# IMPROVING MEDICAL DIAGNOSIS IN RARE DISEASES

EDITED BY: Natália Duarte Linhares, Alfredo Brusco and Kathleen M. Gorman

PUBLISHED IN: Frontiers in Genetics and Frontiers in Pediatrics

frontiers Research Topics

# frontiers

**Frontiers eBook Copyright Statement**

The copyright in the text of individual articles in this eBook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this eBook is the property of Frontiers.

Each article within this eBook, and the eBook itself, are published under the most recent version of the Creative Commons CC-BY licence. The version current at the date of publication of this eBook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or eBook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714
ISBN 978-2-83250-248-8
DOI 10.3389/978-2-83250-248-8

## 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: