA kit, Qiagen, Germany) and fresh or FFPE tissue samples (GentraPuregene tissue kit, Qiagen, Germany), according to the manufacturer's instructions.

Techniques for the Detection of GNAS MAS-Related Genetic Variants

To investigate the presence of heterozygous gain-of-function *GNAS* variants c.604C > T/p.Arg202Cys and c.605G > A/p.Arg202His, commonly reported in the literature as Arg201 (in accordance with the guidelines recommended by the Human Genome Variation Society (HGVS), http://www.hgvs.org/mutnomen/, for a uniform and an unequivocal description of sequence variants in DNA and protein sequences, the nucleotide and protein numbering used in the present article was based on the Locus Reference Genomic (LRG) sequence, https://www.lrg-sequence.org/, covering the *GNAS* transcript NM_001077488.2; the *GNAS* mutations database available at http://databases.lovd.nl/shared/genes/GNAS) associated with the MAS/FD phenotype, different experimental techniques were used: Sanger sequencing, allele-specific PCR (AS-PCR), allele-specific PCR-based TaqMan genotyping with co-amplification at lower denaturation temperature (COLD-MAMA PCR), and dPCR.

Detection limits (LOD) of each method, defined as the lowest concentration that can be reliably distinguished from healthy controls, are a consequence of assay specificity, sample quality/handling, and PCR inhibitors. For the empirical determination of detection limits, we applied the strategy of analyzing no DNA template controls (NTCs, contamination monitors), wild-type negative controls (WTCs, false-positive monitors), and a serial dilution of mutant positive controls in a background of wild-type DNA (MTCs, positive controls for thresholding and making limit of detection calls).

The *GNAS* exon 8 was amplified using the specific couple of primers 5'-ACTCTGAGCCCTCTTTCCAA-3' and 5'-GGTAAC AGTTGGCTTACTGG-3' (thermal conditions: 94°C/5', 40 cycles at 94°C/30", 58°C/30", and 72°C/30", 72°C/5'). Sanger sequencing was performed using the AmpliTaqBigDye Terminator kit and the 3110xl Genetic Analyzer (Applied Biosystems, Foster City, CA, USA).

Amplicons subjected to sequencing were also used as template for AS-PCRs, visualized and analyzed on 3% agarose gels, that was shown to be able to discriminate wild-type and mutant alleles using sequence-specific primers (c.604C 5′-GATTCCAG AAGTCAGGACACG-3′, c.604T 5′-GATTCCAGAAGTCAGGA CACA-3′, c. 605G 5′-GATTCCAGAAGTCAGGACAC-3′, and c.605° 5′-GATTCCAGAAGTCAGGACAT-3′) under stringent thermal conditions (94°C/3', 40 cycles at 94°C/15", 67°C/15", and 72°C/15", 72°C/3').

The GNAS locus fragment bearing MAS/FD -associated variants was investigated by COLD-MAMA PCR, as previously described (de Sanctis et al., 2017).

The dPCR was performed with the QuantStudio™ 3D Digital PCR System platform (all products by ThermoFisher Scientific, Carlsbad, CA, USA) and predesigned FAM-labelled TaqMan® Mutation Detection Assays, GNAS_27895_mu (c.605G > A), and GNAS_27887_mu (c.604C > T). Briefly, 1 µl of 5 ng/µl DNA (A260/280 ratios between 1.7 and 1.9) diluted in 14 µl of reaction mixture (8 µl QuantStudio™ 3D Digital PCR Master Mix, 1.6 µl TaqMan® Mutation Detection Assay, and 4.4 µl nuclease-free water) were loaded into QuantStudio™ 3D Digital PCR 20K Chips and amplified (thermal conditions: 1 cycle at 96°C/10', 40 cycles at 66°C/1', and 98°C/45', 1 cycle at 60°C/1'). Fluorescence on chips was revealed by the QuantStudio™ 3D Instrument, and collected raw data were analyzed by the QuantStudio™ 3D AnalysisSuite™ Software. The software assesses reliable data and displays quality indicators, based upon loading, signal, and noise characteristics, for each chip. This quality control is based on the number of partitions that exceed the total number of wells filled correctly, and to get a precise quantification, we settled a threshold of 10,000 data points with a manually fixed quality threshold over 0.6. Performance experiments conducted for both assays determined the appropriate fluorescence thresholds in FAM relative fluorescence units (RFU) to discriminate between positive calls (blue dots) and no target control calls (yellow dots) (**Figure 1**): GNAS_27,895_mu RFU > 3,000 and GNAS_27,887_mu RFU > 5,000. The output data are reported as copies/µl (cpm) detected on the chip by the instrument that represents the number of observed mutant alleles in the reagent mixture plus DNA. The absolute quantification of mutated alleles can be deduced from the cpm by applying the following formula: cpm X 0.0033 ng (mass of the human genome) X 3.34 (dilution factor for 15 µl of sample plus reaction mixture) = ng/µl.

The specificity and the sensitivity of the dPCR technique were assessed testing NTCs, WTCs, and MTCs, previously identified by Sanger sequencing, and a calibrator curve, in triplicate, of serially diluted mutated samples (representing 100% MTCs), ranging from 50% to 3% and from 25% to 1.5% of c.604 or c.605 heterozygously mutated samples, respectively, representing the relative mutation abundance (RMA). Cutoff values, or rather copies of mutated DNA per 1 µl of the tested sample, cpm, to diagnose the presence of a mutation and the RMA in each patient's sample were then inferred by comparing observed cpm values with those detected by the calibrator curve.

RESULTS

The bulk of experiments performed allowed to obtain a molecular diagnosis in MAS/FD subjects that were still missing because of the low detection rate of the Sanger sequencing, thanks to the setup of the dPCR, a highly sensitive and specific technique to detect low-abundance somatic single-nucleotide *GNAS* alterations. To validate results produced with the dPCR, we also

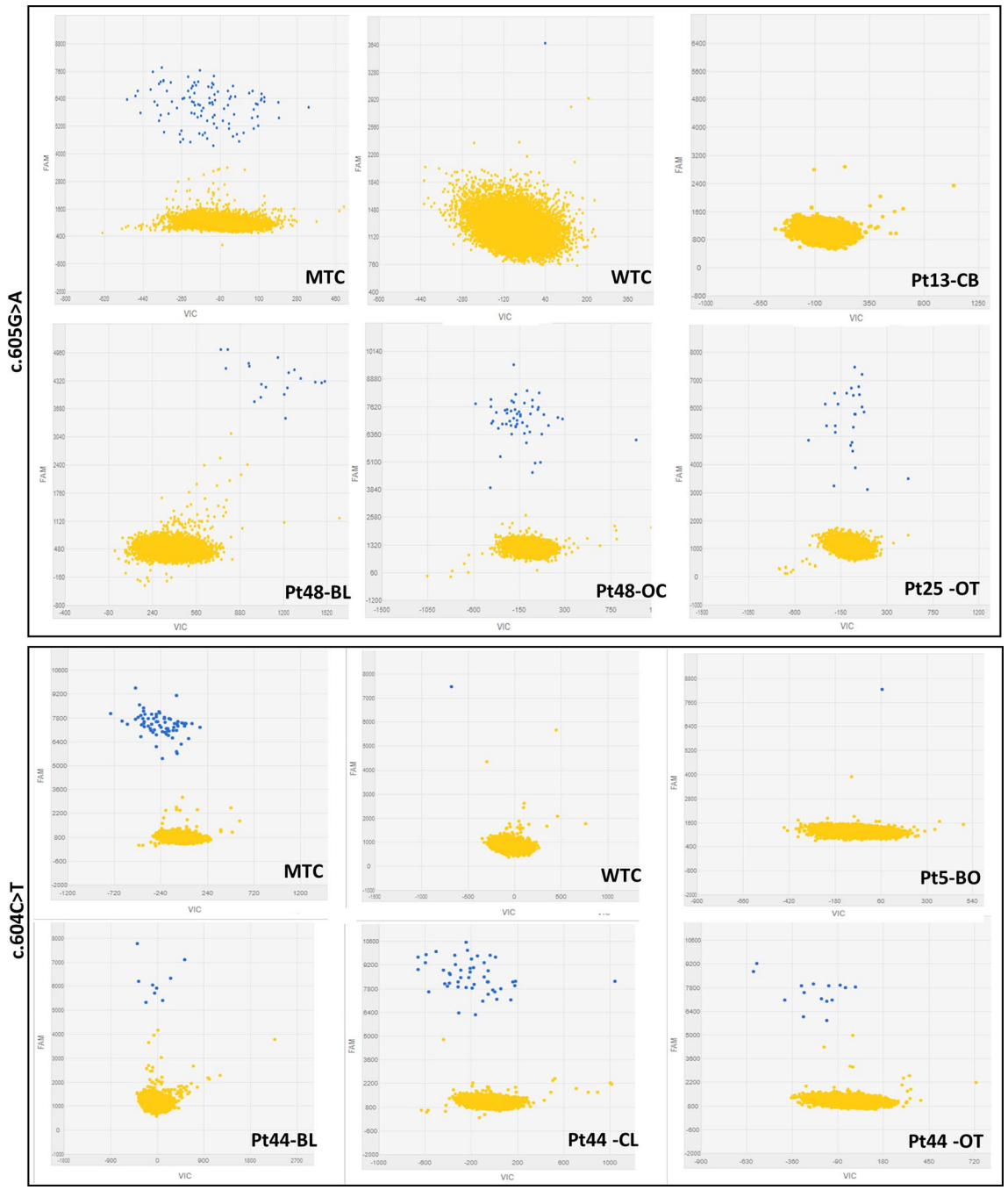

FIGURE 1 | Representative 2-D scatterplots obtained from the raw imaging data analysis of rare mutation detection analysis using the QuantStudio™ 3D AnalysisSuite™ Software (upper panel: GNAS_27895_mu assay; lower panel: GNAS_27877_mu assay). Blue dots represent the FAM™ reporter dye signal that allows to identify and count mutant alleles present in the nanoscale-sized reaction wells of the chip. Yellow dots are the assay control signal and represent the absence of FAM™ signal in part of the nanoscale-sized reaction wells of the chip.

compared this novel approach with additional available methods to search for somatic genetic alterations, the AS-PCR and the COLD-MAMA PCR.

We started our investigation with a careful analysis of Sanger electropherograms obtained by both calibrator curves and MAS/FD patients, showing that the mutated allele was clearly detectable only in case of samples containing elevated amounts of mutated DNA (LOD > 0.1 ng/μl) but only conceivable for lower RMAs (**Figure 2**; **Table 2**). As a matter of fact, the high background noise was a side effect that made it impossible to discriminate real calls from nonspecific signals, thus leading to the failure to establish patients' genotype conclusively.

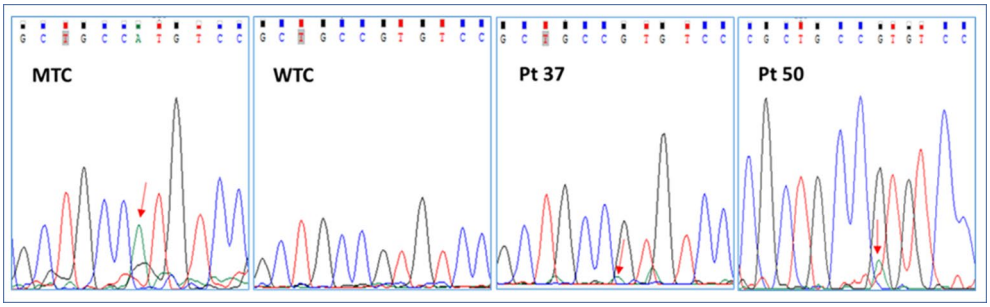

FIGURE 2 | Representative electropherograms of the GNAS c.605G > A mutation screening by Sanger sequencing of a mutant control (MTC), a wild-type control (WTC), and selected patients (pt37-BO-RJ and pt50-OC). Red arrows highlight the position of the c.605A nucleotide.

In particular, the Sanger sequencing detected somatic mutations only in 6 of 79 tested DNA samples (5/15 BO, 33.3%; 0/16 CB/FB, 0%; 1/11 OC/OCL/OT, 9%; 0/37 BL, 0%) belonging to 4 of 54 different patients (7.4% detection rate) (**Supplementary Table 2**; **Tables 1** and **2** panel A). No patient affected by the c.604C > T variant was found. On the other hand, four patients (pt2-BO, pt3-BO, pt37-BO-LJ+M+RJ, and pt46-OC) obtained a positive diagnosis for the c.605G > A variant. The pathogenetic variant was found by Sanger only in DNA samples extracted from affected biopsies. These data further highlighted the technical limitations in performing a rare mutation investigation by Sanger in DNA samples from whole blood.

The use of allele-specific primers allowed to slightly improve the detection threshold of MAS/FD-related variants (LOD > 0.03 ng/μl) as it discovered 13 samples/66 with the c.605A variant, belonging to 7 of 46 different patients: pt2-BO, pt3-BO, pt28-BL, pt37- BO-LJ (left jaw)+RJ (right jaw), pt52-BL+CB, pt46-OC, and pt48-BL+OC), but the sensitivity of the AS-PCR (19.6% detection rate) was still too low to perform an accurate molecular diagnosis of MAS in most patients (**Supplementary Table 2**; **Tables 1** and **2** panel A; **Figure 3**). All Sanger-positive samples were confirmed, and three novel positive patients were found. It is noteworthy that, contrary to Sanger, this method allowed to discover a mutation even in three blood samples (5/10 BO, 40%; 1/16 CB/FB, 6.25%; 4/10 OC/OCL/OT, 40%; 3/33 BL, 9.1%) (**Table 2** panel A). No false-positive/negative results were obtained, and the interindividual and intraindividual, or rather different tissues from the same patient, differences in RMAs were the same as observed by dPCR, thus excluding an artifact or carryover contamination.

Forty samples from 24 patients were prepared for the COLD-MAMA PCR but, because of the insufficient amount of available DNA, only 24 samples from 14 patients for the c.604C > T and 26 samples from 14 patients for the c.605G > A were successfully analyzed, demonstrating the presence of the T-mutated allele in 2/24 samples (pt41-BL and pt44-OCL) and of the A-mutated allele in 12/26 samples (pt6-BO, pt37-BL+BO-LJ+RJ, pt46-OC+OT, pt47-OT, pt48BL+OC, pt49-OC, pt50-OC+OCL), showing the presence of a MAS/FD causative genetic variant in 9 of 14 patients with an overall detection rate of 64.3% (3/3 BO, 100%; 0/2 CB/FB, 0%; 7/10 OC/OCL/OT, 70%; 2/11 BL, 18.1%) and a LOD > 0.02 ng/μl (**Supplementary Table 2**; **Tables 1** and **2** panel A).

The first phase of dPCR data analysis, performed using the QuantStudio™ 3D AnalysisSuite™ Software, relied on manual examination and thresholding 2-D scatterplots from control samples (NTCs, WTCs, and MTCs) (**Figure 1**). Neither false-positive nor false-negative was observed among tested WTCs (c.604T n = 9, c.605A n = 13) and MTCs (c.604T n = 3, c.605A n = 6), confirming the good performance of both assays. Thus, the dPCR allowed to discriminate all tested MTCs from WTCs.

From calibrator curves of serially diluted MTCs, mimicking various RMAs, we determined the LOD of 0.01 ng/μl, reference cpm values were used as cutoffs for inferring the presence of a *GNAS* somatic mutation in patients and the RMA in each patient's sample (**Figure 4**). According to the *t* test (95% CI), all points of the serial curve resulted differently from 0% RMA samples, and statistically significant cutoffs to detect the positiveness of a sample were 0.391 ± 0.074 (P = 0.0028) for the c.604T and 0.464 ± 0.012 (P = 0.0004) for the c.605A alleles. Note that TaqMan® Mutation Detection Assays showed distinct lower detection thresholds, 3% RMA for the GNAS_27887_mu (p.Arg202Cys) and 1.5% RMA for the GNAS_27895_mu (p.Arg202His), but the best fitting explanation for this slightly contrasting performance was the use of DNA serial dilutions prepared from MTCs with a different starting concentration and, consequently, a different RMA (0.13 ng/μl and RMA 12.11 ± 0.023 and 0.19 ng/μl and RMA 17.47 ± 0.177, respectively) rather than a different conduct of assays or a setup failure.

During the analysis of patient samples, the short way to call a positive sample was by visualizing the 2-D plot and comparing cutoff cpm values, considering nonoverlapping error bars, between WTCs and unknown samples, whereas a *t*-test was used to determine the correct RMA subcluster attribution of positive samples.

The dPCR allowed identification of the mutation in 36/79 samples (10/15 BO, 66.6%; 3/16 CB/FB, 18.75%; 10/11 OC/OCL/OT, 90.9%; 14/37 BL, 37.8%) in 23/54 patients (42.6% detection rate), of whom 2 patients (pt41-BL and pt44-BL+OCL+OT) carried the c.604T variant and 21 patients (pt1-BO, pt2-BO, pt3-BO, pt5-BO, pt6-BO, pt12-BO, pt19-BL, pt22-BL, pt25-BL, pt28-BL, pt51-BL, pt53-BL, pt37-BL+BO, pt39-BL+CB, pt52-BL+CB, pt43-BL+FB, pt46-OC+OT, pt47-OC+OT, pt48-BL+OC, pt49-BL+OC, and pt50-OC+OCL) the c.605A variant (**Supplementary Table 2**; **Tables 1** and **2** panel A). In particular

TABLE 1 | Table resuming molecular data of the studied cohort of MAS patients. Cpm, copies/µl. Abs quant, absolute quantification.

		c.604C>T								c.605G>A							
PT ID	**TISSUE**	**SANGER**	**AS-PCR**	**dPCR**	**dPCR cpm**	**dPCR absquant (ng/µl)**	**dPCR RMA**	**COLD-MAMA PCR**	**COLD-MAMA RMA**	**SANGER**	**AS-PCR**	**dPCR**	**dPCR cpm**	**dPCR absquant (ng/µl)**	**dPCR RMA**	**COLD-MAMA PCR**	**COLD-MAMA RMA**
1	**BO**	WT	WT	WT	0.181	0.00				WT	WT	**MUT**	**0.892**	**0.01**	**>6**		
2	**BO**	WT	WT	WT	0.000	0.00				**MUT**	**MUT**	**MUT**	**8.712**	**0.10**	**>50**		
3	**BO**	WT	WT	WT	0.095	0.00				**MUT**	**MUT**	**MUT**	**17.626**	**0.19**	**100**		
4	**BO**	WT	WT	WT	0.000	0.00				WT	WT	WT	0.169	0.00			
5	**BO**	WT	nd	WT	0.000	0.00		nd		WT	nd	**MUT**	**6.594**	**0.07**	**50**	nd	
6	**BO**	WT	WT	WT	0.000	0.00		WT		WT	**MUT**	**MUT**	**3.604**	**0.04**	**25**	**MUT**	**3**
7	**BO**	WT	WT	WT	0.000	0.00				WT	WT	WT	0.000	0.00			
8	**BO**	WT	WT	WT	0.086	0.00				WT	WT	WT	0.353	0.00			
12	**BO**	WT	nd	WT	0.260	0.00		nd		WT	nd	**MUT**	**3.196**	**0.04**	**25**	nd	
9	**BO**	WT	nd	WT	0.255	0.00		nd		WT	nd	WT	0.084	0.00		nd	
	BL	WT	WT	WT	0.166	0.00		WT		WT	WT	**MUT**	**0.622**	**0.01**	**>6**	**MUT**	**6**
	BO-LJ	WT	WT	WT	0.195	0.00		WT		**MUT**	**MUT**	**MUT**	**32.654**	**0.36**	**100**	**MUT**	**>40**
37	**BO-M**	WT	nd	WT	0.000	0.00		nd		**MUT**	nd	**MUT**	**22.108**	**0.24**	**100**	nd	
	BO-RJ	WT	WT	WT	0.169	0.00		WT		WT	**MUT**	**MUT**	**2.121**	**0.02**	**<25**	**MUT**	**3**
	BO-S	WT	nd	WT	0.188	0.00		nd		**MUT**	nd	**MUT**	**22.904**	**0.25**	**100**	nd	
45	**BL**	WT	WT	WT	0.087	0.00				WT	WT	WT	0.312	0.00			
	BO	WT	WT	WT	0.092	0.00				WT	WT	WT	0.105	0.00			
10	**CB**	WT	WT	WT	0.152	0.00				WT	WT	WT	0.343	0.00			
11	**CB**	WT	WT	WT	0.237	0.00				WT	WT	WT	0.436	0.00			
13	**CB**	WT	WT	WT	0.105	0.00				WT	WT	WT	0.100	0.00			
14	**CB**	WT	WT	WT	0.000	0.00				WT	WT	WT	0.185	0.00			
	BL	WT	WT	WT	0.000	0.00				WT	WT	WT	0.178	0.00			
38	**CB-WT**	WT	WT	WT	0.183	0.00				WT		WT	0.233	0.00			
	CB-YT	WT	WT	WT	0.105	0.00				WT		WT	0.416	0.00			
	CB-GT	WT	WT	WT	0.173	0.00				WT		WT	0.191	0.00			
39	**BL**	WT	WT	WT	0.319	0.00				WT	WT	**MUT**	**0.509**	**0.01**	**6**		
	CB	WT	WT	WT	0.000	0.00				WT	WT	**MUT**	**1.251**	**0.01**	**<12.5**		
52	**BL**	WT	WT	WT	0.112	0.00		nd		WT	**MUT**	**MUT**	**0.598**	**0.01**	**6**	nd	
	CB	WT	WT	WT	0.000	0.00		nd		WT	**MUT**	**MUT**	**6.208**	**0.07**	**50**	nd	
40	**BL**	WT	WT	WT	0.000	0.00		WT		WT	WT	WT	0.096	0.00		WT	
	CB	WT	WT	WT	0.320	0.00		nd		WT	WT	WT	0.139	0.00		WT	
	BL	WT	WT	**MUT**	**2.249**	**0.02**	**<12.5**	**MUT**	**<2.5**	WT	WT	WT	0.254	0.00		WT	
41	**CB-GT**	WT	WT	WT	0.104	0.00				WT	WT	WT	0.078	0.00			
	CB-WT	WT	WT	WT	0.000	0.00		nd		WT	WT	WT	0.336	0.00		nd	
	CB-YT	WT	WT	WT	0.000	0.00				WT	WT	WT	0.180	0.00			
42	**BL**	WT	WT	WT	0.223	0.00		WT		WT	WT	WT	0.083	0.00		WT	
	FB	WT	WT	WT	0.265	0.00		WT		WT	WT	WT	0.190	0.00		WT	
43	**BL**	WT	WT	WT	0.188	0.00				WT	WT	**MUT**	**0.873**	**0.01**	**>6**		
	CB	WT	WT	WT	0.000	0.00				WT	WT	WT	0.089	0.00			
	FB	WT	WT	WT	0.000	0.00				WT	WT	**MUT**	**1.633**	**0.02**	**>12.5**		

(Continued)

TABLE1 | Continued

		c.604C>T								c.605G>A							
PT ID	**TISSUE**	**SANGER**	**AS-PCR**	**dPCR**	**dPCR cpm**	**dPCR absquant (ng/µl)**	**dPCR RMA**	**COLD-MAMA PCR**	**COLD-MAMA RMA**	**SANGER**	**AS-PCR**	**dPCR**	**dPCR cpm**	**dPCR absquant (ng/µl)**	**dPCR RMA**	**COLD-MAMA PCR**	**COLD-MAMA RMA**
35	**OT**	WT	nd	WT	0.000	0.00				WT	nd	WT	0.244	0.00			
	BL	WT	WT	**MUT**	**0.777**	**0.01**	<3	WT		WT	WT	WT	0.293	0.00		WT	
44	**OCL**	WT	WT	**MUT**	**4.995**	**0.06**	**<25**	**MUT**	**<2.5**	WT	WT	WT	0.256	0.00		WT	
	OT	WT	WT	**MUT**	**1.485**	**0.02**	>6	WT		WT	WT	WT	0.184	0.00		WT	
46	**OC**	WT	WT	WT	0.083	0.00		WT		**MUT**	**MUT**	**MUT**	**2.969**	**0.03**	**<25**	**MUT**	**10**
	OT	WT	WT	WT	0.182	0.00		WT		WT	WT	**MUT**	**0.520**	**0.01**	**6**	**MUT**	**<0.5**
	BL	WT	WT	WT	0.000	0.00		WT		WT	WT	WT	0.299	0.00		WT	
47	**OC**	WT	WT	WT	0.102	0.00		WT		WT	WT	**MUT**	**1.135**	**0.01**	**<12.5**	WT	
	OT	WT	WT	WT	0.084	0.00		WT		WT	WT	**MUT**	**1.164**	**0.01**	**<12.5**	**MUT**	**3**
48	**BL**	WT	WT	WT	0.334	0.00		WT		WT	**MUT**	**MUT**	**5.347**	**0.06**	**<50**	**MUT**	**16**
	OC	WT	WT	WT	0.000	0.00		WT		WT	**MUT**	**MUT**	**10.169**	**0.11**	**>50**	**MUT**	**>40**
49	**BL**	WT	WT	WT	0.101	0.00		WT		WT	WT	**MUT**	**0.465**	**0.01**	**3**	WT	
	OC	WT	WT	WT	0.192	0.00		WT		WT	WT	**MUT**	**1.578**	**0.02**	**>12.5**	**MUT**	**26**
50	**OC**	WT	WT	WT	0.412	0.00		WT		WT	**MUT**	**MUT**	**3.312**	**0.04**	**25**	**MUT**	**<0.5**
	OCL	WT	WT	WT	0.205	0.00		WT		WT	**MUT**	**MUT**	**7.275**	**0.08**	**>50**	**MUT**	**2.5**
15	**BL**	WT	WT	WT	0.000	0.00				WT	WT	WT	0.191	0.00			
16	**BL**	WT	WT	WT	0.115	0.00				WT	WT	WT	0.000	0.00			
17	**BL**	WT	WT	WT	0.154	0.00				WT	WT	WT	0.264	0.00			
18	**BL**	WT	WT	WT	0.101	0.00				WT	WT	WT	0.000	0.00			
19	**BL**	WT	WT	WT	0.117	0.00		WT		WT	WT	**MUT**	**0.815**	**0.01**	**>6**	WT	
20	**BL**	WT	WT	WT	0.097	0.00				WT	WT	WT	0.119	0.00			
21	**BL**	WT	WT	WT	0.089	0.00				WT	WT	WT	0.308	0.00			
22	**BL**	WT	WT	WT	0.000	0.00				WT	WT	**MUT**	**0.644**	**0.01**	**6**		
23	**BL**	WT	WT	WT	0.000	0.00				WT	WT	WT	0.219	0.00			
24	**BL**	WT	WT	WT	0.000	0.00				WT	WT	WT	0.000	0.00			
25	**BL**	WT	WT	WT	0.000	0.00		nd		WT	WT	**MUT**	**2.252**	**0.02**	**>12.5**	nd	
26	**BL**	WT	WT	WT	0.026	0.00				WT	WT	WT	0.000	0.00			
27	**BL**	WT	WT	WT	0.000	0.00				WT	WT	WT	0.000	0.00			
28	**BL**	WT	WT	WT	0.000	0.00		nd		WT	**MUT**	**MUT**	**2.170**	**0.02**	**>12.5**	nd	
29	**BL**	WT	WT	WT	0.000	0.00				WT	WT	WT	0.285	0.00			
31	**BL**	WT	WT	WT	0.095	0.00		WT		WT	WT	WT	0.383	0.00		WT	
32	**BL**	WT	WT	WT	0.277	0.00		nd		WT	WT	WT	0.000	0.00		WT	
33	**BL**	WT	WT	WT	0.000	0.00				WT	WT	WT	0.190	0.00			
34	**BL**	WT	WT	WT	0.000	0.00				WT	WT	WT	0.000	0.00			
51	**BL**	WT	nd	WT	0.000	0.00		nd		WT	nd	**MUT**	**1.216**	**0.01**	**12.5**	nd	
53	**BL**	WT	nd	WT	0.258	0.00		nd		WT	nd	**MUT**	**1.043**	**0.01**	**<12.5**	nd	
54	**BL**	WT	nd	WT	0.000	0.00		nd		WT	nd	WT	0.246	0.00		nd	
30	**BL**	WT	nd	WT	0.000	0.00		nd		WT	nd	WT	0.171	0.00		nd	
36	**BL**	WT	WT	WT	0.150	0.00				WT	WT	WT	0.202	0.00			

RMA, relative mutation amount; BO, bone; BL, blood; LJ, left jaw; M, maxilla; RJ, right jaw; S, symphysis; CB, cutaneous biopsy; FB, fibroblast; OCL, ovarian cyst liquid; OC, ovarian cyst; OT, ovarian tissue; nd, not determined; WT, wild-type. Bold data highlight results obtained in mutated patients.

TABLE 2 | Tables resuming the detection rate of MAS-related c.604T and c.605A *GNAS* activating variants.

	Sanger	AS-PCR	COLD-MAMA PCR	dPCR
A				
BO	5/15; **33.3%**	5/10; **50%**	3/3; **100%**	10/15; **66.6%**
CB/FB	0/16; **0%**	1/16; **6.25%**	0/2; **0%**	3/16; **18.75%**
OC/OCL/OT	1/11; **9%**	4/10; **40%**	7/10; **70%**	10/11; **90.9%**
BL	0/37; **0%**	3/33; **9.1%**	2/11; **18.1%**	14/37; **37.8%**
Patients (overall DR)	4/54; **7.4%**	9/46; **19.6%**	9/14; **64.3%**	23/54; **42.6%**
B				
BO	0/3; **0%**	3/3; **100%**	3/3; **100%**	3/3; **100%**
CB/FB	0/2; **0%**	0/2; **0%**	0/2; **0%**	0/2; **0%**
OC/OCL/OT	1/10; **10%**	4/10; **40%**	7/10; **70%**	10/10; **100%**
BL	0/11; **0%**	1/11; **9%**	2/11; **18.1%**	6/11; **54.5%**
Patients (overall DR)	2/14; **14.3%**	5/14; **35.7%**	9/14; **64.3%**	10/14; **71.4%**

BO, bone; CB/FB, cutaneous biopsy/fibroblasts; OC/OCL/OT, ovarian cyst/ovarian cyst liquid/ovarian tissue; BL, blood; DR, detection rate.
Panel A reports all patients whereas panel B reports only patients tested by all of the four available techniques. Percentages values are highlighted in bold.

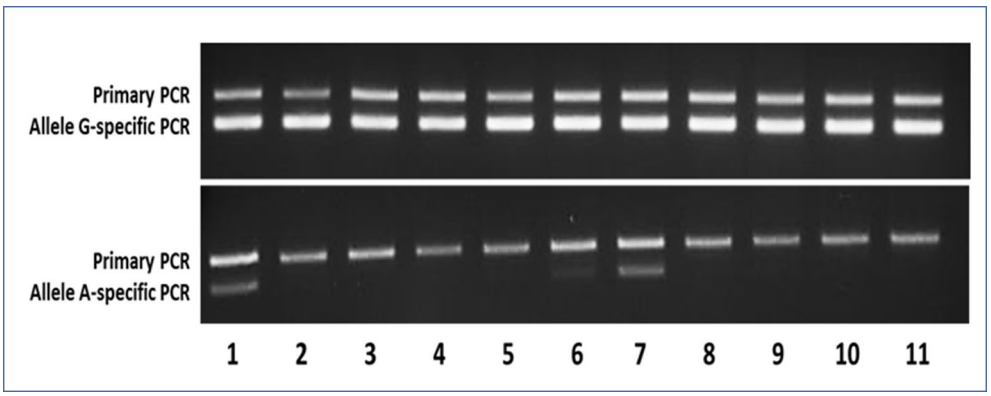

FIGURE 3 | Representative agarose gels of the AS-PCR for the detection of the c.605G > A variant showing both the aspecific primary PCR on top and the allele-specific nested PCR below amplicon bands. In the upper gel, as expected, the G-specific amplicon was appreciable for each sample, whereas in the lower gel, the A-specific amplicon was detectable only in samples bearing the mutant A allele. Samples 1–9 were MAS patients (1–2 = pt46-OC+OT; 3–5 = pt47-BL+OC+OT; 6–7 = pt48-BL+OT; 8–9 = pt49-BL+OT), whereas samples 10 and 11 were WTCs.

and importantly, the technique was able to identify a mutation in 37.8% of peripheral blood samples.

The comparison of data obtained in seven mutated patients from whom we collected both blood and affected tissues (pt37, pt39, pt43, pt 44, pt48, pt49, and pt52) ensured the ability of dPCR to detect MAS/FD-associated alleles in gDNA extracted from blood samples with an RMA sufficient to be ascertained by the system.

Curiously, patient 41's cutaneous biopsy resulted negative for alterations, but we found the mutation in two independently collected blood samples by both dPCR and COLD-MAMA PCR. In patient 43, we found the mutated allele both in the blood and in cultured fibroblasts but not in the DNA extracted from the cutaneous biopsy (**Supplementary Table 2**; **Table 1**).

DISCUSSION

The work presented in this paper aimed at setting up a novel technology, the dPCR, to perform the molecular diagnosis of MAS/FD by the identification of rare somatic activating variants of the *GNAS* gene. The detection of low-level somatic variants in MAS/FD patients is a challenge for the laboratory because it relies on the discrimination between two highly similar sequences, of which the wild-type one is significantly more abundant in the sample. Molecular procedures currently used in our laboratory, such as Sanger sequencing and AS-PCR, do not reach the needed sensibility, thus failing to detect rare mutations. For this reason, most MAS/FD patients often lack a molecular confirmation of their disease.

Up to now, scientists developed several methods to overcome technical limitations associated with performing a successful genetic diagnosis of MAS/FD (i.e., enrichment methods with subsequent restriction enzyme digestion, RFLP, or with peptide nucleic acid probes, PNA, and next-generation sequencing, NGS, with or without PNA) (Schwindinger et al., 1992; Liang et al., 2011; Hannon et al., 2003). Selective enrichment methods reached good levels in low-abundance variant identification, but their execution is particularly cumbersome and time-consuming, as the requirement of a large number of PCR cycles for the selective amplification of the mutated allele leads to an elevated risk to create PCR artifacts and cross-contamination. An opportunity to overcome these limitations would be an approach such as the

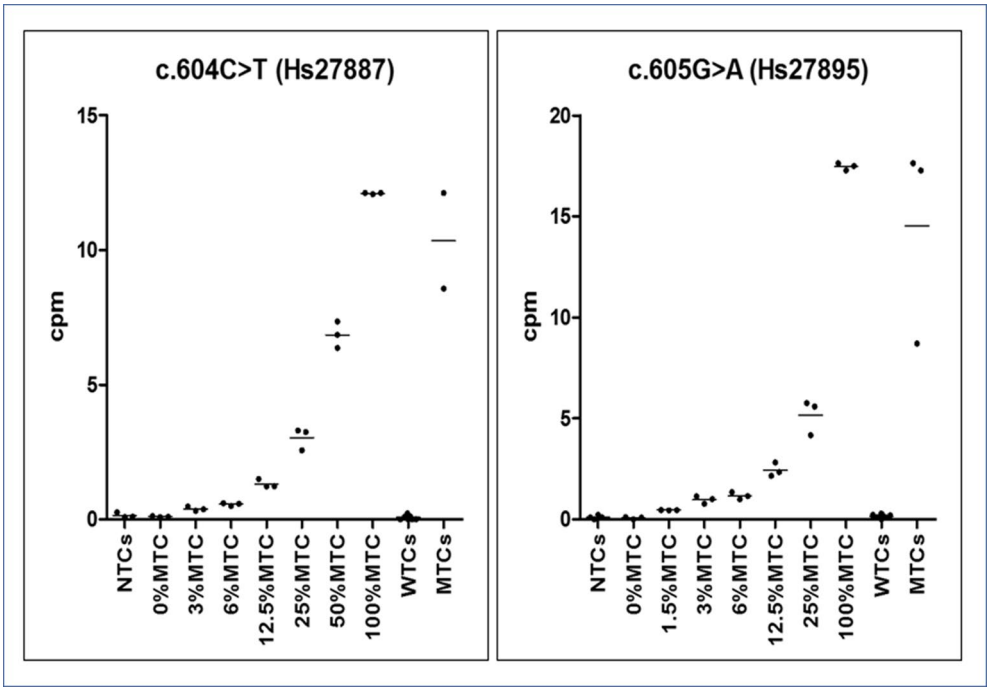

FIGURE 4 | Graphs of dPCR cpm calibrator curves. On the Y-axis, the copies per µl (cpm) of mutated allele are reported, whereas on the X-axis, the sample type (no template ctrls - NTCs. wild-type template ctrls - WTCs and mutated template ctrls - MTCs). Observed reference cpm (mean of triplicates ± SD/% RMAs) were 0.102 ± 0.017/0%, 0.391 ± 0.074/3%, 0.563 ± 0.052/6%, 1. 321 ± 0.157/12.5%, 3.035 ± 0.411/25%, 6.855 ± 0.494/50%, and 12.11 ± 0.023/100% for the c.604T variant and 0.068 ± 0.06/0%, 0.464 ± 0.012/1.5%, 0.976 ± 0.189/3%, 1.171 ± 0.181/6%, 2.44 ± 0.346/12.5%, 5.17 ± 0.873/25%, and 17.47 ± 0.177/100% for the c.605A variant. All points of the serial dilution resulted significantly different (CI 95%) from the 0%MTC.

deep-sequencing NGS (depth of sequence coverage approximately 800X) with molecular bar code technique, although it is more expensive and still far from clinical settings.

During the setup phase, the dPCR allowed to discriminate all tested WTCs and MTCs. It showed an extremely high sensitivity, being able to detect even few copies of mutated alleles in DNA samples extracted from different tissues, including blood samples.

To validate the dPCR approach, we analyzed our samples also by Sanger sequencing, AS-PCR, and COLD-MAMA PCR. Positive results obtained by Sanger, AS-PCR, and COLD-MAMA PCR were in full agreement with those by dPCR.

As expected, the direct sequencing of *GNAS* exon 8, confirmed to be the less sensitive method, was able to reach in our cohort only a 7.4% detection rate, which means in only 4 of the 54 analyzed patients. Mutated alleles were found exclusively in a small subset of biopsies, mainly from bone, whereas no mutation was detected in blood samples. The AS-PCR allowed to slightly improve the overall detection rate, raising it to 19.6% (7 of 46 analyzed patients) and to identify three positive patients directly from the blood-extracted DNA, two of them confirmed in the affected tissue as well. The COLD-MAMA PCR reached an overall 64.3% detection rate (9 of 14 analyzed patients) but needed, as the AS-PCR, a large amount of starting material with respect to the dPCR that limited the opportunity to test all of the patients of our cohort and to compare final diagnoses. This method identified almost all mutated tissue samples but in only 2 of the 11 blood-extracted samples (18.1% detection rate). Finally, the dPCR obtained an overall 42.6% detection rate by finding 23 mutated patients of the 54 tested ones. In particular, this latter technique discovered low amounts of mutated alleles in 14/37 blood samples, with a detection rate of 37.8%. (**Table 2** panel A). We consider noteworthy to show the success rate of the dPCR because of the fact that this technique allowed to find false-negative (FN) patients by the other methods (19 Sanger FN, 10 AS-PCR FN, and 1 COLD-MAMA PCR FN).

In the present work, we investigated the two most common mutations affecting the amino acid residue R202 (previously reported as R201), the p.R202H and p.R202C, that account for most of diagnosed MAS-associated genetic defects. In ultra-rare cases, additional substitutions have been reported at position 202, p.R202S/L/G, and at position 228, p.Q228L/R/K/H (Candeliere et al., 1997; Riminucci et al., 1999; Lietman et al., 2005; Idowu et al., 2007). The existance of such exceptional variants and the possibility to discover novel activating GNAS defects might explain for the portion of MAS patients still missing a confirmative molecular diagnosis that includes both true- and false-negative patients.

To compare the efficiency of the different analytical methods in detecting mutated MAS/FD patients, we analyzed separately the cluster of samples and patients investigated by all four of them and determined the detection rate for patients and by single sample type. This analysis showed an increase in sensitivity leading to the improvement of the detection rate in the considered subcohort of 14 cases (14.3% for Sanger, 35.7% AS-PCR, 64.3%

for COLD-MAMA PCR, and 71.4% for dPCR), indicating dPCR as the most performing approach. When we considered the testing of specific tissues, we observed that all techniques, but Sanger, found the mutated allele in bone samples, whereas in ovarian tissues and blood, we noticed a progressive improvement in finding rare variants from the lowest by Sanger to the highest by dPCR (**Table 2** panel B). Moreover, the dPCR approach showed to be less susceptible to poor sample quality and to the presence of inhibitors because it demonstrated to be able to successfully analyze all of the cohort of samples, including very small amounts of DNA after long-term storage and FFPE-extracted DNAs. In particular, the other methods failed in some cases, which were reported as "not determined" in **Table 2**, because of the reduced amount or the poor quality of material available for testing.

Detecting mutated alleles from blood sample was the final and more important goal of our study because we wanted to obtain an analytical method for the noninvasive testing of MAS/FD patients, overcoming the need for invasive procedures like biopsies. Checking the detection rates in the subcohort of DNAs from peripheral blood, dPCR achieved the highest performance with respect to the other techniques (54.5% dPCR versus 18.1% COLD-MAMA PCR versus 9% AS-PCR versus 0% Sanger) (**Table 2** panel B). In seven patients, results were cross-confirmed by the genotype found in tissue specimens (pt37, pt39, pt43, pt44, pt48, pt49, and pt52) (**Supplementary Table 2**; **Table 1**). We agree with Narumi and colleagues' (2013) observation that the relative mutation abundance in blood is strikingly variable, it does not reflect distribution and extension of affected lesions, and it does not correlate with disease severity and progression, as already observed in other congenital syndromes caused by somatic genetic variants (i.e., Proteus and megalencephaly syndromes) (Lindhurst et al., 2011; Rivière et al., 2012; Narumi et al., 2013).

Our calibrator curves, being derived from a serial dilution of mutated patients and not cloned DNA, allowed a relative rather than an absolute quantification of the RMA in samples, and we used them to define reference cutoff cpm values to discriminate between wild-type and mutated samples. We do not consider this fact as a limit of the present study that was aimed at improving the number of successful molecular diagnosis of MAS/FD because the absolute RMA showed no real utility from a clinical point of view and it did not affect the dPCR performance because this method carries out an absolute quantitation of nucleic acid samples based on PCR amplification and absolute count of single template molecules.

Data from previously published papers showed that the dPCR is an extremely competitive method. For example, Lumbroso and colleagues (2004) studied a very big series of MAS patients and determined the frequency of observed mutated patients in several different tissues by RFLP analysis. The dPCR reached a better performance in the blood (37.8% dPCR versus 21% RFLP) and ovarian tissues (90.9% by dPCR versus 65% by RFLP) (Lumbroso et al., 2004).

If we consider ours and Lumbroso's data on tissue samples, 10 of 15 (66.6%) and 9 of 11 (82%) bone biopsies, 2 of 14 (14.3%) and 3 of 11 (27%) cutaneous biopsies, 1 of 2 (50%) and 0 fibroblasts, 5 of 5 (100%) and 3 of 8 (38%) ovarian cysts, 2 of 2 (100%) and 13 of 19 (68%) ovarian cyst liquids and 3 of 4 (75%) and 10 of 13 (77%) ovarian tissues were positive by RFLPs and dPCR, respectively (Narumi et al., 2013). Even if ours and Lumbroso's are different techniques, this comparison allowed to appreciate the different detection rates of mutated alleles related to the specific tested tissue and to confirm, in accordance with Vasilev and colleagues (2014), that tissues atypically/rarely involved in MAS may have the *GNAS* mutation in mosaic distribution (Narumi et al., 2013; Vasilev et al., 2014).

Bone was the best tissue to identify MAS/FD rare variants, which were found both in patients showing isolated FD and the classic MAS triad. In both our samples of ovarian cystic fluid, we identified the mutated allele, confirming its usefulness also in case of isolated precocious puberty. The systematic study of the ovarian cystic fluid would be advisable for its major implications for both disease progression and treatment.

Our data are also in line with the observation that some samples may remain apparently negative because the biopsy missed loci where mutated cells are confined, and that specimens like skin presented an additional difficulty because of the low proportion of melanocytes (Weinstein et al., 1991; Schwindinger et al., 1992; Lumbroso et al., 2004; Vasilev et al., 2014). In our series, this hypothesis is confirmed by the unexpected results obtained in two intriguing cases presenting discrepancies between blood and cutaneous biopsies. In patient 41, the cutaneous biopsy resulted in negative but blood-extracted DNA was positive for a *GNAS* mutation. An artifact was excluded as we confirmed this positive result in two independently collected and extracted blood samples by both dPCR and COLD-MAMA PCR. The best explanation for this finding is the coring of a tissue mainly composed of normal cells. This is further supported by our findings in patient 43 where a mutated allele was detected both in blood and in cultured fibroblasts but not in the skin.

This work included a wide and heterogeneous case series, reflecting different clinical presentations, as well as various sample types. Even if a negative genetic investigation does not rule out the clinical diagnosis of MAS/FD, the detection of pathogenetic variants associated with MAS/FD is fundamental, in particular in partial forms, as it allows early diagnosis.

In conclusion, our data clearly confirm the difficulty encountered in performing a successful molecular diagnosis of *GNAS* somatic activating variants, even from tissue specimens. The strengths of the dPCR approach were 1) being less susceptible to poor quality sample issues (short amplicons), leading to successful amplification also of degraded DNA; 2) being less sensitive to the presence of inhibitors (sample dilution); 3) very low associated risk of cross-contamination (no preamplification cycles required); and 4) being cost-effective (1-day processing time and appropriate cost of reagent). Overall, the dPCR demonstrated to be a sensitive, accurate, and specific analytical tool that we propose should be used as the first-line choice for the molecular diagnosis of MAS/FD.

DATA AVAILABILITY

Data generated and/or analyzed during this study are included in this published article, and they are available from the corresponding author on reasonable request.

ETHICS STATEMENT

Informed consent was obtained from all patients (or legal guardians for minors) and relatives included in the present study. All procedures were performed in compliance with relevant legislation and institutional guidelines and were approved by the IRCCS Fondazione CàGrandaOspedale Maggiore Policlinico institutional committee.

AUTHOR CONTRIBUTIONS

FE conceived and designed the project, analyzed and interpreted data, and was a major contributor in writing the manuscript. MB, IG, MM, AP, and PB acquired and analyzed data. GM conceived and designed the project, followed patients, interpreted data, and was a major contributor in writing the manuscript. LS and MA followed patients and were a minor contributor in writing the manuscript. All authors read and approved the final manuscript.

FUNDING

This work was supported by the Italian Ministry of Health under Grant GR-2009-1608394; Fondazione IRCCS Ca' Granda Policlinico Ospedale Maggiore under Grant Ricerca Corrente Funds.

ACKNOWLEDGMENTS

The authors are members and acknowledge the Euro-Pseudohypoparathyroidism network (EuroPHP) and the EUCID.net (COST action BM1208 on imprinting disorders; www.imprinting-disorders.eu).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2019.00862/full#supplementary-material

REFERENCES

Albright, F., Butler, A. M., Hampton, A. O., and Smith, P. (1937). Syndrome characterized by osteitis fibrosa disseminated, areas of pigmentation and endocrine dysfunction, with precocious puberty in females: report of five cases. *N. Engl. J. Med.* 216, 727–746. doi: 10.1056/NEJM193704292161701

Boyce, A. M., Turner, A., Watts, L., Forestier-Zhang, L., Underhill, A., Pinedo-Villanueva, R., et al. (2017). Improving patient outcomes in fibrous dysplasia/McCune-Albright syndrome: an international multidisciplinary workshop to inform an international partnership. *Arch. Osteoporos.* 12 (1), 21. doi: 10.1007/s11657-016-0271-6

Candeliere, G. A., Roughley, P. J., and Glorieux, F. H. (1997). Polymerase chain reaction-based technique for the selective enrichment and analysis of mosaic arg201 mutations in G alpha s from patients with fibrous dysplasia of bone. *Bone* 21 (2), 201–206. doi: 10.1016/S8756-3282(97)00107-5

de Sanctis, L., Galliano, I., Montanari, P., Matarazzo, P., Tessaris, D., and Bergallo, M. (2017). Combining Real-Time COLD- and MAMA-PCR TaqMan Techniques to Detect and Quantify R201 GNAS Mutations in the McCune-Albright Syndrome. *Horm. Res. Paediatr.* 87 (5), 342–349. doi: 10.1159/000463384

Hannon, T. S., Noonan, K., Steinmetz, R., Eugster, E. A., Levine, M. A., and Pescovitz, O. H. (2003). Is McCune-Albright syndrome overlooked in subjects with fibrous dysplasia of bone? *J. Pediatr.* 142 (5), 532–538. doi: 10.1067/mpd.2003.153

Idowu, B. D., Al-Adnani, M., O'Donnell, P., Yu, L., Odell, E., Diss, T., et al. (2007). A sensitive mutation-specific screening technique for GNAS1 mutations in cases of fibrous dysplasia: the first report of a codon 227 mutation in bone. *Histopathology* 50 (6), 691–704. doi: 10.1111/j.1365-2559.2007.02676.x

Kalfa, N., Philibert, P., Ecochard, A., Hannon, T., Lumbroso, S., and Sultan, C. (2006). Searching for somatic mutations in McCune-Albright syndrome: a comparative study of the peptidic nucleic acid versus the nested PCR method based on 148 samples. *Eur. J. Endocrinol.* 155, 839–843. doi: 10.1530/eje.1.02301

Karadag, A., Riminucci, M., Bianco, P., Cherman, N., Kuznetsov, S. A., Nguyen, N., et al. (2004). A novel technique based on a PNA hybridization probe and FRET principle for quantification of mutant genotype in fibrous dysplasia/McCune-Albright syndrome. *Nucleic Acids Res.* 32 (7), e63. doi: 10.1093/nar/gnh059

Kuznetsov, S. A., Cherman, N., Riminucci, M., Collins, M. T., Robey, P. G., and Bianco, P. (2008). Age-dependent demise of GNAS-mutated skeletal stem cells and "normalization" of fibrous dysplasia of bone. *J. Bone. Miner. Res.* 23 (11), 1731–1740. doi: 10.1359/jbmr.080609

Landis, C. A., Masters, S. B., Spada, A., Pace, A. M., Bourne, H. R., and Vallar, L. (1989). GTPase inhibiting mutations activate the alpha chain of Gs and stimulate adenylyl cyclase in human pituitary tumours. *Nature* 340 (6236), 692–696. doi: 10.1038/340692a0

Liang, Q., Wei, M., Hodge, L. A., Fanburg-Smith, J. C., Nelson, A., Miettinen, M., et al. (2011). Quantitative analysis of activating alpha subunit of the G protein (Gsα) mutation by pyrosequencing in fibrous dysplasia and other bone lesions. *J. Mol. Diagn.* 13 (2), 137–142. doi: 10.1016/j.jmoldx.2010.10.003

Lietman, S. A., Ding, C., and Levine, M. A. (2005). A highly sensitive polymerase chain reaction method detects activating mutations of the GNAS gene in peripheral blood cells in McCune-Albright syndrome or isolated fibrous dysplasia. *J. Bone. Joint. Surg. Am.* 87 (11), 2489–2494. doi: 10.2106/JBJS.E.00160

Lindhurst, M. J., Sapp, J. C., Teer, J. K., Johnston, J. J., Finn, E. M., Peters, K., et al. (2011). A mosaic activating mutation in AKT1 associated with the Proteus syndrome. *N. Engl. J. Med.* 365 (7), 611–619. doi: 10.1056/NEJMoa1104017

Lumbroso, S., Paris, F., Sultan, C., and European Collaborative Study. (2004). Activating Gsalpha mutations: analysis of 113 patients with signs of McCune-Albright syndrome–a European Collaborative Study. *J. Clin. Endocrinol. Metab.* 89 (5), 2107–2113. doi: 10.1210/jc.2003-031225

Matarazzo, P., Lala, R., Andreo, M., Einaudi, S., Altare, F., Viora, E., et al. (2006). McCune-Albright syndrome: persistence of autonomous ovarian hyperfunction during adolescence and early adult age. *J. Pediatr. Endocrinol. Metab.* Suppl 2, 607–617.

McCune, D. J. (1936). Osteitis fibrosa cystica: the case of a nine year old girl who also exhibits precocious puberty, multiple pigmentation of the skin and hyperthyroidism. *Am. J. Dis. Child.* 52, 743–744.

Narumi, S., Matsuo, K., Ishii, T., Tanahashi, Y., and Hasegawa, T. (2013). Quantitative and sensitive detection of GNAS mutations causing McCune-Albright syndrome with next generation sequencing. *PLoS One* 8 (3), e60525. doi: 10.1371/journal.pone.0060525

Qin, L., Wang, J., Tian, X., Yu, H., Truong, C., Mitchell, J. J., et al. (2016). Detection and Quantification of Mosaic Mutations in Disease Genes by Next-Generation Sequencing. *J. Mol. Diagn.* 18 (3), 446–453. doi: 10.1016/j.jmoldx.2016.01.002

Riminucci, M., Fisher, L. W., Majolagbe, A., Corsi, A., Lala, R., De Sanctis, C., et al. (1999). A novel GNAS1 mutation, R201G, in McCune-Albright syndrome. *J. Bone Miner. Res.* 14 (11), 1987–1989. doi: 10.1359/jbmr.1999.14.11.1987

Rivière, J. B., Mirzaa, G. M., O'Roak, B. J., Beddaoui, M., Alcantara, D., Conway, R. L., et al. (2012). De novo germline and postzygotic mutations in AKT3, PIK3R2 and PIK3CA cause a spectrum of related megalencephaly syndromes. *Nat. Genet.* 44 (8), 934–940. doi: 10.1038/ng.2331

Robinson, C., Collins, M. T., and Boyce, A. M. (2016). Fibrous dysplasia/McCune-Albright syndrome: clinical and translational perspectives. *Curr. Osteoporos. Rep.* 14, 178–186. doi: 10.1007/s11914-016-0317-0

Salpea, P., and Stratakis, C. A. (2014). Carney complex and McCune Albright syndrome: an overview of clinical manifestations and human molecular

genetics. *Mol. Cell Endocrinol.* 386 (1-2), 85–91. doi: 10.1016/j.mce.2013.08.022

Schwindinger, W. F., Francomano, C. A., and Levine, M. A. (1992). Identification of a mutation in the gene encoding the subunit of the stimulatory G protein of adenylyl cyclase in McCune-Albright syndrome. *Proc. Natl. Acad. Sci. USA* 89, 5152–5156. doi: 10.1073/pnas.89.11.5152

Sykes, P. J., Neoh, S. H., Brisco, M. J., Hughes, E., Condon, J., and Morley, A. A. (1992). Quantitation of targets for PCR by use of limiting dilution. *Biotechniques* 13 (3), 444–449.

Tessaris, D., Boyce, A. M., Zacharin, M., Matarazzo, P., Lala, R., De Sanctis, L., et al. (2006). Growth hormone-Insulin-like growth factor 1 axis hyperactivity on bone fibrous dysplasia in McCune-Albright Syndrome. *Clin. Endocrinol. (Oxf)* 89 (1), 56–64. doi: 10.1111/cen.13722

Tessaris, D., Corrias, A., Matarazzo, P., De Sanctis, L., Wasniewska, M., Messina, M. F., et al. (2012). Thyroid abnormalities in children and adolescents with McCune-Albright syndrome. *Horm. Res. Paediatr.* 78 (3), 151–157. doi: 10.1159/000342641

Vasilev, V., Daly, A. F., Thiry, A., Petrossians, P., Fina, F., Rostomyan, L., et al. (2014). McCune-Albright syndrome: a detailed pathological and genetic analysis of disease effects in an adult patient. *J. Clin. Endocrinol. Metab.* 99 (10), E2029–E2038. doi: 10.1210/jc.2014-1291

Vogelstein, B., and Kinzler, K. W. (1999). Digital PCR. *Proc. Natl. Acad. Sci. USA* 96 (16), 9236–9241. doi: 10.1073/pnas.96.16.9236

Wei, Z., Shu, C., Zhang, C., Huang, J., and Cai, H. (2017). A short review of variants calling for single-cell-sequencing data with applications. *Int. J. Biochem. Cell Biol.* 92, 218–226. doi: 10.1016/j.biocel.2017.09.018

Weinstein, L. S., Shenker, A., Gejman, P. V., Marino, M. J., Friedman, E., and Spiegel, A. M. (1991). Activating mutations of the stimulatory G protein in the McCune-Albright syndrome. *N. Engl. J. Med.* 325, 1688–1695. doi: 10.1056/NEJM199112123252403

Zhu, Z., Qui, S., Shao, K., and Hou, Y. (2018). Progress and challenges of sequencing and analyzing circulating tumor cells. *Cell Biol. Toxicol.* 34, 405–415. doi: 10.1007/s10565-017-9418-5

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Genetics

CASE REPORT
published: 06 November 2019
doi: 10.3389/fgene.2019.01086

Heterozygous Deletion of the *SHOX* Gene Enhancer in two Females With Clinical Heterogeneity Associating With Skewed XCI and Escaping XCI

*Yixi Sun[1,2,3], Yuqin Luo[1,2,3], Yeqing Qian[1,2,3], Min Chen[1,2,3], Liya Wang[1,2,3], Hongge Li[1,2,3], Yu Zou[4] and Minyue Dong[1,2,3]**

[1] Department of Reproductive Genetics, Women's Hospital, School of Medicine, Zhejiang University, Hangzhou, China, [2] Key Laboratory of Reproductive Genetics, Ministry of Education, Zhejiang University, Hangzhou, China, [3] Key Laboratory of Women's Reproductive Health of Zhejiang Province, Zhejiang University, Hangzhou, China, [4] Department of Diagnostic Radiology, Women's Hospital, School of Medicine, Zhejiang University, Hangzhou, China

OPEN ACCESS

Edited by:
Yueqiu Tan,
Central South University,
China

Reviewed by:
Tsutomu Ogata,
Hamamatsu University School of Medicine,
Japan
Gudrun Rappold,
Heidelberg University,
Germany
Liming Tao,
Broad Institute, United States

***Correspondence:**
Minyue Dong,
dongmy@zju.edu.cn

Specialty section:
This article was submitted to Genetic Disorders, a section of the journal Frontiers in Genetics

Received: 06 June 2019
Accepted: 09 October 2019
Published: 06 November 2019

Citation:
Sun Y, Luo Y, Qian Y, Chen M, Wang L, Li H, Zou Y and Dong M (2019) Heterozygous Deletion of the SHOX Gene Enhancer in two Females With Clinical Heterogeneity Associating With Skewed XCI and Escaping XCI. Front. Genet. 10:1086. doi: 10.3389/fgene.2019.01086

Skewed X-chromosome inactivation (XCI) plays an important role in the phenotypic heterogeneity of X-linked disorders. However, the role of skewed XCI in XCI-escaping gene *SHOX* regulation is unclear. Here, we focused on a heterozygous deletion of *SHOX* gene enhancer with clinical heterogeneity. Using SNP array, we detected that the female proband with Leri-Weill dyschondrosteosis (LWD) carried an 857 kb deletion on Xp22.3 (encompassing *SHOX* enhancer) and a 5,707 kb large-fragment deletion on Xq25q26. XCI analysis revealed that the X-chromosome with the Xq25q26 large-fragment deletion was completely inactivated, which forced the complete activation of the other X-chromosome carrying *SHOX* enhancer deletion. While the Xp22.3 deletion locates on the escaping XCI region, under the combined action of skewed XCI and escaping XCI, transcription of *SHOX* gene was mainly from the activated X-chromosome with *SHOX* enhancer defect, involving in the formation of LWD phenotype. Interestingly, this *SHOX* enhancer deletion was inherited from her healthy mother, who also demonstrated completely skewed XCI. However, the X-chromosome with *SHOX* enhancer deletion was inactivated, and the normal X-chromosome was activated. Combing with escaping XCI, her phenotype was almost normal. In summary, this study was a rare report of *SHOX* gene enhancer deletion in a family with clinical heterogeneity due to skewed inactivation of different X-chromosomes, which can help in the genetic counseling and prenatal diagnosis of disorders in females with *SHOX* defect.

Keywords: *SHOX* gene enhancer, Leri-Weill dyschondrosteosis, skewed X-chromosome inactivation (XCI), clinical heterogeneity, *HUMARA* assay, escaping X-chromosome inactivation (XCI)

INTRODUCTION

The short stature homeobox gene (*SHOX*), locating in the pseudoautosomal region (PAR1) of the short arm of the X and Y chromosomes, is one of the major growth genes in humans. In 1997, the *SHOX* gene was linked with the occurrence of short stature in Turner syndrome for the first time (Rao et al., 1997). Subsequently, *SHOX* haploinsufficiency has been demonstrated in individuals exhibiting different phenotypes, ranging from idiopathic short stature (ISS) to Lėri-Weill dyschondrosteosis

(LWD) (Fukami et al., 2016). In LWD, the classic clinical features observed are short stature and Madelung deformity, which are also characterized by abnormal alignment of the radius, ulna, and carpal bones of the wrist (Seki et al., 2014). Additionally, a loss of both copies of *SHOX* results in the occurrence of Langer mesomelic dysplasia (LMD), which is a more severe disorder (Shears et al., 2002; Zinn et al., 2002).

The genetic defects underlying *SHOX* haploinsufficiency include copy-number variations (CNVs) and mutations, which not only occur in the coding region, but also in the regulatory elements of the *SHOX* gene (Binder, 2011). Among the regulatory elements, highly conserved non-coding DNA elements (CNEs) located several hundred kilobases downstream of *SHOX*, they have previously been identified as enhancers (Sabherwal et al., 2007). Numerous studies have shown that the *SHOX* gene enhancer plays an important role in the regulation of this gene to achieve optimal transcriptional efficiency. Furthermore, certain defects in the coding region along with deletions in the *SHOX* gene enhancer can also cause ISS or LWD, but the combined proportion in which both need to occur has not been well defined, since highly variable phenotypes are observed even if the same mutations run with in a family (Binder et al., 2004). Thus, a detailed study to determine the cause of these phenotypic differences was required.

X-chromosome inactivation (XCI) plays an important role in the phenotypic heterogeneity of X-linked disorders in females (Plenge et al., 2002; Renault et al., 2011; Sankaran et al., 2015; Torres and Puig, 2017; Juchniewicz et al., 2018). Due to skewed XCI, the same mutation in X-linked genes may result in the different phenotypes (Orstavik, 2009). While it was reported that skewed XCI might affect the phenotypes of patients with Xp22.3 complex rearrangement (Suzuki et al., 2016). However, the role of skewed XCI in the XCI-escaping gene *SHOX* was unclear. In this study, we have evaluated a case in which deletion of the *SHOX* enhancer was observed among members of a family with clinical heterogeneity due to differently skewed XCI. In this family, the proband with heterozygous deletion of the *SHOX* enhancer was a patient with LWD, although her mother harboring the same deletion was found to be almost healthy.

CASE PRESENTATION

The proband II2 is a 35-year-old Chinese woman with a short stature (150 cm, −1.1SD), bilateral Madelung deformity, bowing of the radius, and mesomelia affecting the arms, in particular, which was a typical LWD phenotype (**Figure 1A**). However, her parents and other family members were found to be healthy. It was observed that the mother was healthy with a height of 163cm (+1.1SD) and completely normal arms (**Figure 1B**).

The proband II2 arriving at our hospital was a pregnant female who was in her 14th week of gestation requesting for a non-invasive prenatal testing (NIPT). Triple tests on chromosomes 13, 18, and 21 showed a low risk in NIPT; however, a 6 Mb deletion in the long arm of the X-chromosome (Z-score was −19.43) was detected, which was anticipated to have a maternal origin (**Figure 2A**). In order to confirm the risk of this genetic disorder to the fetus, amniocentesis was subsequently conducted. Karyotyping and SNP array were performed and showed the presence of an 857 kb deletion in Xp22.33 (chrX: 784,064–1,640,746) in the male fetus III1. To verify the results, SNP array was also performed on samples of the pregnant woman II2 and her husband II1. Surprisingly, not only a deletion of 857 kb in Xp22.33, but also a 5,707 kb deletion (chrX: 127,915,006–133,621,667) in Xq25q26.3 was observed in the woman, the location of which was consistent with the result of the fetal NIPT (**Figures 2B, C** and **Supplementary Table 2**). SNP array results have been deposited in Gene Expression Omnibus (GEO), the accession number is GSE138489, as appended below: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc = GSE138489. To explore the reason underlying the phenotypic differences between the proband II2 with LWD and her healthy mother I2, we also performed a PCR-based *HUMARA* assay to assess XCI patterns.

Finally, the fetus III1 with Xp22.33 deletion was born, his length was 50cm (−0.2 SD), weight 3.3 kg (+0 SD). On the latest examination at 8 months, his developmental milestones were almost normal, and his length was 72 cm (+0.3 SD), weight 11.7 kg (+2.2 SD).

This study was carried out in accordance with the recommendations of Ethics Committee of Women's Hospital, School of Medicine Zhejiang University, and an informed consent was acquired from all the participants of this study in accordance with the Declaration of Helsinki. The study protocol was approved by the Review Board of the Women's Hospital, School of Medicine, Zhejiang University in China.

MATERIALS AND METHODS

Materials and methods were in **Supplementary Materials** and **Supplementary Table 1**.

RESULTS

In order to explore the possible causes of the disease, we analyzed the two deletions (857 kb and 5,707 kb).The 5,707 kb deletion in Xq25q26.3 included 29 Online Mendelian Inheritance in Man (OMIM) genes. Among these, 9 genes (bold) were found to be morbid and might be associated with many X-linked diseases (**Supplementary Table 2**) but not with skeletal development. Additionally, the 857 kb deletion in Xp22.33 included 11 OMIM genes in the pseudoautosomal region (PAR1) of the X and Y chromosomes (**Figures 2B–D** and **Supplementary Table 2**). Interestingly, the *SHOX* gene enhancer, which is closely associated with skeletal development, was also located in the deletion region of Xp22.33. Further analysis revealed that it was located 164 kb downstream of the *SHOX* gene (chrX: 585,079–620,146) and included the evolutionarily conserved non-coding DNA element 9 (CNE9), which was the *SHOX* gene enhancer. We speculated that the 857 kb deletion in Xp22.33 had contributed to the *SHOX* gene mutation resulting in the LWD phenotype in case of the proband.

Although no other patients were exhibiting the LWD phenotype in her family, we detected the gene dosage of *CRLF2*

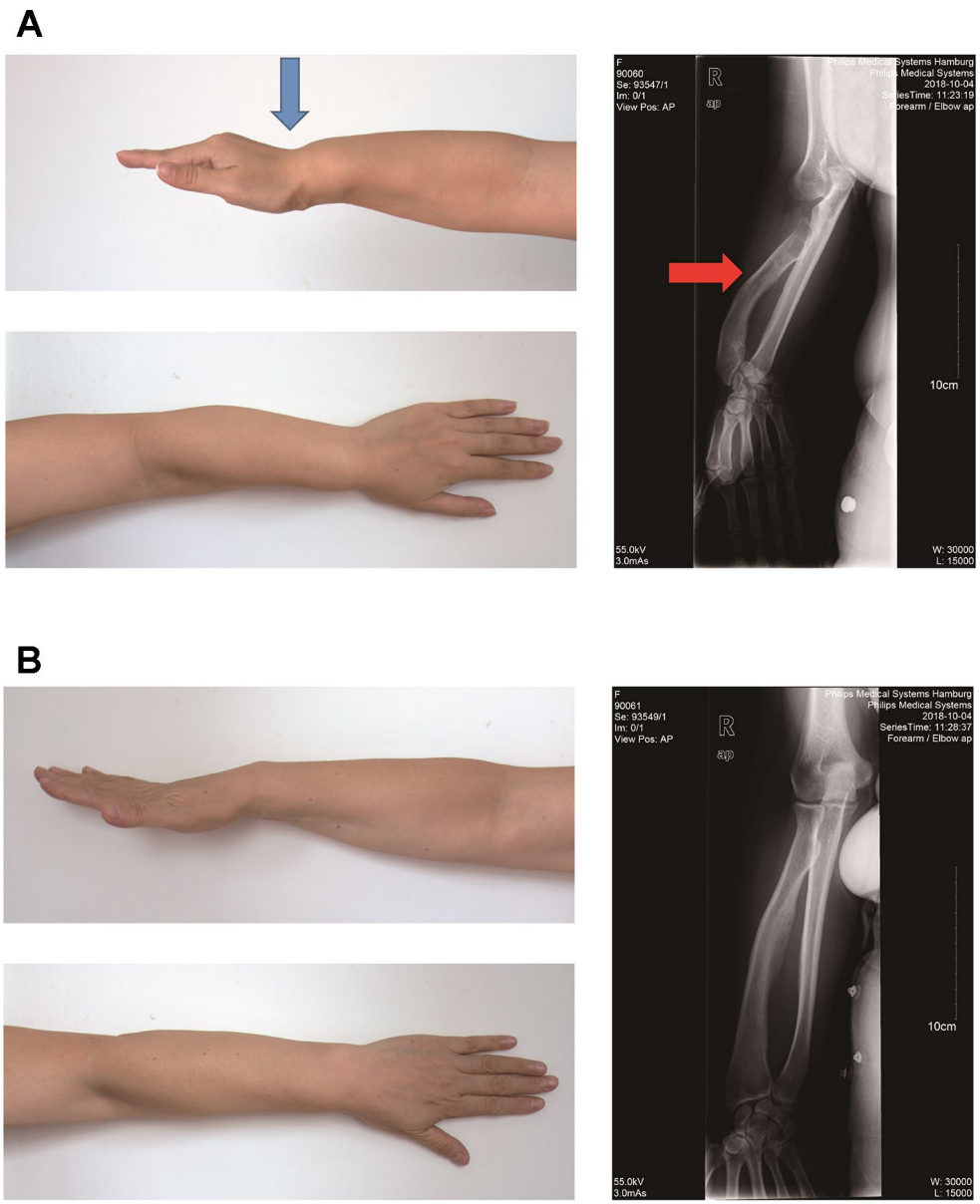

FIGURE 1 | Photograph of the forearm of the proband II2 and mother I2. **(A)**. Photograph of the forearm of the proband II2 showing a Madelung deformity and bowing of the radius (blue and red arrows). This photograph shows the curve and shortening of the forearm, hand and wrist, the structure of which appears like that of a dinner fork. **(B)**. Photograph of the forearm of the mother I2 showing a normal phenotype.

and CNE9 (the *SHOX* gene enhancer in the deletion region of Xp22.33), *AIFM1* and *FRMD7* (deletion region of Xq25q26.3) of the female proband and her parents by qPCR (**Figure 3**). The Xp22.33 regions were both located in PAR1 of the X and Y chromosomes. Data obtained from a normal female was used as the control, and the dosages of CNE9 and *CRLF2* were normal in I1 and II1, and half in I2, II2, and III1 (**Figures 3A, B**). On the other hand, the Xq25q26.3 region existed only in the X chromosome. The gene dosages of *AIFM1* and *FRMD7* were normal in female I2, and half in I1, II1, II2, and III1 (**Figures 3C**, **D**). Above all, the Xp22.33 deletion in the proband II2 was inherited from her healthy mother I2, and the Xq25q26.3 deletion occurred *de novo*.

We further explored the reason underlying the phenotypic differences between the proband II2 with LWD and her healthy mother I2. As presented in **Figure 4**, a PCR-based *HUMARA* assay was performed to assess XCI patterns in II2 and I2. After digestion with the methylation-sensitive restriction enzyme HpaII, it was found that only the inactive X-chromosome could synthesize PCR products. The origin of the inactivated X-chromosome was determined by segregation analysis. The undigested PCR product of II2 gave two peaks of 274 bp and 280

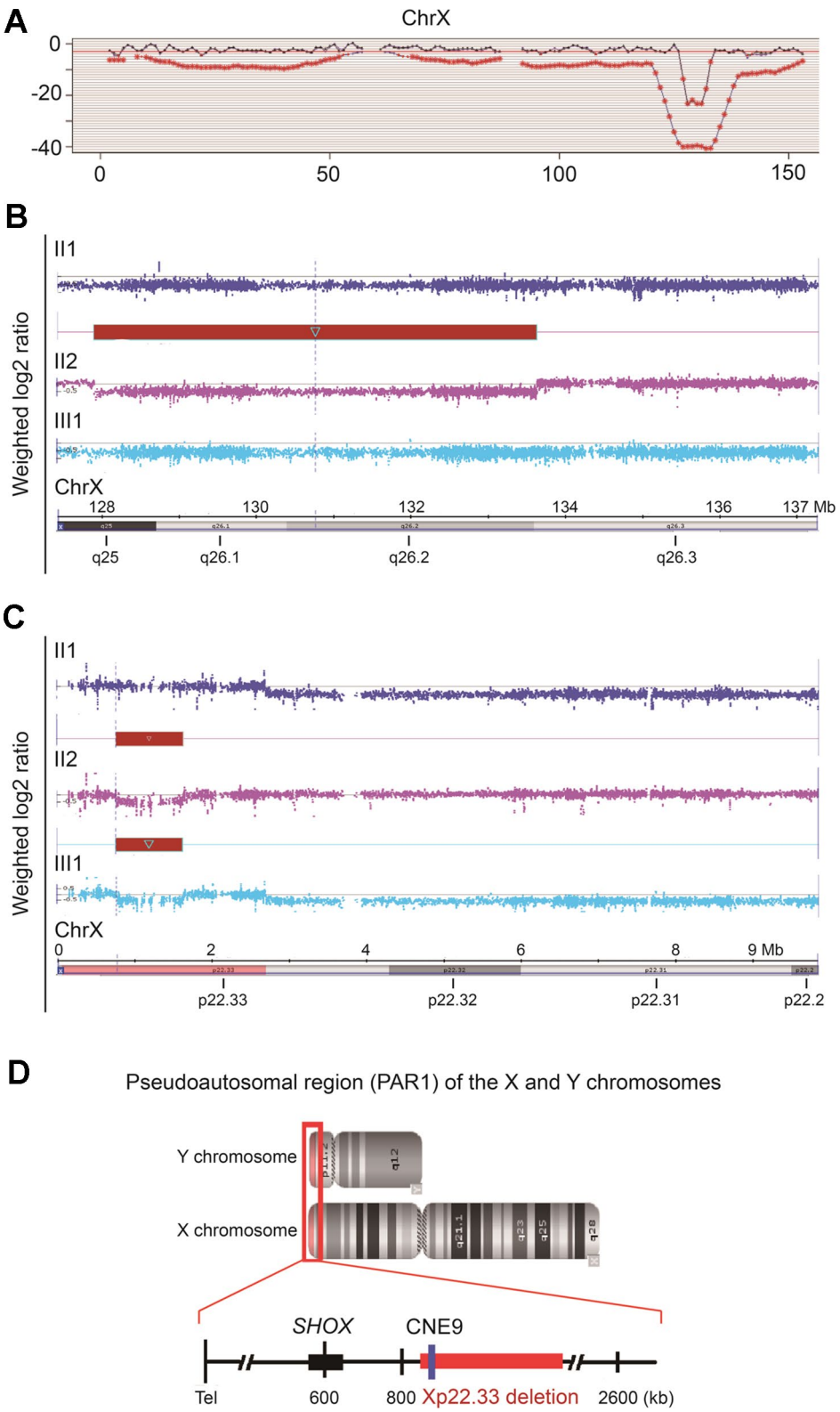

FIGURE 2 | Continued

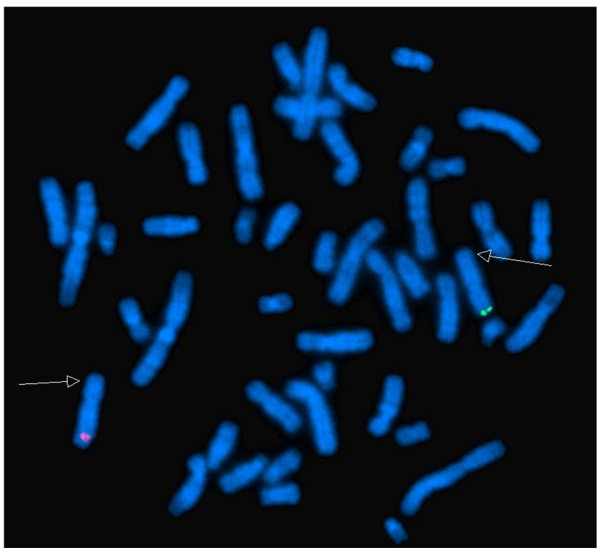

FIGURE 2 | Characteristics of the 5,707 kb and 857 kb deletions of Xq25q26.3 and Xp22.33, respectively. **(A)**. Results of cfDNA screening of the pregnant proband. cfDNA screening study of the maternal plasma, illustrating an uncertain 6 Mb deletion in the long arm of the X-chromosome (128M–133M), Z score = −19.43. **(B** and **C)**. SNP array analysis of the fetus (III1) and the couples (II1, II2). **(B)**. The red bar indicates a heterozygous 5,707 kb deletion in Xq25q26.3 (chrX: 127,915,006–133,621,667) in the fetus (III1). **(C)**. The red bar indicates a heterozygous 857 kb deletion in Xp22.33 (chrX: 784,064–1,640,746) in the fetus (III1) and the pregnant proband (II2). **(D)**. Pseudoautosomal region (PAR1) of the X and Y chromosomes. The 857 kb deletion in Xp22.33 (red bar) was located 164 kb downstream of the *SHOX* gene (chrX: 585,079–620,146), including the evolutionarily conserved CNE9 (blue bar), which was the *SHOX* gene enhancer. **(E)**. For the proband II2, FISH experiment showed that the Xp22.3 deletion (detected by the probe RP11-1119O18, Spectrum Green) and the Xq25q26 deletion (detected by the probe RP11-313D19, Spectrum Red) located on the different X chromosomes, respectively.

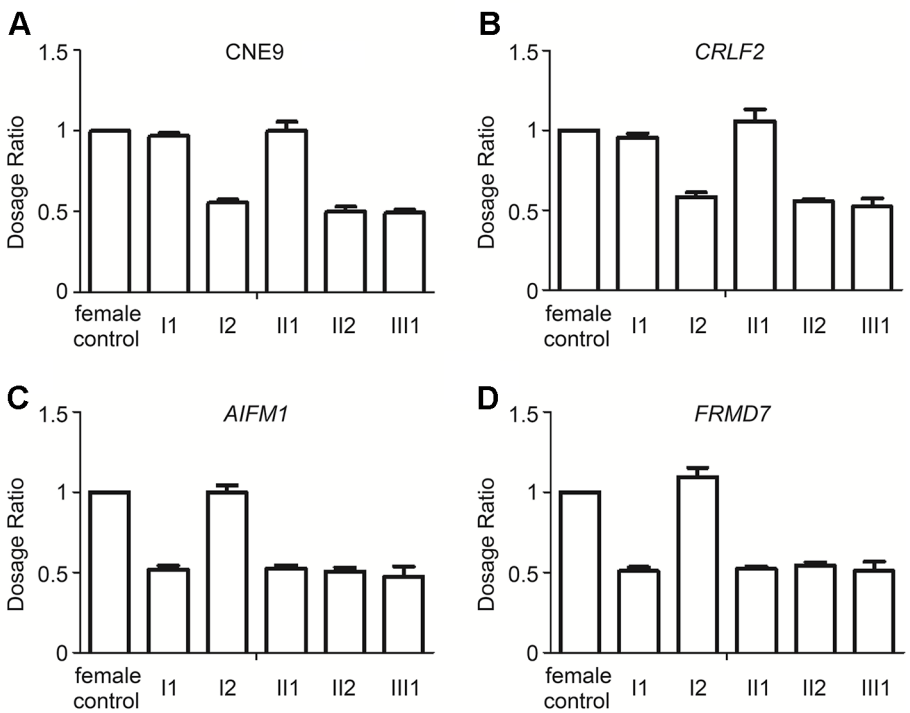

FIGURE 3 | Relative ratio of the Xp22.33 (CNE9 and *CRLF2*) and Xq25q26.3 (*AIFM1* and *FRMD7*) regions in the family by qPCR. **(A** and **B)**. Xp22.33 region (CNE9 and *CRLF2*) was located in PAR1 of the X and Y chromosomes, including the *SHOX* enhancer. Dosages in normal female were equal, Xp22.33 region dosages in I1 and II1 were normal, and those in I2, II2, III1 with Xp22.33 deletion were half. **(C** and **D)** The *AIFM1* and *FRMD7* were present only in the X-chromosome. The *AIFM1* and *FRMD7* dosage in I2 were normal, while that in I1, II1, II2, III1 with Xq25q26.3 deletion was half. From the above data, it was elucidated that the Xp22.33 deletion in the proband II2 was inherited from her healthy mother I2, and the Xq25q26.3 deletion occurred *de novo*.

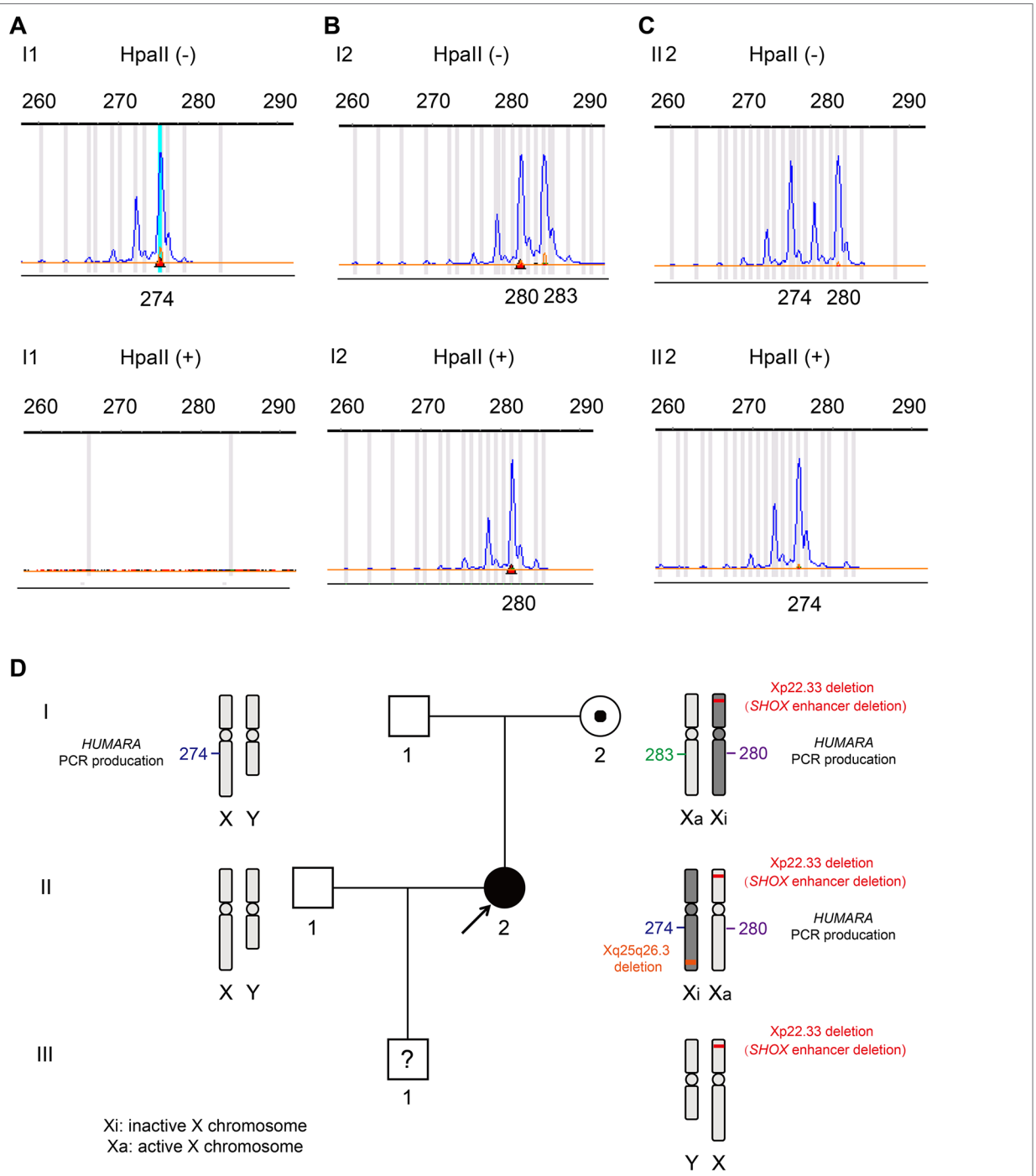

FIGURE 4 | X-chromosome inactivation (XCI) pattern and linkage analyses were based on the polymorphic CAG repeat in exon 1 of the gene for androgen receptor (*HUMARA*). **(A)**. A peak of 274 bp for *HUMARA* was observed by assaying the undigested PCR product of I1 and no peaks were observed for the HpaII digested product. **(B)**. The undigested PCR product of I2 gave two peaks of 280 and 283 bp, each, while only one peak of 280 bp was observed with the HpaII digested product. I2 exhibited 100% skewing of XCI, and the inactivated X-chromosome was linked with the 280 bp peak of the *HUMARA* PCR products. **(C)**. The undigested PCR product of the proband II2 gave two peaks of 274 and 280 bp. One X-chromosome linked with the 280 bp peak of *HUMARA* was inhibited from the mother I2 and the other from the father I1. The product of HpaII digestion gave only one peak of 274 bp. II2 also demonstrated 100% skewing of XCI, but the inactivated X-chromosome was linked with the 274 bp peak of AR, which is different from that of the mother I2. **(D).** Schematic diagram of Xp22.33 and Xq25q26.3 deletions, and *HUMARA* PCR products in the pedigree. It can be seen that Xp22.33 and Xq25q26.3 deletions in II2 are located on different X-chromosomes. Xp22.33 deletion (*SHOX* enhancer deletion) of I2 occurred in the X-chromosome, whose allele was linked with the 280 bp peak of *HUMARA* PCR. The X-chromosome was inactivated and delivered to the proband I2, but her X-chromosome was activated.

bp, respectively. A single peak representing 274 bp was observed for the HpaII-digested product (**Figure 4B**), which indicated the inactivated allele inherited from her father I1, in whose samples one peak of 274 bp was assayed by PCR. The assay of the undigested PCR product of the mother I2 gave two peaks of 280 bp and 283 bp. A single peak of 280 bp was obtained by assaying the HpaII-digested product, which was different from that of the inactivated allele of II2. The above results showed that complete (100%) skewing of XCI was found in the proband II2 and her mother I2, which was different from that of the inactivated X-chromosome (**Figures 4C, D**). Segregation analysis revealed that the normal X-chromosome from her father I1 was completely inactivated (**Figure 4A**).

Figure 4D gives aschematic diagram of Xp22.33 and Xq25q26.3 deletions, and PCR product of *HUMARA* in the pedigree. Only Xp22.33 deletion (not Xq25q26.3 deletion) derived from II2 was seen in III1. Meanwhile, the FISH detection showed that the deletions of Xp22.33 and Xq25q26.3 in II2 were located on different X-chromosomes (**Figure 2E**). Both, Xp22.33 deletion (*SHOX* enhancer deletion) and 280 bp peak of *HUMARA* PCR products in II2, were derived from I2, which were linked to the same X-chromosome. However, the allele which gave a peak of 280 bp after *HUMARA* PCR was found to be activated in II2, but inactivated in I2 (**Figures 4B, C**).

DISCUSSION

Microdeletions in the region downstream of *SHOX* have been reported as the most common genetic defects in patients with LWD (Benito-Sanz et al., 2006; Marchini et al., 2016). Several evolutionarily conserved non-coding elements located downstream of *SHOX* (for example CNE4, CNE5, ECR1, and ECS4/CNE9), are known to act as enhancers (Benito-Sanz et al., 2005; Fukami et al., 2006; Huber et al., 2006; Rappold et al., 2007), and interactions between the *SHOX* gene and CNEs have been verified using *in vitro* and *in vivo* assays (Kenyon et al., 2011; Verdin et al., 2015). Moreover, the deletion of our case also included the limb enhancer with 563 basepair (bp) (chrX: 827,128–827,691), which had specific activity in the limb regions where SHOX functions, and it also contributed to the pathogenicity of deletions downstream of *SHOX* (Skuplik et al., 2018). In addition, microdeletions in the region that is further downstream of the previously known CNEs have recently been identified in patients with LWD features or short stature (Chen et al., 2009). Therefore, we concluded that the Xp22.33 deletion resulted in *SHOX* gene defect and contributed to the manifestation of abnormal skeletal phenotype in the proband. Interestingly, the proband was a patient with LWD, but her mother harboring the Xp22.33 deletion was almost normal.

Individuals with *SHOX* defect (for example CNE9 deletion) have a phenotype ranging from normal to LWD (Benito-Sanz et al., 2012; Bunyan et al., 2013; Marchini et al., 2016). Phenotypic differences have been explained by several hypotheses. For example, as a genetic modifier, *CYP26C1* variants can effect clinical manifestations of *SHOX* deficiency (Montalbano et al., 2016). For another example, it has described skewed X-inactivation patterns, which were caused by X-chromosomal rearrangements, can also effect phenotypes of Xp22.3 defect (Suzuki et al., 2016; Ogushi et al., 2019), which gives us an inspiration.

This further indicated the presence of differently skewed XCI in the proband and her mother, which suggested that clinical heterogeneity resulting from the deletion of the *SHOX* gene enhancer was caused by the skewed XCI. While *SHOX* is also an XCI-escaping gene (Carrel and Willard, 2005; Blaschke and Rappold, 2006; Sun et al., 2017). Skewed XCI and escaping XCI are both involved in our case. Expressions of the XCI-escaping genes on two X chromosome were not always equal, XCI-escaping genes in active X chromosome (Xa) were frequently in the state of preferential expression. LAURA CARREL et al. has indicated that the XCI-escaping gene REP1 was also expressed from the inactive chromosome (Xi), but the level of expression relative to Xa was reduced, it showed that skewed XCI were also existed in XCI-escaping genes (Carrel and Willard, 1999). In addition, Nathalie Fieremans et al. has also indicated that escaping XCI was often partial and incomplete with a lower expression from the inactive X chromosome. Skewed XCI of the XCI-escaping genes *DDX3X* and *SMC1A* were revealed in intellectual disability female patients (Fieremans et al., 2016). So we speculated that the *SHOX* enhancer was similar, based on the skewed XCI and escaping XCI theory, we can explain the current case reasonably.

According to one-hit of skewed XCI, defect in a single X-chromosome results in its inactivation, in order to remedy the defect. A large proportion of female carriers of severe X-linked disorders are asymptomatic and have severely skewed XCI, probably because of selectively mediated favorable skewing, thus suggesting preferential X-inactivation against the chromosome that harbors the mutation (Orstavik, 2009), for example ATR-X syndrome (Gibbons et al., 1992), dyskeratosis congenital (Ferraris et al., 1997), X-linked agammaglobulinemia (Moschese et al., 2000), and severe combined immunodeficiency (Li et al., 1998). In our study, X-chromosomes with deletion of the *SHOX* gene enhancer were completely inactivated in the proband's mother I2, and the activated chromosomes were normal. Meanwhile, due to XCI-escaping, we speculated that transcription of SHOX was mainly from the activated X chromosome, which was normal. So it was observed that I2 was almost healthy and did not exhibit severe LWD phenotype.

According to the two-hit model of skewed XCI, if both the X-chromosomes are defective, the mutated X-chromosome with a more harmful mutation will be completely inactivated to avoid its adverse reaction, forcing the other to become active (Plenge et al., 2002; Orstavik, 2009). The hypothesis of "female X-linked two-hit model" has been used to support studies on multiple disorders, including X-linked intellectual disability (Plenge et al., 2002), *MECP2* duplication (Fieremans et al., 2014), and the Wiskott-Aldrich syndrome (Daza-Cajigal et al., 2013).

In our study, a deletion of 857 kb (including deletion of the *SHOX* gene enhancer) was found on one of the X-chromosomes of the proband II2, and another 5,707 kb deletion on the other. The larger deletion was likely more harmful. Hence, the X-chromosome with the 5,707 kb deletion was totally inactivated, and the other one with the *SHOX* gene enhancer deletion (857 kb deletion) was completely

activated. Meanwhile, due to locating on XCI-escaping region, transcription of *SHOX* was partially from the inactive chromosome, but it was mainly origin from the active X-chromosome with *SHOX* gene enhancer deletion. Combing effect of skewed XCI and escaping XCI, the proband II2 presented severe LWD phenotype.

No skewed XCI were observed in the male. Perhaps the fetus III1 was too young to present abnormal phenotypes. It was reported that, during childhood, there was probably no relevant additional loss of height in patients with *SHOX* defect (Binder and Rappold, 1993), while mesomelic disproportion of the skeleton with shortening of the extremities can be evident first until in school-aged children and increase with age in frequency and severity (Ross et al., 2001). Although the phenotype of III1 is now normal, it doesn't mean that it would be always normal, and final phenotype should be observed for long time.

XCI values depend on the tissues, but a general concordance of XCI patterns was observed among tissues from the same person (Bittel et al., 2008). Recently, Wen-Bin He et al. indicated that XCI pattern of amniocytes can predict the risk of dystrophinopathy in fetal carriers of DMD mutations (He et al., 2019). Therefore, we speculated XCI analysis in amniocytes cells may also contribute to predicting the phenotype of fetus with SHOX defect.

In summary, here we described a rare case of XCI–escaping gene *SHOX* enhancer deletion in a family with obvious clinical heterogeneity due to skewing inactivation of different X-chromosomes. Furthermore, we underlined the key role of skewed XCI and escaping XCI in the phenotype of X-linked disorders in females. It can help in the genetic counseling and prenatal diagnosis of disorders in females with *SHOX* defect.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: GSE138489 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc = GSE138489).

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Review Board of the Women's Hospital, School of Medicine, Zhejiang University in China. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin. Written informed consent was obtained from the individual(s), and minor(s)' legal guardian/next of kin, for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

MD designed the study. NIPT was carried out by HL; YS, and LW contributed the qPCR detection. YL and YQ performed the karyotyping, FISH and SNP array. XCI analysis was performed by YS and MC. YS and MD wrote the draft manuscript. YZ provided the imagological examination. All co-authors provided feedback on the estimates and contributed to the subsequent versions of the manuscript. All authors read and approved the final version of the manuscript.

FUNDING

This study was supported by the National Natural Science Foundation of China (Grant Nos.81801441 and 81300532), the Key Research and Development Program of the Zhejiang province (Grant No.2019C03025), the National Key Research and Development Program of China (Grant Nos.2016YFC1000703 and 2018YFC1002702).

ACKNOWLEDGMENTS

We also thank Dr. Jiong Gao (BGI Genomics, BGI-Shenzhen, Shenzhen 518083, China) for his assistance in the preparation of this manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2019.01086/full#supplementary-material

REFERENCES

Benito-Sanz, S., del Blanco, D. G., Aza-Carmona, M., Magano, L. F., Lapunzina, P., Argente, J., et al. (2006). PAR1 deletions downstream of SHOX are the most frequent defect in a Spanish cohort of Leri-Weill dyschondrosteosis (LWD) probands. *Hum. Mutat.* 27 (10), 1062. doi: 10.1002/humu.9456

Benito-Sanz, S., Royo, J. L., Barroso, E., Paumard-Hernandez, B., Barreda-Bonis, A. C., Liu, P., et al. (2012). Identification of the first recurrent PAR1 deletion in Leri-Weill dyschondrosteosis and idiopathic short stature reveals the presence of a novel SHOX enhancer. *J. Med. Genet.* 49 (7), 442–450. doi: 10.1136/jmedgenet-2011-100678

Benito-Sanz, S., Thomas, N. S., Huber, C., Gorbenko del Blanco, D., Aza-Carmona, M., Crolla, J. A., et al. (2005). A novel class of Pseudoautosomal region 1 deletions downstream of SHOX is associated with Leri-Weill dyschondrosteosis. *Am. J. Hum. Genet.* 77 (4), 533–544. doi: 10.1086/449313

Binder, G. (2011). Short stature due to SHOX deficiency: genotype, phenotype, and therapy. *Horm. Res. Paediatr.* 75 (2), 81–89. doi: 10.1159/000324105

Binder, G., and Rappold, G. A. (1993). "SHOX Deficiency Disorders," in *GeneReviews((R))*. Eds. Adam, M. P., Ardinger, H. H., Pagon, R. A., Wallace, S. E., Bean, L. J. H., Stephens, K., and Amemiya., A.(Seattle (WA : University of Washington)).

Binder, G., Renz, A., Martinez, A., Keselman, A., Hesse, V., Riedl, S. W., et al. (2004). SHOX haploinsufficiency and Leri-Weill dyschondrosteosis: prevalence and growth failure in relation to mutation, sex, and degree of wrist deformity. *J. Clin. Endocrinol. Metab.* 89 (9), 4403–4408. doi: 10.1210/jc.2004-0591

Bittel, D.C., Theodoro, M.F., Kibiryeva, N., Fischer, W., Talebizadeh, Z., and Butler, M.G. (2008). Comparison of X-chromosome inactivation patterns in multiple tissues from human females. *J. Med. Genet.* 45 (5), 309–313. doi: 10.1136/jmg.2007.055244.

Blaschke, R. J., and Rappold, G. (2006). The pseudoautosomal regions, SHOX and disease. *Curr. Opin. Genet. Dev.* 16 (3), 233–239. doi: 10.1016/j.gde.2006.04.004

Bunyan, D. J., Baker, K. R., Harvey, J. F., and Thomas, N. S. (2013). Diagnostic screening identifies a wide range of mutations involving the SHOX gene, including a common 47.5 kb deletion 160 kb downstream with a variable phenotypic effect. *Am. J. Med. Genet. A* 161A (6), 1329–1338. doi: 10.1002/ajmg.a.35919

Carrel, L., and Willard, H. F. (1999). Heterogeneous gene expression from the inactive X chromosome: an X-linked gene that escapes X inactivation in some human cell lines but is inactivated in others. *Proc. Natl. Acad. Sci. U.S.A.* 96 (13), 7364–7369. doi: 10.1073/pnas.96.13.7364

Carrel, L., and Willard, H. F. (2005). X-inactivation profile reveals extensive variability in X-linked gene expression in females. *Nature* 434 (7031), 400–404. doi: 10.1038/nature03479

Chen, J., Wildhardt, G., Zhong, Z., Roth, R., Weiss, B., Steinberger, D., et al. (2009). Enhancer deletions of the SHOX gene as a frequent cause of short stature: the essential role of a 250 kb downstream regulatory domain. *J. Med. Genet.* 46 (12), 834–839. doi: 10.1136/jmg.2009.067785

Daza-Cajigal, V., Martinez-Pomar, N., Garcia-Alonso, A., Heine-Suner, D., Torres, S., Vega, A. K., et al. (2013). X-linked thrombocytopenia in a female with a complex familial pattern of X-chromosome inactivation. *Blood Cells Mol. Dis.* 51 (2), 125–129. doi: 10.1016/j.bcmd.2013.04.004

Ferraris, A. M., Forni, G. L., Mangerini, R., and Gaetani, G. F. (1997). Nonrandom X-chromosome inactivation in hemopoietic cells from carriers of dyskeratosis congenita. *Am. J. Hum. Genet.* 61 (2), 458–461. doi: 10.1016/S0002-9297(07)64075-0

Fieremans, N., Bauters, M., Belet, S., Verbeeck, J., Jansen, A. C., Seneca, S., et al. (2014). De novo MECP2 duplications in two females with intellectual disability and unfavorable complete skewed X-inactivation. *Hum. Genet.* 133 (11), 1359–1367. doi: 10.1007/s00439-014-1469-6

Fieremans, N., Van Esch, H., Holvoet, M., Van Goethem, G., Devriendt, K., Rosello, M., et al. (2016). Identification of Intellectual Disability Genes in Female Patients with a Skewed X-Inactivation Pattern. *Hum. Mutat.* 37 (8), 804–811. doi: 10.1002/humu.23012

Fukami, M., Kato, F., Tajima, T., Yokoya, S., and Ogata, T. (2006). Transactivation function of an approximately 800-bp evolutionarily conserved sequence at the SHOX 3' region: implication for the downstream enhancer. *Am. J. Hum. Genet.* 78 (1), 167–170. doi: 10.1086/499254

Fukami, M., Seki, A., and Ogata, T. (2016). SHOX Haploinsufficiency as a Cause of Syndromic and Nonsyndromic Short Stature. *Mol. Syndromol.* 7 (1), 3–11. doi: 10.1159/000444596

Gibbons, R. J., Suthers, G. K., Wilkie, A. O., Buckle, V. J., and Higgs, D. R. (1992). X-linked alpha-thalassemia/mental retardation (ATR-X) syndrome: localization to Xq12-q21.31 by X inactivation and linkage analysis. *Am. J. Hum. Genet.* 51 (5), 1136–1149.

He, W. B., Du, J., Xie, P. Y., Zhou, S., Zhang, Y. X., Lu, G. X., et al. (2019). X-chromosome inactivation pattern of amniocytes predicts the risk of dystrophinopathy in fetal carriers of DMD mutations. *Prenat. Diagn.* 39 (8), 603–608. doi: 10.1002/pd.5473

Huber, C., Rosilio, M., Munnich, A., Cormier-Daire, V., and French, S. G. M. (2006). High incidence of SHOX anomalies in individuals with short stature. *J. Med. Genet.* 43 (9), 735–739. doi: 10.1136/jmg.2006.040998

Juchniewicz, P., Kloska, A., Tylki-Szymanska, A., Jakobkiewicz-Banecka, J., Wegrzyn, G., Moskot, M., et al. (2018). Female fabry disease patients and X-chromosome inactivation. *Gene* 641, 259–264. doi: 10.1016/j.gene.2017.10.064

Kenyon, E. J., McEwen, G. K., Callaway, H., and Elgar, G. (2011). Functional analysis of conserved non-coding regions around the short stature hox gene (shox) in whole zebrafish embryos. *PloS One* 6 (6), e21498. doi: 10.1371/journal.pone.0021498

Li, S. L., Ting, S. S., Lindeman, R., Ffrench, R., and Ziegler, J. B. (1998). Carrier identification in X-linked immunodeficiency diseases. *J. Paediatr. Child Health* 34 (3), 273–279. doi: 10.1046/j.1440-1754.1998.00216.x

Marchini, A., Ogata, T., and Rappold, G. A. (2016). A track record on SHOX: from basic research to complex models and therapy. *Endocr. Rev.* 37 (4), 417–448. doi: 10.1210/er.2016-1036

Montalbano, A., Juergensen, L., Roeth, R., Weiss, B., Fukami, M., Fricke-Otto, S., et al. (2016). Retinoic acid catabolizing enzyme CYP26C1 is a genetic modifier in SHOX deficiency. *EMBO Mol. Med.* 8 (12), 1455–1469. doi: 10.15252/emmm.201606623

Moschese, V., Orlandi, P., Plebani, A., Arvanitidis, K., Fiorini, M., Speletas, M., et al. (2000). X-chromosome inactivation and mutation pattern in the Bruton's tyrosine kinase gene in patients with X-linked agammaglobulinemia. *Italian XLA Collaborative Group. Mol. Med.* 6 (2), 104–113.

Ogushi, K., Hattori, A., Suzuki, E., Shima, H., Izawa, M., Yagasaki, H., et al. (2019). DNA Methylation Status of SHOX-Flanking CpG Islands in Healthy Individuals and Short Stature Patients with Pseudoautosomal Copy Number Variations. *Cytogenet. Genome Res.* 158, 56–62. doi: 10.1159/000500468

Orstavik, K. H. (2009). X chromosome inactivation in clinical practice. *Hum. Genet.* 126 (3), 363–373. doi: 10.1007/s00439-009-0670-5

Plenge, R. M., Stevenson, R. A., Lubs, H. A., Schwartz, C. E., and Willard, H. F. (2002). Skewed X-chromosome inactivation is a common feature of X-linked mental retardation disorders. *Am. J. Hum. Genet.* 71 (1), 168–173. doi: 10.1086/341123

Rao, E., Weiss, B., Fukami, M., Rump, A., Niesler, B., Mertz, A., et al. (1997). Pseudoautosomal deletions encompassing a novel homeobox gene cause growth failure in idiopathic short stature and Turner syndrome. *Nat. Genet.* 16 (1), 54–63. doi: 10.1038/ng0597-54

Rappold, G., Blum, W. F., Shavrikova, E. P., Crowe, B. J., Roeth, R., Quigley, C. A., et al. (2007). Genotypes and phenotypes in children with short stature: clinical indicators of SHOX haploinsufficiency. *J. Med. Genet.* 44 (5), 306–313. doi: 10.1136/jmg.2006.046581

Renault, N. K., Renault, M. P., Copeland, E., Howell, R. E., and Greer, W. L. (2011). Familial skewed X-chromosome inactivation linked to a component of the cohesin complex, SA2. *J. Hum. Genet.* 56 (5), 390–397. doi: 10.1038/jhg.2011.25

Ross, J. L., Scott, C. Jr., Marttila, P., Kowal, K., Nass, A., Papenhausen, P., et al. (2001). Phenotypes Associated with SHOX Deficiency. *J. Clin. Endocrinol. Metab.* 86 (12), 5674–5680. doi: 10.1210/jcem.86.12.8125

Sabherwal, N., Bangs, F., Roth, R., Weiss, B., Jantz, K., Tiecke, E., et al. (2007). Long-range conserved non-coding SHOX sequences regulate expression in developing chicken limb and are associated with short stature phenotypes in human patients. *Hum. Mol. Genet.* 16 (2), 210–222. doi: 10.1093/hmg/ddl470

Sankaran, V. G., Ulirsch, J. C., Tchaikovskii, V., Ludwig, L. S., Wakabayashi, A., Kadirvel, S., et al. (2015). X-linked macrocytic dyserythropoietic anemia in females with an ALAS2 mutation. *J. Clin. Invest.* 125 (4), 1665–1669. doi: 10.1172/JCI78619

Seki, A., Jinno, T., Suzuki, E., Takayama, S., Ogata, T., and Fukami, M. (2014). Skeletal deformity associated with SHOX deficiency. *Clin. Pediatr. Endocrinol.* 23 (3), 65–72. doi: 10.1297/cpe.23.65

Shears, D. J., Guillen-Navarro, E., Sempere-Miralles, M., Domingo-Jimenez, R., Scambler, P. J., and Winter, R. M. (2002). Pseudodominant inheritance of Langer mesomelic dysplasia caused by a SHOX homeobox missense mutation. *Am. J. Med. Genet.* 110 (2), 153–157. doi: 10.1002/ajmg.10421

Skuplik, I., Benito-Sanz, S., Rosin, J. M., Bobick, B. E., Heath, K. E., and Cobb, J. (2018). Identification of a limb enhancer that is removed by pathogenic deletions downstream of the SHOX gene. *Sci. Rep.* 8 (1), 14292. doi: 10.1038/s41598-018-32565-1

Sun, Y. X., Zhang, Y. X., Zhang, D., Xu, C. M., Chen, S. C., Zhang, J. Y., et al. (2017). XCI-escaping gene KDM5C contributes to ovarian development *via* downregulating miR-320a. *Hum. Genet.* 136 (2), 227–239. doi: 10.1007/s00439-016-1752-9

Suzuki, E., Shima, H., Toki, M., Hanew, K., Matsubara, K., Kurahashi, H., et al. (2016). Complex X-chromosomal rearrangements in two women with ovarian dysfunction: implications of chromothripsis/chromoanasynthesis-dependent and -independent origins of complex genomic alterations. *Cytogenet. Genome Res.* 150 (2), 86–92. doi: 10.1159/000455026

Torres, R. J., and Puig, J. G. (2017). Skewed X inactivation in Lesch-Nyhan disease carrier females. *J. Hum. Genet.* 62 (12), 1079–1083. doi: 10.1038/jhg.2017.88

Verdin, H., Fernandez-Minan, A., Benito-Sanz, S., Janssens, S., Callewaert, B., De Waele, K., et al. (2015). Profiling of conserved non-coding elements upstream of SHOX and functional characterisation of the SHOX cis-regulatory landscape. *Sci. Rep.* 5, 17667. doi: 10.1038/srep17667

Zinn, A. R., Wei, F., Zhang, L., Elder, F. F., Scott, C. I. Jr., Marttila, P., et al. (2002). Complete SHOX deficiency causes Langer mesomelic dysplasia. *Am. J. Med. Genet.* 110 (2), 158–163. doi: 10.1002/ajmg.10422

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Genetics

ORIGINAL RESEARCH
published: 14 November 2019
doi: 10.3389/fgene.2019.01016

In-Frame Variants in *STAG3* Gene Cause Premature Ovarian Insufficiency

Wen-Juan Xiao[1†], *Wen-Bin He*[1,2,3†], *Ya-Xin Zhang*[1], *Lan-Lan Meng*[2,3], *Guang-Xiu Lu*[1,2,3], *Ge Lin*[1,2,3], *Yue-Qiu Tan*[1,2,3*‡] *and Juan Du*[1,2,3*‡]

[1] *Institute of Reproduction and Stem Cell Engineering, Central South University, Changsha, China,* [2] *Genetic Center, Reproductive and Genetic Hospital of CITIC-Xiangya, Changsha, China,* [3] *NHC Key Laboratory of Human Stem Cell and Reproductive Engineering, Central South University, Changsha, China*

OPEN ACCESS

Edited by:
Fan Jin,
Zhejiang University, China

Reviewed by:
Ye Wang,
Sun Yat-sen University, China
Xu Chenming,
Shanghai Jiao Tong University, China

***Correspondence:**
Yue-Qiu Tan
tanyueqiu@csu.edu.cn
Juan Du
tandujuan@csu.edu.cn

†These authors have contributed equally to this work and share first authorship

‡These authors have contributed equally to this work and share senior authorship

Specialty section:
This article was submitted to Genetic Disorders, a section of the journal Frontiers in Genetics

Received: *17 April 2019*
Accepted: *24 September 2019*
Published: *14 November 2019*

Citation:
Xiao W-J, He W-B, Zhang Y-X, Meng L-L, Lu G-X, Lin G, Tan Y-Q and Du J (2019) In-Frame Variants in STAG3 Gene Cause Premature Ovarian Insufficiency. Front. Genet. 10:1016. doi: 10.3389/fgene.2019.01016

Premature ovarian insufficiency (POI) is a severe clinical syndrome defined by ovarian dysfunction in women less than 40 years old who generally manifest with infertility, menstrual disturbance, elevated gonadotrophins, and low estradiol levels. *STAG3* is considered a genetic aetiology of POI, which facilitates entry of REC8 into the nucleus of a cell and plays an essential role in gametogenesis. At present, only six truncated variants associated with POI have been reported; there have been no reports of an in-frame variant of *STAG3* causing POI. In this study, two novel homozygous in-frame variants (c.877_885del, p.293_295del; c.891_893dupTGA, p.297_298insAsp) in *STAG3* were identified in two sisters with POI from a five-generation consanguineous Han Chinese family. To evaluate the effects of these two variants, we performed fluorescence localization and co-immunoprecipitation analyses using *in vitro* cell model. The two variants were shown to be pathogenic, as neither STAG3 nor REC8 entered nuclei, and interactions between mutant STAG3 and REC8 or SMC1A were absent. To the best of our knowledge, this is the first report on in-frame variants of *STAG3* that cause POI. This finding extends the spectrum of variants in *STAG3* and sheds new light on the genetic origins of POI.

Keywords: premature ovarian insufficiency, *STAG3* gene, in-frame variant, fluorescence localization, co-immunoprecipitation

INTRODUCTION

Premature ovarian insufficiency (POI), also known as premature ovarian failure, refers to hypergonadotropic hypogonadism in women younger than 40 years and is one of the major causes of female infertility, affecting at least 1–3% of adult women in the world (Bannatyne et al., 1990). Apart from menstrual disturbance (amenorrhea or oligomenorrhea for at least 4 months), the main symptoms of POI are reduced levels of estradiol and elevated plasma levels of follicle stimulating hormone (FSH) (>25 mIU/ml on two occasions, > 4 weeks apart) (Shelling, 2010; Gowri et al., 2015; Tucker et al., 2016; Webber et al., 2016). The etiology of POI is complex, such as that for autoimmune diseases, chemotherapy, and pelvic surgery, among which genetic defect was reported to be related to POI (Franca et al., 2019).

POI is a condition with extremely high genetic heterogeneity. A set of genes has been reported to be responsible for POI, including X-linked genes (e.g., *BMP15* and *FMR1*) and autosomal genes

(e.g., *HFM1*, *FIGLA*, *FOXL2*, *STAG3*, and *FSHR*) (Rossetti et al., 2017). Stromal antigen 3 (*STAG3*) is a meiosis-specific gene that is restrictedly expressed in testes and ovaries in humans, and it plays an important role in gametogenesis and fertility (Houmard et al., 2009; Nogues et al., 2009; Garciacruz et al., 2010; Caburet et al., 2014). Variants of *STAG3* are rare, and only six truncated variants, three frameshift variants, two nonsense variants, and one splicing variant (Caburet et al., 2014; Stabej et al., 2016; Colombo et al., 2017; He et al., 2018a; Franca et al., 2019) (**Table 1**) associated with POI have been reported to date. However, it remains enigmatic whether in-frame variants of *STAG3* can cause POI and female infertility.

In this study, we identified two novel homozygous in-frame variants of *STAG3* in a consanguineous Han Chinese family with POI by whole-exome sequencing (WES). Moreover, we performed *in silico* and *in vitro* functional analyses to reveal the relationship between the two variants and POI. To the best of our knowledge, this is the first report of in-frame variants of *STAG3* associated with POI, and the first example of performing *in vitro* functional analyses to functionally characterize in-frame variants of *STAG3*.

MATERIALS AND METHODS

Patients and Blood Sampling

Four members of a five-generation consanguineous Han Chinese family with POI, namely, two affected sisters (V-1; V-2, **Figure 1A**) and two unaffected second-cousin parents (IV-1; IV-2, **Figure 1A**) participated in this study. Peripheral blood samples from the four individuals were collected, and genomic DNA (gDNA) was extracted using a QIAamp DNA Blood Midi Kit (Qiagen, Hilden, Germany). The four family members provided their written informed consent to participate in this study and agreed with the publication of this case report. This study was reviewed and approved by the ethics committee of the Reproductive & Genetic Hospital of CITIC-Xiangya of Central South University, China.

WES, Variant Filtering, and Sanger Sequencing

The gDNA from the proband was subjected to WES following protocols described previously (He et al., 2018b). Whole exomes were captured using Agilent SureSelect version 4 (Agilent Technologies, Santa Clara, CA) and sequenced on a HighSeq2000 sequencing platform (Illumina, San Diego, California, USA). WES data analysis was performed using the Genome Analysis Toolkit (GATK) and consisted of removing adaptors, mapping WES raw reads to a reference human genome sequence (NCBI GRCh37, reference genome Hg19) using the Burrows–Wheeler Aligner (BWA), eliminating PCR duplicates, and sorting sequences using Picard (http://broadinstitute.github.io/picard/). Variant identification was performed using the GATK package according to the recommended best practices, including base recalibration variant calling with Haplotype Caller and variant quality score recalibration and annotation using ANNOVAR software.

All candidate variants needed to meet the following criteria: (i) absent or occurring at a frequency less than 0.01 in any of the following databases: 1,000 Genomes Variant Database, Exome Aggregation Consortium (ExAC), or NHLBI-GO Exome Sequencing Project (GO-ESP); (ii) homozygous variants were considered with priority; (iii) predicted to be deleterious; and (iv) relevant to the phenotype

TABLE 1 | The currently reported phenotypes and genotypes of *STAG3* gene in POI patients with 46,XX.

Patients	Genotype	Phenotype						References
		Menstrual history	FSH (mIU/ml)	LH (mIU/ml)	Estradiol (Pg/ml)	Ultrasonographic examination	Age at diagnosis (years)	
1	c.877_885del (p.293_295del) (Hom) c.891_893dupTGA (p.297_298insAsp) (Hom)	Primary amenorrhea	51	11.08	<10	Uterus and ovaries not visualized	18	This study
2	c.877_885del (p.293_295del) (Hom)c.891_893dupTGA (p.297_298insAsp) (Hom)	Primary amenorrhea	72.19	18.58	<10	Small uterus and small ovaries	21	This study
3	c.291dupC (p.Asn98Glnfs*2) (Het)c.1950C > A (p.Tyr650*) (Het)	Primary amenorrhea	89	37	<13	Infantile uterus; ovaries were not visualized	21	(Franca et al., 2019)
4	c.677C > G (p.Ser227*) (Hom)	Primary amenorrhea	48	23	19	NA	NA	(Colombo et al., 2017)
5	c.1573+5G > A (p.Leu490Thrfs*10) (Hom)	Primary amenorrhea	48.69	26	normal	Streak gonads and small uterus	19	(He et al., 2018a)
6	c.1573+5G > A (p.Leu490Thrfs*10) (Hom)	Primary amenorrhea	48.38	25.51	normal	NA	NA	
7	c.1947_48dupCT (p.Tyr650Serfs*22) (Hom)	Primary amenorrhea	136	31	<37	Normal uterus; bilateral streak gonads	NA	(Stabej et al., 2016)
8	c.1947_48dupCT (p.Tyr650Serfs*22) (Hom)	Primary amenorrhea	130	62	<37	Small uterus; streak gonads	NA	
9-12	c.968delC (p.Gln188Argfs*8) (Hom)	Primary amenorrhea	>45	>18	<22	Bilateral streak gonads	17–20 (4 sisters)	(Caburet et al., 2014)

POI, premature ovarian insufficiency; FSH, follicle stimulating hormone; LH, luteinizing hormone; Hom, homozygote variant; Het, heterozygote variant; NA, data not available.

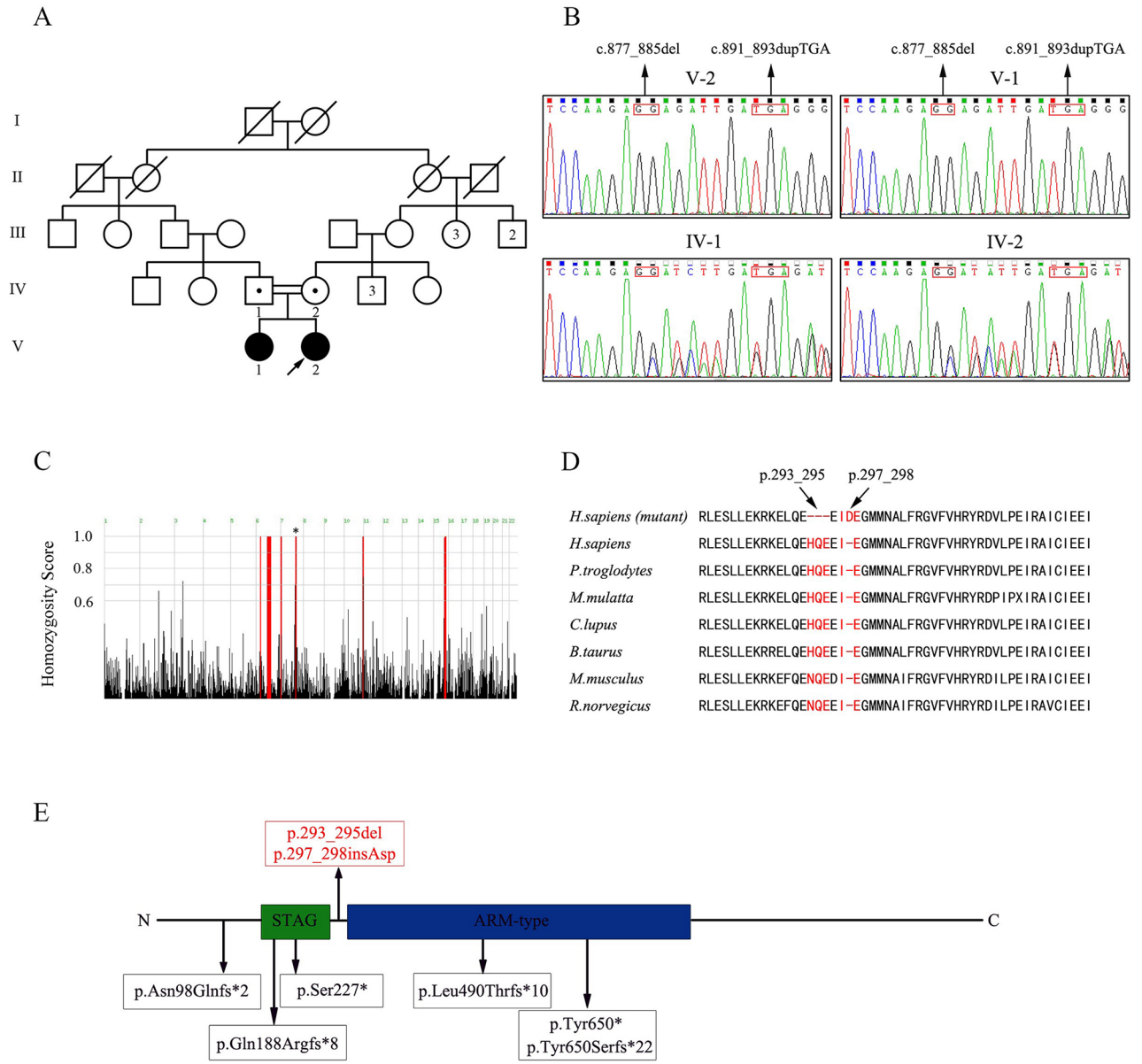

FIGURE 1 | Pedigree of the consanguineous Han Chinese family with POI-associated *STAG3* variants. **(A)** Two second cousins (IV-1; IV-2) in generation four married each other and had two daughters (V-1; V-2) with POI. The number of siblings is written inside each symbol. The proband is marked with a black arrow. Filled symbols indicate other affected members. Open symbols indicate unaffected members. Heterozygous carriers are indicated with a dot in the middle of each symbol. Numbers are allotted to the family members whose DNA samples were used in this study. **(B)** Based on sequence chromatograms for *STAG3* (NM_001282716: c.877_885del; c.891_893dupTGA) in this family, the unaffected parents (IV-1; IV-2) are heterozygous carriers of two variants of *STAG3*, whereas the two affected sisters (V-1; V-2) are homozygous. **(C)** Homozygosity mapping of the proband. Homozygous regions with a remarkable signal are indicated in red. The asterisk indicates the location of *STAG3*. **(D)** Cross-species alignment reveals strong evolutionary conservation at positions 293_295 and 297_298 in *STAG3* (NP_036579.2). **(E)** The STAG3 protein consists of 1,225 amino acids, including a STAG domain and an ARM-type domain. The two variants discovered in this study are shown at the top of the figure, whereas previously reported variants are displayed at the bottom of the figure.

shared by the patients based on comprehensive expression data (expression in ovary), and model organism data (a female-lacking-oocytes phenotype presented in animal models).

Specific PCR primers targeting two variants of the *STAG3* gene (NM_001282716) were designed and validated. The sequences were as follows: 5′-AACCACATGCAGAGGGGTTAT-3′ and 5′-TCCAGCTGCATTAATTCTGGGA-3.′

Plasmid Construction

The full-length coding sequence of *STAG3* was amplified from control human blood cDNA, and the *STAG3* product was cloned into pDsRed2-N1, resulting in STAG3 followed by DsRed2 protein, as STAG3-WT-DsRed2. Additionally, the cDNA encoding $FLAG_3$-STAG3 was cloned into pcDNA3.1, namely, $FLAG_3$-STAG3-WT. Variants c.877_885del and

c.891_893dupTGA were separately or simultaneously introduced into the STAG3-WT-DsRed2 and $FLAG_3$-STAG3-WT plasmid vectors using a Mut Express II Fast Mutagenesis Kit V2 (Vazyme, Guangzhou, China) to achieve site-directed mutagenesis. Altogether, six mutants were generated (Mut 1 indicates STAG3-p.293_295del-DsRed2 or $FLAG_3$-STAG3-p.293_295del; Mut 2 indicates STAG3-p.297_298insAsp-DsRed2 or $FLAG_3$-STAG3-p.297_298insAsp; Mut 3 indicates STAG3-p.293_295del and p.297_298insAsp-DsRed2 or $FLAG_3$-STAG3-p.293_295del and p.297_298insAsp). Moreover, cDNAs encoding REC8 and SMC1A were separately cloned into pEGFP-N1 and pcDNA3.1-HA. All expression constructs were sequenced to exclude unintended PCR-generated variants and confirm the presence of the desired variants.

Cell Culture

Chinese hamster ovary (CHO) cells and human embryonic kidney (HEK) 293 cells were cultured in Dulbecco's modified Eagle medium nutrient mixture F-12 (Ham) (DMEM/F12) (Gibco, Gaithersburg, MD, USA) with 10% fetal bovine serum (FBS) (Gibco, Gaithersburg, MD, USA) at 37°C in a humidified 5% CO_2 incubator.

Fluorescence Localization

According to the manufacturer's instructions for Lipofectamine 3000 (Thermo Fisher Scientific, Carlsbad, CA, USA), CHO cells were transfected with expression vectors containing wild-type *STAG3-DsRed2* and *REC8-EGFP* plasmids or mutated *STAG3-DsRed2* and *REC8-EGFP* plasmids. CHO cells transfected with empty vector (EV; without *STAG3*) and *REC8-EGFP* plasmids were used as negative controls. CHO cells were fixed for fluorescence localization after culture for 48 h. 4',6-Diamidino-2-phenylindole (DAPI) (Beyotime) was used to visualize DNA. All fluorescent images were captured on an orthostatic fluorescence microscope (Olympus, Japan).

Co-Immunoprecipitation (Co-IP)

Co-IP was performed as described previously (Wolf et al., 2018). Briefly, HEK293 cells were transfected with expression vectors containing wild-type *$FLAG_3$-STAG3* and *REC8-EGFP* plasmids (or SMC1A-HA plasmids) or mutated *$FLAG_3$-STAG3* and *REC8-EGFP* plasmids (or SMC1A-HA plasmids), according to the manufacturer's instructions for Lipofectamine 3000 use. Untransfected HEK293 cells were used as a negative control. After incubation for 36 h, the culture medium was supplemented with 0.2 μg/ml nocodazole (Sigma, Louis, MO, USA) for prometaphase arrest of the cells. Total protein was extracted using RIPA lysis buffer (Thermo Fisher Scientific) after 12 h, and then, western blotting was performed to detect the expression of STAG3, REC8, and SMC1A using anti-FLAG, anti-GFP, and anti-HA (Proteintech) antibodies, respectively, at a 1:1,000 dilution. Next, immunoprecipitation was performed using Dynabeads Protein G (Invitrogen) following the manufacturer's instructions.

RESULTS

Clinical Features of the Affected Individuals

The proband (V-2, **Figure 1A**) initially went to the clinic because of primary amenorrhea and was diagnosed with POI at 18 years old. Her breast development was at Tanner stage I, although she had a normal height (160 cm) and weight (50 kg). The results of hormonal studies showed that her serum FSH level was high (51 mIU/ml, normal range 2–22 mIU/ml), whereas her serum estradiol level was low (< 10 pg/ml, normal range 30–190 pg/ml), and her anti-Müllerian hormone (AMH) level was too low to be detected. She had a normal level of luteinizing hormone (LH; 11.08 mIU/ml, normal range 1–19 mIU/ml). Her uterus and ovaries were not visualized upon ultrasonographic examination, and she had an abnormal vulva without labia minora. The proband did not accept hormonal replacement therapy (HRT) because of the increased prevalence of invasive breast cancer after HRT (Biscup, 2003).

The proband's affected elder sister (V-1, **Figure 1A**) showed normal growth and development but presented with POI when she was 21 years old. Similar to the proband, she went to the hospital because of primary amenorrhea. Her serum FSH level was elevated (72.19 mIU/ml, normal range 2–22 mIU/ml), whereas her serum estradiol level was low (< 10 pg/ml, normal range 30–190 pg/ml), and her LH level was normal (18.58 mIU/ml, normal range 1–19 mIU/ml). Ultrasonographic examination showed that her uterus was anteflexed and small (23×10×20 mm). Both ovaries were also small: the left ovary was 14×10×11 mm, and the right one was 15×11×13 mm, consistent with the size of streak gonads. After receiving HRT with progesterone for 3 months, she showed normal menses, and her uterus grew (36×20×35 mm) to a size close to that of a normal adult-sized uterus.

Additionally, both sisters had a normal karyotype (46, XX) and FMR1 CGG repeats, and no associated adrenal or thyroid autoimmune disorders were found.

Identification of Two Homozygous In-Frame Variants of *STAG3*

WES of the proband (V-2, **Figure 1A**) resulted in approximately 139,631,574 raw reads with a mean 182.73-fold depth in the target regions, revealing high-quality sequencing (**Supplementary Table 1**). For rare inherited diseases, the frequency of possible pathogenic variants should be very low in a healthy population. Therefore, the results were screened against minor allele frequency (MAF) > 1% in public SNP and indel databases for variants predicted to be deleterious or to result in loss of function. Homozygous variants were preferentially considered candidates because the patients were from a consanguineous family. Finally, we focused on the phenotypic relevance for POI in the proband. Only two homozygous variants of *STAG3* (NM_001282716: c.877_885del, p.293_295del; c.891_893dupTGA, p.297_298insAsp) fulfilled these criteria (**Supplementary Table 2**).

The two homozygous in-frame *STAG3* variants were confirmed by Sanger sequencing and were also detected in her affected sister

(**Figure 1B**). Their unaffected parents were heterozygous carriers of the two variants (**Figure 1B**). In addition, homozygosity mapping indicated that the two *STAG3* variants were not located in heterozygous regions (**Figure 1C**) and that positions p.293_295 and p.297_298 are highly conserved across different species (**Figure 1D**).

According to the American College of Medical Genetics and Genomics (ACMG) standards and guidelines for the interpretation of variations (Richards et al., 2015), the two in-frame variants of STAG3 (c.877_885del, p.293_295del; c.891_893dupTGA, p.297_298insAsp) are both classified as likely pathogenic variants. Therefore, we speculated that both homozygous *STAG3* variants were candidates as causes of POI.

Effects of Mutant STAG3 Proteins on REC8 Localization in CHO Cells

Co-expression of wild-type STAG3 and REC8 is necessary for REC8 to enter nuclei. A fluorescence localization analysis was performed to evaluate the effects of the two variants. As expected, REC8 localized to nuclei, as did the will-type STAG3 protein, when CHO cells were transfected with expression vectors containing wild-type *STAG3* and *REC8* plasmids; (**Figure 2A**). In contrast, REC8 localized exclusively in the cytoplasm (**Figure 2A**) when CHO cells were co-transfected with the *REC8* construct and empty vector or mutant *STAG3* construct. In the latter scenario, nearly all of the mutant STAG3 proteins were excluded from nuclei (**Figure 2A**).

Lack of Interactions Between Mutant STAG3 and REC8 or SMC1A

We then performed a Co-IP analysis to determine the pathogenicity of the two variants. As shown, wild-type STAG3 interacted with both REC8 and SMC1A as previously reported (Prieto et al., 2001; Wolf et al., 2018) (**Figure 2B**). Additionally, when we directly examined the ability of the three transiently expressed mutant STAG3 proteins (p.293_295del; p.297_298insAsp; p.293_295del, and p.297_298insAsp) to associate with REC8 or SMC1A, the results revealed that there was no interaction between the three mutant STAG3 proteins and REC8 or SMC1A (**Figure 2B**).

DISCUSSION

In the present study, two novel homozygous in-frame variants (c.877_885del, p.293_295del; c.891_893dupTGA,

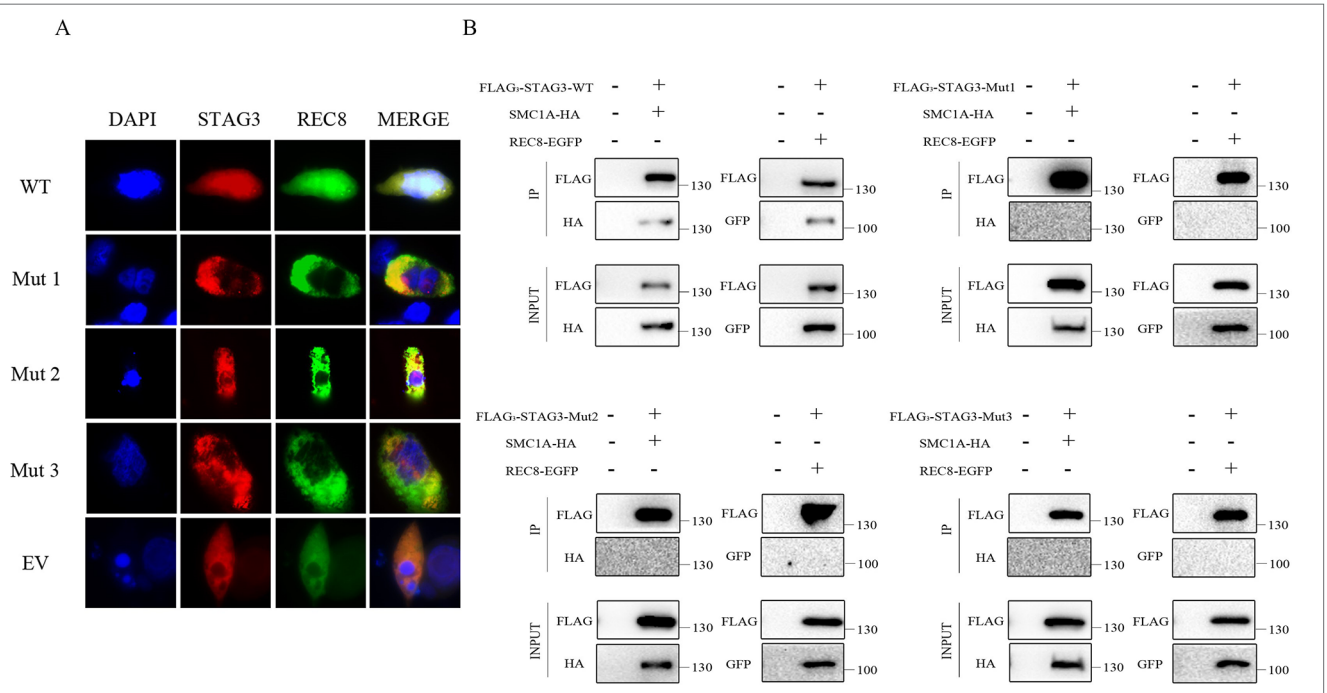

FIGURE 2 | Fluorescence localization and Co-IP analyses of the two in-frame variants of *STAG3*. **(A)** CHO cells were transiently transfected with a plasmid encoding STAG3-DsRed2 and REC8-EGFP, mutant STAG3-DsRed2 and REC8-EGFP, or an empty vector and REC8-EGFP. After culture for 48 h, cells were fixed for fluorescence localization. Blue fluorescence indicates DAPI staining in nuclei. The three mutant STAG3 proteins were restricted to the cytoplasm and could not assist REC8 in entering nuclei, resulting in aberrant localization of REC8. **(B)** HEK293 cells were transiently transfected with the plasmid encoding $FLAG_3$-STAG3 (or mutant $FLAG_3$-STAG3) and REC8-EGFP, or $FLAG_3$-STAG3 (or mutant $FLAG_3$-STAG3) and SMC1A-HA. Untransfected HEK293 cells were used as the negative control. The culture medium was supplemented with nocodazole for prometaphase arrest of the cells after 36 h. Total protein was extracted after 12 h for Co-IP analysis. Wild-type STAG3 interacted with both REC8 and SMC1A, whereas there was no interaction between the three mutant STAG3 proteins with REC8 or SMC1A. WT indicates the plasmid encoding the wild type STAG3, Mut 1 indicates the plasmid encoding c.877_885del (p.293_295del) STAG3, Mut 2 indicates the plasmid encoding c.891_893dupTGA (p.297_298insAsp) STAG3, Mut 3 indicates the plasmid simultaneously encoding c.877_885del (p.293_295del) and c.891_893dupTGA (p.297_298insAsp) STAG3, and EV indicates the empty vector (without *STAG3*).

p.297_298insAsp) of *STAG3* were identified in a consanguineous Han Chinese family with POI. A fluorescence localization analysis was employed to evaluate the effects of the two variants and revealed that the mutant STAG3 proteins led to aberrant localization of REC8. Furthermore, interactions between the mutant STAG3 proteins and REC8 or SMC1A were absent, which was confirmed by Co-IP analysis. Therefore, both in-frame variants were shown to be deleterious and associated with POI in this family.

The STAG3 protein contains two domains: the STAG domain and the armadillo (ARM)-type domain. The ARM-type domain is located after the STAG domain, which is predicted to interact with nucleic acid or another protein (Caburet et al., 2014) (**Figure 1E**). At present, six *STAG3* variants have been reported in five families with POI (Caburet et al., 2014; Stabej et al., 2016; Colombo et al., 2017; He et al., 2018a; Franca et al., 2019) (**Table 1**). All six variants are truncated variants that severely disrupt the STAG domain and/or ARM-type domain. Even c.1573+5G > A is a splice site variant, with RT-PCR revealing that the resulting mutant STAG3 protein is truncated (p.Leu490Thrfs*10) (He et al., 2018a), leading to destruction of the ARM-type domain. In our study, both variants (p.293_295del and p.297_298insAsp) are in-frame variants, and neither localizes to the STAG or ARM-type domain; both variants are positioned between these two domains (**Figure 1E**). Therefore, the STAG and ARM-type domains are not disrupted, in theory. However, the p.293_295 and p.297_298 sites are highly conserved across species (**Figure 1D**), and non-truncating variants located between two domains have also been reported to impair protein function and lead to disease (Xia et al., 2015). Therefore, we postulate that the two variants are pathogenic, and that the region connecting the STAG and ARM-type domains is significant for STAG3 function.

To verify the pathogenicity of the two variants, we performed *in vitro* functional analyses to determine whether STAG3 protein function was impaired. *STAG3* encodes a meiosis-specific subunit of cohesin, a multiprotein complex that plays an essential role in the proper pairing of chromosomes, sister chromatid cohesion, and chromosome segregation (Caburet et al., 2014; Wolf et al., 2018). Furthermore, STAG3 has been reported to interact with the cohesion subunits REC8 and SMC1A, and the entry of REC8 into nuclei requires STAG3 (Prieto et al., 2001; Wolf et al., 2018). The results of the fluorescence localization analysis in this study showed that none of the three mutant STAG3 proteins could assist REC8 in entering nuclei and that nearly all of the STAG3 mutants were restricted to the cytoplasm. In addition, there were no interactions between the mutant STAG3 proteins and REC8 or SMC1A observed in the Co-IP analysis. Therefore, we propose that the two in-frame variants of *STAG3* lead to STAG3 dysfunction and are responsible for POI in the two study patients.

The proband's affected sister is sterile and has a small uterus and streak gonads, which is consistent with patients with POI caused by other *STAG3* variants (**Table 1**). In contrast, the uterus and ovaries of the proband in our study were not visualized by ultrasonography (**Table 1**). The difference in phenotype between the proband and her affected sister might be explained by the proband's affected sister receiving HRT before coming to our hospital. Her uterus became larger after 3 months of HRT, which is consistent with patients with POI caused by variants in *MCM8* and *BRCA2* (Tenenbaumrakover et al., 2015; Weinbergshukron et al., 2018). Both *MCM8* and *BRCA2* have been shown to be associated with meiosis and female infertility (Moynahan et al., 2001; Rodriguezmari et al., 2011; Gambus and Blow, 2013), and *STAG3* is a meiosis-specific gene involved in female infertility. Furthermore, *MCM8* and *BRCA2* have been associated with ovarian dysgenesis in humans (Moynahan et al., 2001; Rodriguezmari et al., 2011; Gambus and Blow, 2013). Female *stag3*$^{-/-}$ mice also present with severe and very early ovarian dysgenesis (Caburet et al., 2014). Therefore, we suggest that *STAG3* might also be related to ovarian dysgenesis in humans.

In summary, to the best of our knowledge, we are the first researchers to report in-frame variants of *STAG3* that cause POI and to verify the role of in-frame *STAG3* variants in POI pathogenicity through *in vitro* functional analyses of cell model. Our findings extend the mutational and phenotypic spectrums of *STAG3* and have important implications for genetic counseling of patients with POI. However, to confirm the relationship between *STAG3* and ovarian dysgenesis in humans, further study with a large sample size is needed.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this manuscript will be made available by the authors, without undue reservation, to any qualified researcher.

AUTHOR CONTRIBUTIONS

JD and Y-QT designed the study. W-JX, W-BH, Y-XZ and L-LM performed the variant analysis of *STAG3*. W-JX and W-BH carried out the evaluation of the pathogenicity of variations by *in vitro* cell model functional analyses. G-XL and GL worked on the clinical study. W-JX, W-BH, Y-QT and JD wrote the paper. All authors read and approved the final manuscript.

FUNDING

This work was supported by grants from the National Key Research & Developmental Program of China (2018YFC1004900), the National Natural Science Foundation of China (81771645 and 81971447), the Hunan Provincial Natural Science Foundation of China (2019JJ51006), the Scientific Research Foundation of the Health Committee of Hunan Province (C2019193), the science and technology major project of the ministry of science and technology of Hunan Province, China (2017SK1030), and the Scientific Research Foundation of Reproductive and Genetic Hospital of CITIC-Xiangya (YNXM-201915, YNXM-201913, YNXM-201912, YNXM-201916).

ACKNOWLEDGMENTS

We thank all of the donors who participated in this program and all of our colleagues at Reproductive and Genetic Hospital of CITIC-Xiangya.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2019.01016/full#supplementary-material

REFERENCES

Bannatyne, P., Russell, P., and Shearman, R. P. (1990). Autoimmune oophoritis: a clinicopathologic assessment of 12 cases. *Int J. Gynecol. Pathol.* 9 (3), 191–207. doi: 10.1097/00004347-199007000-00001

Biscup, P. (2003). Risks and benefits of long-term hormone replacement therapy. *Am. J. Health-syst. Pharm.* 60 (14), 1419–1425. doi: 10.1093/ajhp/60.14.1419

Caburet, S., Arboleda, V. A., Llano, E., Overbeek, P. A., Barbero, J. L., Oka, K., et al. (2014). Mutant Cohesin in Premature Ovarian Failure. *N. Engl. J. Med.* 370 (10), 943–949. doi: 10.1056/NEJMoa1309635

Colombo, R., Pontoglio, A., and Bini, M. (2017). A STAG3 missense mutation in two sisters with primary ovarian insufficiency. *Eur. J. Obstet. Gynecol. Reprod. Biol.* 216, 269–271. doi: 10.1016/j.ejogrb.2017.08.005

Franca, M. M., Nishi, M. Y., Funari, M. F. A., Lerario, A. M., Baracat, E. C., Hayashida, S. A. Y., et al. (2019). Two rare loss-of-function variants in the STAG3 gene leading to primary ovarian insufficiency. *Eur. J. Med. Genet.* 62 (3), 186–189. doi: 10.1016/j.ejmg.2018.07.008

Gambus, A., and Blow, J. J. (2013). Mcm8 and Mcm9 form a dimeric complex in Xenopus laevis egg extract that is not essential for DNA replication initiation. *Cell. Cycle.* 12 (8), 1225–1232. doi: 10.4161/cc.24310

Garciacruz, R., Brieno, M. A., Roig, I., Grossmann, M., Velilla, E., Pujol, A., et al. (2010). Dynamics of cohesin proteins REC8, STAG3, SMC1β and SMC3 are consistent with a role in sister chromatid cohesion during meiosis in human oocytes. *Hum. Reprod.* 25 (9), 2316–2327. doi: 10.1093/humrep/deq180

Gowri, V., Al, S. M., Al-Farsi, F. A., Al-Busaidi, N. A., Dennison, D., Al, K. S., et al. (2015). Aetiological profile of women presenting with premature ovarian failure to a single tertiary care center in Oman. *Post. Reprod. Health.* 21 (2), 63–68. doi: 10.1177/2053369115587419

He, W., Banerjee, S., Meng, L., Du, J., Gong, F., Huang, H., et al. (2018a). Whole-exome sequencing identifies a homozygous donor splice-site mutation in STAG3 that causes primary ovarian insufficiency. *Clin. Genet.* 93 (2), 340–344. doi: 10.1111/cge.13034

He, W., Tu, C., Liu, Q., Meng, L., Yuan, S., Luo, A., et al. (2018b). DMC1 mutation that causes human non-obstructive azoospermia and premature ovarian insufficiency identified by whole-exome sequencing. *J. Med. Genet.* 55 (3), 198–204. doi: 10.1136/jmedgenet-2017-104992

Houmard, B. S., Small, C., Yang, L., Naluaicecchini, T., Cheng, E., Hassold, T. J., et al. (2009). Global gene expression in the human fetal testis and ovary. *Biol. Reprod.* 81 (2), 438–443. doi: 10.1095/biolreprod.108.075747

Moynahan, M. E., Pierce, A. J., and Jasin, M. (2001). BRCA2 is required for homology-directed repair of chromosomal breaks. *Mol. Cell* 7 (2), 263–272. doi: 10.1016/S1097-2765(01)00174-5

Nogues, C., Fernandez, C., Rajmil, O., and Templado, C. (2009). Baseline expression profile of meiotic-specific genes in healthy fertile males. *Fertil. Steril.* 92 (2), 578–582. doi: 10.1016/j.fertnstert.2008.06.034

Prieto, I., Suja, J. A., Pezzi, N., Kremer, L., Martineza, C., Rufas, J. S., et al. (2001). Mammalian STAG3 is a cohesin specific to sister chromatid arms in meiosis I. *Nat. Cell Biol.* 3 (8), 761–766. doi: 10.1038/35087082

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., et al. (2015). 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.* 17 (5), 405–424. doi: 10.1038/gim.2015.30

Rodriguezmari, A., Wilson, C., Titus, T. A., Canestro, C., Bremiller, R., Yan, Y., et al. (2011). Roles of brca2 (fancd1) in oocyte nuclear architecture, gametogenesis, gonad tumors, and genome stability in zebrafish. *PLoS Genet* 7 (3), e1001357. doi: 10.1371/journal.pgen.1001357

Rossetti, R., Ferrari, I., Bonomi, M., and Persani, L. (2017). Genetics of primary ovarian insufficiency. *Clin. Genet.* 91 (2), 183–198. doi: 10.1111/cge.12921

Shelling, A. N. (2010). Premature ovarian failure. *Reprod.* 140 (5), 633–641. doi: 10.1530/REP-09-0567

Stabej, P. L. Q., Williams, H. J., James, C., Tekman, M., Stanescu, H., Kleta, R., et al. (2016). STAG3 truncating variant as the cause of primary ovarian insufficiency. *Eur. J. Hum. Genet.* 24 (1), 135–138. doi: 10.1038/ejhg.2015.107

Tenenbaumrakover, Y., Weinbergshukron, A., Renbaum, P., Lobel, O., Eideh, H., Gulsuner, S., et al. (2015). Minichromosome maintenance complex component 8 (MCM8) gene mutations result in primary gonadal failure. *J. Med. Genet.* 52 (6), 391–399. doi: 10.1136/jmedgenet-2014-102921

Tucker, E. J., Grover, S., Bachelot, A., Touraine, P., and Sinclair, A. H. (2016). Premature ovarian insufficiency: new perspectives on genetic cause and phenotypic spectrum. *Endocr. Rev.* 37 (6), 609–635. doi: 10.1210/er.2016-1047

Webber, L., Davies, M., Anderson, R. A., Bartlett, J., Braat, D., Cartwright, B., et al. (2016). ESHRE Guideline: management of women with premature ovarian insufficiency. *Hum. Reprod.* 31 (5), 926–937. doi: 10.1093/humrep/dew027

Weinbergshukron, A., Rachmiel, M., Renbaum, P., Gulsuner, S., Walsh, T., Lobel, O., et al. (2018). Essential role of brca2 in ovarian development and function. *N. Engl. J. Med.* 379 (11), 1042–1049. doi: 10.1056/NEJMoa1800024

Wolf, P. G., Ramos, A. C., Kenzel, J., Neumann, B., and Stemmann, O. (2018). Studying meiotic cohesin in somatic cells reveals that Rec8-containing cohesin requires Stag3 to function and is regulated by Wapl and sororin. *J. Cell Sci.* 131 (11), jcs212100. doi: 10.1242/jcs.212100

Xia, H., Huang, X., Guo, Y., Hu, P., He, G., Deng, X., et al. (2015). Identification of a novel MYO15A mutation in a chinese family with autosomal recessive nonsyndromic hearing loss. *PLoS One* 10 (8), e0136306. doi: 10.1371/journal.pone.0136306

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Genetics

CASE REPORT
published: 25 November 2019
doi: 10.3389/fgene.2019.01201

Whole-Exome Sequencing Revealed Mutations of *MED12* and *EFNB1* in Fetal Agenesis of the Corpus Callosum

*Ying Jiang[1], Ye-Qing Qian[1], Meng-Meng Yang[1], Qi-Tao Zhan[1], Yuan Chen[1], Fang-Fang Xi[1], Matthew Sagnelli[2], Min-Yue Dong[1], Bai-Hui Zhao[1] and Qiong Luo[1]**

[1] Department of Obstetrics, Women's Hospital, School of Medicine, Zhejiang University, Hangzhou, China, [2] University of Connecticut School of Medicine, Farmington, CT, United States

Agenesis of the corpus callosum (ACC) is a birth defect in which the corpus callosum is either partially or completely missing. With recent advances in prenatal ultrasound, detection of ACC in obstetric practices is becoming more common. Etiologies of ACC include chromosome errors, genetic factors, prenatal infections, and other factors related to the prenatal environment. In an effort to elucidate more about the genetic influence in the pathogenesis of ACC, we identified, through whole-exome sequencing (WES), two gene mutations in two families with complete agenesis of the corpus callosum. These two mutations are located on chromosome X: one is a hemizygous missense mutation c.3746T>C (p. L1249P) in the gene mediator complex subunit 12 (*MED12*); the other one is a heterozygous missense mutation c.128+5G>C in gene ephrin B1 (*EFNB1*). Historically, early diagnosis of complete ACC during pregnancy has been difficult; however, WES has provided us with a creative avenue of diagnosis, combining identification of genetic mutations with prenatal imaging.

Keywords: whole-exome sequencing, *de novo* mutation, fetal agenesis of the corpus callosum, Sanger sequencing, prenatal diagnosis

OPEN ACCESS

Edited by:
Yueqiu Tan,
Central South University, China

Reviewed by:
Yetao Xu,
Nanjing Medical University, China
Xi Xia,
Peking University Shenzhen Hospital, China

***Correspondence:**
Qiong Luo
luoq@zju.edu.cn

Specialty section:
This article was submitted to Genetic Disorders, a section of the journal Frontiers in Genetics

Received: *09 June 2019*
Accepted: *30 October 2019*
Published: *25 November 2019*

Citation:
Jiang Y, Qian Y-Q, Yang M-M, Zhan Q-T, Chen Y, Xi F-F, Sagnelli M, Dong M-Y, Zhao B-H and Luo Q (2019) Whole-Exome Sequencing Revealed Mutations of MED12 and EFNB1 in Fetal Agenesis of the Corpus Callosum. Front. Genet. 10:1201. doi: 10.3389/fgene.2019.01201

INTRODUCTION

The corpus callosum (CC) is one of the five main cerebral commissures connecting the left and right cerebral hemispheres. Corpus callosum abnormalities are separated into three categories: complete agenesis, partial agenesis, and dysgenesis (Palmer and Mowat, 2014). The prevalence of agenesis of the corpus callosum (ACC) is around 1.4 per 10,000 live births (Glass et al., 2008). ACC may be associated with intellectual disabilities, epilepsy, behavioral difficulties, and cognitive impairments (Moutard et al., 2003). Many couples decide to terminate the pregnancy on the basis of ultrasound (US)/magnetic resonance imaging (MRI) findings in ACC. However, follow-up studies indicate that 71.2% of isolated ACC cases have normal intelligence, 13.6% have borderline or moderate intellectual disabilities, and 15.2% have severe intellectual disability (Santo et al., 2012). These findings highlight the importance of improving the precision of the diagnosis of ACC as well as moving the window of diagnosis earlier in pregnancy.

Studies have shown that the majority of ACC cases are caused by *de novo* dominant mutations (De Ligt et al., 2012). The Netrin receptor *DCC* (*Deleted in colorectal cancer*) gene is a critical factor in corpus callosum development in mice and works through orienting callosal axons at the midline (Fazeli et al., 1997). *DCC* mutations have been shown to cause isolated ACC, including the reported mirror movement (MM) phenotype, in humans (Marsh et al., 2017). Additionally, studies have

shown that fibroblast growth factor receptor/glial fibrillary acidic protein (*Fgfr1/Gfap*) play an important role in corpus callosum formation (Smith et al., 2006). Recently, whole-exon sequencing has provided the opportunity for molecular genetic screening of rare human diseases.

In the present study, we performed whole-exome sequencing in two families who had children who were affected by complete ACC. With whole-exon sequencing and Sanger sequencing, we identified two single-nucleotide missense mutations which are likely related to the pathogenesis of ACC.

CLINICAL REPORT

Proband 1

A fetal ultrasound of a 21-year-old G1P0 female at 26 weeks gestation revealed agenesis of the corpus collosum and a right lateral ventricle about 1.0 cm in width. A subsequent brain MRI performed at 28 weeks and 5 days gestation, determined by last menstrual period (LMP), confirmed complete callosal agenesis and showed mild ventriculomegaly. Both lateral ventricles had a width of approximately 11 mm, were parallel, and teardrop-shaped. The third ventricle was slightly widened and uplifted. No obvious corpus callosum signals were observed in each section. The cisterna magna was about 9 mm in width, and the vermis cerebelli was present (**Figures 1A–C**). A female fetus was delivered at 31 weeks. There is no anatomic pathology data for the fetus.

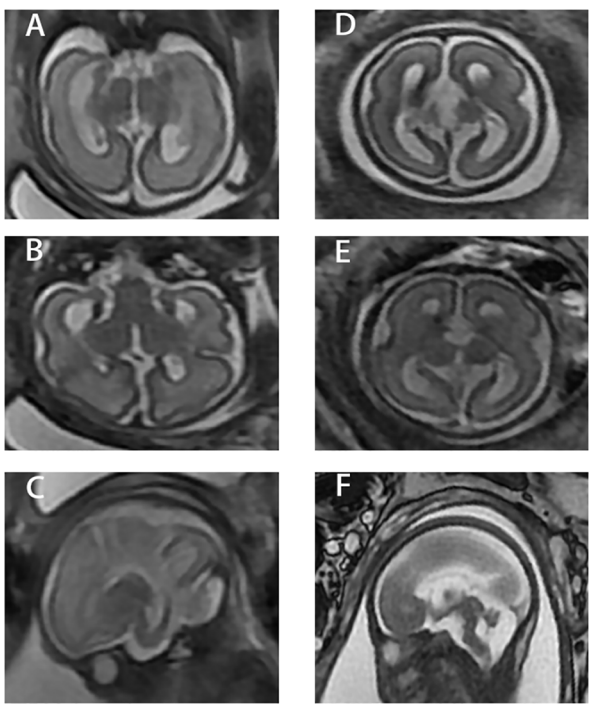

FIGURE 1 | Three different sections of the two cases. **(A–C)** Case 1. **(D–F)** Case 2. 1) **(A)** and **(D)** are coronal images from fetal MRI. 2) **(B)** and **(E)** are axial images from fetal MRI. 3) **(C)** and **(F)** are sagittal images from fetal MRI.

Proband 2

A 33-year-old G1P0 female presented at 26 weeks gestation for a fetal ultrasound which demonstrated agenesis of the corpus callosum. A subsequent brain MRI performed at 27 weeks, determined by LMP, confirmed complete callosal agenesis complicated by variant Dandy–Walker malformation. Specifically, the posterior horn was about 0.6 cm in width, the bilateral lateral ventricles were parallel, and the third ventricle was widened and uplifted. There was no obvious corpus callosum in each section. The posterior cranial fossa was about 0.8 cm in width and the fourth ventricle was about 0.9 cm in width. The superior vermis was present while the lower vermis was absent (**Figures 1D–F**). The male fetus was delivered at 27 weeks, 5 days gestation. There is no anatomic pathology data for the fetus.

Chromosomal array (CMA) result of both the probands using CytoScan™ HD whole-genome SNP array (Affymetrix, USA) showed no disease-related copy number variations (CNVs). The reporting threshold of the copy number result was set at 500 kb with marker count ≥50 for gains and 200 kb with marker count ≥50 for losses.

MATERIALS AND METHODS

Ethics Approval for the Investigation

The Ethics Committee of the Women's Hospital, Zhejiang University, School of Medicine (Hangzhou, China) approved our study. The investigation conforms to the principles outlined in the Helsinki Declaration. Written informed consent was obtained from the patients and their spouses whose child's tissue was collected.

DNA Extraction

A 5-ml aliquot of peripheral whole blood was collected from the parents and EDTA was added for anticoagulation. Additionally, a 1 × 1 × 1-cm medial thigh muscle sample was gathered from the induced fetus. Genomic DNA from 2 ml of peripheral whole blood was extracted using a Qiagen Blood DNA mini kit (Qiagen®, Hilden, Germany). DNA from muscle was extracted with a Genomic DNA Purification Kit (Invitrogen, cat. K0512, USA) and then preserved at -20°C. The remaining samples were stored at -80°C.

Whole-Exome Sequencing and Data Analysis

Target enrichment of target region sequences was carried out by the Agilent SureSelect Human Exon Sequence Capture Kit. The sequencing libraries were quantified using the Illumina DNA Standards and Primer Premix Kit (Kapa Biosystems, Boston, MA, USA), massively parallel sequenced using the Illumina HiSeq 2500 platform (Illumina, San Diego, CA, USA), and then massively parallel sequenced again using the Illumina HiSeq 2500 platform (Illumina, San Diego, CA, USA). After sequencing and filtering out low-quality reads, high-quality reads were compared to the human genome reference sequence (GRCh37.p12, hg19). The GATK software was used to call variants.

Data Filtering

In our study, we selected variants in fetuses by complying with the following criteria: 1) mutation frequency ≥0.01, quality <300, and alt/depth <25%; 2) MAF > 0.02 when referring to dbSNP (https://www.ncbi.nlm.nih.gov/projects/SNP/), the 1000 Genomes Project (http://www.1000genomes.org/), and the ExAC Browser (http://exac.broadinstitue.org/); 3) variants located in deep intron regions; 4) variants reported benign by ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/); and 5) depth >20 and alt/all <0.1. Subsequently, the variants left were further selected according to the following criteria: 1) deleterious protein or splicing predictions; 2) variants in a gene responsible for an Online Mendelian Inheritance in Man (OMIM) disease/phenotype associated with the probands; and 3) disease inheritance models: genes associated with autosomal-recessive disease (AR), defined as variants in homozygous or compound heterozygous fetuses with parents who are both heterozygous carriers; genes associated with autosomal-dominant disease (AD), defined as *de novo* variants in fetuses or variants inherited from either of the parents for non-lethal or late-onset diseases; genes associated with X-linked recessive disease (XLR), characterized by hemizygous male fetuses with heterozygous mothers; and genes associated with X-linked dominant disease, characterized by *de novo* variants in fetuses or variants in fetuses that were inherited from either of the parents in the case of non-lethal or late-onset disease. The identified novel mutations were further assessed for possible pathogenicity using PolyPhen-2, SIFT, PROVEAN, HSF, regSNP-intron, and MutationTaster. The selected variants were classified by following the American College of Medical Genetics and Genomics/Association for Molecular Pathology (ACMG/AMP) guidelines. All putative disease-causing variants detected by new-generation sequencing (NGS) were confirmed in probands and their parents by Sanger sequencing. Sequencing was performed on ABI 3500xL (Applied Biosystems, Foster City, CA, USA), and the results were analyzed using DNASTAR software.

RESULTS

Identification of Variants on Two Probands

In proband 1, with the average coverage of 132.28X on targeted regions, we confirmed 95,687 novel genetic variants (GVs) through whole-exome sequencing (WES), while for proband 2, under mean coverage of 71.63X, we identified 79,974 GVs. The summary of WES for the two probands is shown in **Table 1**.

Pathogenic Mutations on Two Probands

In proband 1, we identified 23 SNPs in callosum agenesis-related genes compared to 11 SNPs identified in proband 2. Two pathogenic mutations were identified: chrX: 70349584: T:C (c.3746T>C) in exon 25 (**Figure 2A**) and chrX: 68049752: G:C (c.128+5G>C) in intron 1 (**Figure 2C**); These two mutations were further confirmed by Sanger sequencing (**Figures 2B**, **D**).

TABLE 1 | Summary of whole-exome sequencing on these two probands.

	Proband 1	Proband 2
Total captured region size	60 M	60 M
On-Target Reads (%)	73.98	72.6
% of captured regions with coverage >20	98.6	96.26
Mean coverage of target region	132.28	71.63
Total number of SNPs	348,832	218,408
Total number of INDELs	37,565	21,246
Total number of novel GVs not listed in dbSNP	95,687	79,974
Total number of SNPs in a panel of corpus callosum agenesis-related genes	23	11
Total number of INDELs in a panel of corpus callosum agenesis-related genes	1	3

SNP, single-nucleotide polymorphism; INDEL, insertions and deletions; GV, genetic variants.

MED12: c.3746T>C of proband 1 was inherited from his healthy mother (**Figure 2B**), while *EFNB1*: c.128+5G>C of proband 2 appeared *de novo* (**Figure 2D**).

The T-to-C transversion at position 3746 resulted in a leucine-to-proline substitution at code1249 in the mediator complex subunit12 (*MED12*) protein (p. L1249P) (**Figure 3**), a gene which is highly conserved between different species. However, the missense mutation of *EFNB1* located in intron 1 did not result in amino acid transversion; however, it was located 5 bp behind exon 1 of *EFNB1*. The HSF database strongly supports the pathogenicity of *EFNB1*: c.128+5G>C, indicating it might influence the RNA splicing.

DISCUSSION

In this report, we described two families whose fetuses were diagnosed with complete agenesis of the corpus callosum. With whole-exon sequencing, we identified one hemizygous missense mutation in exon 27 of *MED12* on the X chromosome and one *de novo* heterozygous missense mutation in intron 1 of *EFNB1* on the X chromosome.

Currently, fetal MRI and ultrasound are useful tools in the diagnosis of complete ACC (cACC) in second- and third-trimester pregnancies (Tang et al., 2009). Genetic diagnosis with whole-exon sequencing offers the potential for an earlier diagnosis of cACC accompanied by the option for earlier termination of pregnancy. Though autosomal-dominant, autosomal-recessive, and X-linked patterns have been described as causes of ACC (Edwards et al., 2014), there has been no clear inheritance model found in the majority of cases, Therefore, it is likely that most cases arise from a *de novo* mutation. Some cases of partial ACC (OMIM, 304100) follow an X-linked recessive pattern caused by a mutation in the L1 cell adhesion molecule (*L1CAM*) gene on chromosome Xq28, which only affects male fetus. Additionally, Bassuk and Sherr reported a *de novo* mutation c.427T>G (S143A) of gene *PRICKLE1* in fetal agenesis of the corpus callosum without other possible mutations (Bassuk and Sherr, 2015).

In our study, we identified another X-linked recessive mutation in the *MED12* gene associated with cACC. *MED12*

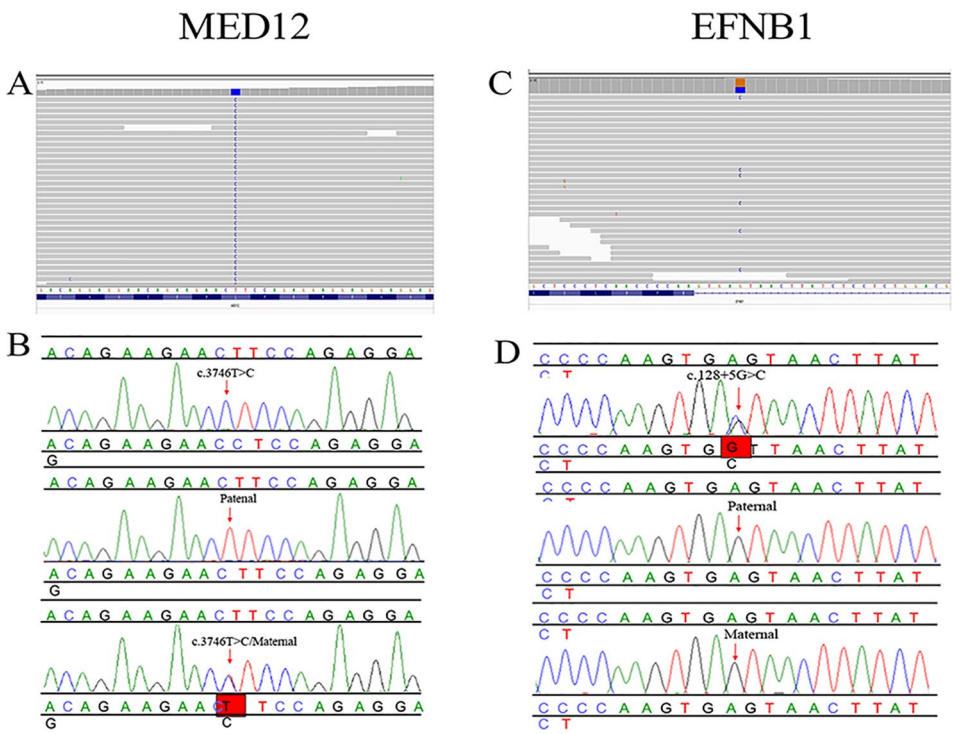

FIGURE 2 | Pedigree and mutation analysis of the family. **(A, C)** Alignment of exon sequences to Hg19 showing SNP in exon 25 of *MED12* and intron 1 of *EFNB1*. **(B, D)** Chromatogram identified by Sanger sequencing. The *top* chromatograms are from the fetus, the *middle two* are from the father, and the *bottom two* are from the mother.

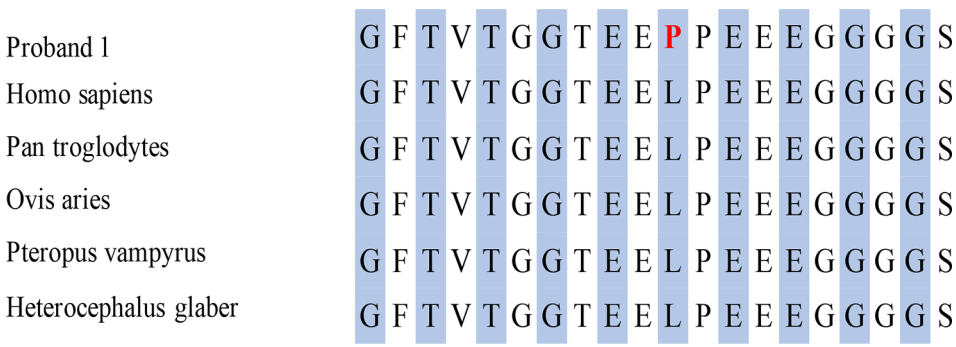

FIGURE 3 | Conserved amino acid sequence of *MED12* and predicted amino acid transversion at position 3746 caused by missense mutation identified in this proband.

encodes one of the subunits of the Mediator complex, and MED12 has a role in RNA polymerase II transcription (Donnio et al., 2017). Hong et al. showed that med12-deficient zebrafish embryos showed defects in brain, neural crest, and kidney development (Hong et al., 2005). To date, at least 39 different mutations of *MED12* have been identified in patients with X-linked intellectual disability (XLID), including Lujan–Fryns syndrome (c.3020A>G, p. N1007S) [OMIM 309520], Opitz–Kaveggia or FG syndrome (FGS1)(c.2881C>T, p. R961W) [OMIM 305450], and the X-linked Ohdo syndrome [OMIM 300895] (Schwartz et al., 2007; Prontera et al., 2016).While these syndromes are caused by mutations along different positions of the *MED12* gene, they may be related with a deficient or absent corpus (Graham and Schwartz, 2013). The *EFNB1* mutation in our study follows an X-linked dominant inheritance pattern and has been mostly described in patients with craniofrontonasal syndrome (CFNS; OMIM 304110) (Wallis et al., 2008; Hogue et al., 2010). Features of the disease include frontonasal dysplasia, craniofacial asymmetry, craniosynostosis, bifid nasal tip, grooved nails, and abnormalities of the thoracic skeleton. It has been shown that *EFNB1* plays an important role in defining the position of

the coronal suture (Twigg et al., 2004). Yet, to our knowledge, these two mutations have not been reported in literature. Our study expended the clinical spectrum of abnormalities associated with mutations in the X chromosome in humans. Since both of these two mutations are located in chromosome X, we inferred that male fetuses are more likely to develop complete agenesis of the corpus callosum. This finding is in accordance with a previous large cohort study on patients with agenesis of the corpus callosum that showed a slightly increased prevalence in males (56% vs. 44%) (Romaniello et al., 2017).

Since we collected samples from the expired fetuses *in utero*, there is no information regarding the intellectual function of these probands. Therefore, it was difficult to connect the mutation identified closely to intellectual disabilities, epilepsy, behavioral difficulties, or cognitive impairments in childhood or adulthood. In future studies, we are planning to carry out a large cohort analysis of agenesis of the corpus callosum with these mutations to further confirm our results and establish a scientific significance for genotype–phenotype correlation.

CONCLUSION

Fast development of high-throughput sequencing provides our clinicians with a new pathway to explore the potential for genetic diagnosis of fetal ACC. Whole exon sequencing combined with Sanger sequencing demonstrated mutations of *MED12* and *EFNB1* associated with cACC and strongly implicates these two genes as one causative explanation for fetal cACC.

DATA AVAILABILITY STATEMENT

The SRA accession number for the Illumina sequencing is PRJNA549835. All other data is available on request.

AUTHOR CONTRIBUTIONS

YJ, Y-QQ and QL designed and wrote the paper. M-MY and Q-TZ collected clinical information, YC, F-FX and M-YD conducted the bioinformatics analysis. B-HZ revised the manuscript. MS provided the language editing. QL contributed to the study design and data interpretation and the manuscript preparation. QL is the guarantor of the work and, as such, has full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

FUNDING

This work was supported by the Key Subjects Group of Reproductive Medicine, School of Medicine, Zhejiang University, the National Nature Science Foundation of China (grant 81571447), and the Medical and Health Technology Program (New Technology Investigation Program) in Zhejiang Province (grant 2019321755).

ACKNOWLEDGMENTS

The authors thank the staff at Women's Hospital of Zhejiang University for technical assistance and facility support.

REFERENCES

Bassuk, A. G., and Sherr, E. H. (2015). A *de novo* mutation in PRICKLE1 in fetal agenesis of the corpus callosum and polymicrogyria. *J. Neurogenet.* 29, 174–177. doi: 10.3109/01677063.2015.1088847

De Ligt, J., Willemsen, M. H., Van Bon, B. W., Kleefstra, T., Yntema, H. G., Kroes, T., et al. (2012). Diagnostic exome sequencing in persons with severe intellectual disability. *N. Engl. J. Med.* 367, 1921–1929. doi: 10.1056/NEJMoa1206524

Donnio, L. M., Bidon, B., Hashimoto, S., May, M., Epanchintsev, A., Ryan, C., et al. (2017). MED12-related XLID disorders are dose-dependent of immediate early genes (IEGs) expression. *Hum. Mol. Genet.* 26, 2062–2075. doi: 10.1093/hmg/ddx099

Edwards, T. J., Sherr, E. H., Barkovich, A. J., and Richards, L. J. (2014). Clinical, genetic and imaging findings identify new causes for corpus callosum development syndromes. *Brain* 137, 1579–1613. doi: 10.1093/brain/awt358

Fazeli, A., Dickinson, S. L., Hermiston, M. L., Tighe, R. V., Steen, R. G., Small, C. G., et al. (1997). Phenotype of mice lacking functional Deleted in colorectal cancer (Dcc) gene. *Nature* 386, 796–804. doi: 10.1038/386796a0

Glass, H. C., Shaw, G. M., Ma, C., and Sherr, E. H. (2008). Agenesis of the corpus callosum in California 1983-2003: a population-based study. *Am. J. Med. Genet. A* 146A, 2495–2500. doi: 10.1002/ajmg.a.32418

Graham, J. M.Jr., and Schwartz, C. E. (2013). MED12 related disorders. *Am. J. Med. Genet. A* 161A, 2734–2740. doi: 10.1002/ajmg.a.36183

Hogue, J., Shankar, S., Perry, H., Patel, R., Vargervik, K., and Slavotinek, A. (2010). A novel EFNB1 mutation (c.712delG) in a family with craniofrontonasal syndrome and diaphragmatic hernia. *Am. J. Med. Genet. A* 152A, 2574–2577. doi: 10.1002/ajmg.a.33596

Hong, S. K., Haldin, C. E., Lawson, N. D., Weinstein, B. M., Dawid, I. B., and Hukriede, N. A. (2005). The zebrafish kohtalo/trap230 gene is required for the development of the brain, neural crest, and pronephric kidney. *Proc. Natl. Acad. Sci. U.S.A.* 102, 18473–18478. doi: 10.1073/pnas.0509457102

Marsh, A. P., Heron, D., Edwards, T. J., Quartier, A., Galea, C., Nava, C., et al. (2017). Mutations in DCC cause isolated agenesis of the corpus callosum with incomplete penetrance. *Nat. Genet.* 49, 511–514. doi: 10.1038/ng.3794

Moutard, M. L., Kieffer, V., Feingold, J., Kieffer, F., Lewin, F., Adamsbaum, C., et al. (2003). Agenesis of corpus callosum: prenatal diagnosis and prognosis. *Childs Nerv. Syst.* 19, 471–476. doi: 10.1007/s00381-003-0781-6

Palmer, E. E., and Mowat, D. (2014). Agenesis of the corpus callosum: a clinical approach to diagnosis. *Am. J. Med. Genet. C Semin. Med. Genet.* 166C, 184–197. doi: 10.1002/ajmg.c.31405

Prontera, P., Ottaviani, V., Rogaia, D., Isidori, I., Mencarelli, A., Malerba, N., et al. (2016). A novel MED12 mutation: Evidence for a fourth phenotype. *Am. J. Med. Genet. A* 170, 2377–2382. doi: 10.1002/ajmg.a.37805

Romaniello, R., Marelli, S., Giorda, R., Bedeschi, M. F., Bonaglia, M. C., Arrigoni, F., et al. (2017). Clinical characterization, genetics, and long-term follow-up of a large cohort of patients with Agenesis of the Corpus Callosum. *J. Child Neurol.* 32, 60–71. doi: 10.1177/0883073816664668

Santo, S., D'antonio, F., Homfray, T., Rich, P., Pilu, G., Bhide, A., et al. (2012). Counseling in fetal medicine: agenesis of the corpus callosum. *Ultrasound Obstet. Gynecol.* 40, 513–521. doi: 10.1002/uog.12315

Schwartz, C. E., Tarpey, P. S., Lubs, H. A., Verloes, A., May, M. M., Risheg, H., et al. (2007). The original Lujan syndrome family has a novel missense mutation (p.N1007S) in the MED12 gene. *J. Med. Genet.* 44, 472–477. doi: 10.1136/jmg.2006.048637

Smith, K. M., Ohkubo, Y., Maragnoli, M. E., Rasin, M. R., Schwartz, M. L., Sestan, N., et al. (2006). Midline radial glia translocation and corpus callosum formation require FGF signaling. *Nat. Neurosci.* 9, 787–797. doi: 10.1038/nn1705

Tang, P. H., Bartha, A. I., Norton, M. E., Barkovich, A. J., Sherr, E. H., and Glenn, O. A. (2009). Agenesis of the corpus callosum: an MR imaging analysis of associated abnormalities in the fetus. *AJNR Am. J. Neuroradiol.* 30, 257–263. doi: 10.3174/ajnr.A1331

Twigg, S. R., Kan, R., Babbs, C., Bochukova, E. G., Robertson, S. P., Wall, S. A., et al. (2004). Mutations of ephrin-B1 (EFNB1), a marker of tissue boundary formation, cause craniofrontonasal syndrome. *Proc. Natl. Acad. Sci. U.S.A.* 101, 8652–8657. doi: 10.1073/pnas.0402819101

Wallis, D., Lacbawan, F., Jain, M., Der Kaloustian, V. M., Steiner, C. E., Moeschler, J. B., et al. (2008). Additional EFNB1 mutations in craniofrontonasal syndrome. *Am. J. Med. Genet. A* 146A, 2008–2012. doi: 10.1002/ajmg.a.32388

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: info@frontiersin.org | +41 21 510 17 00

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership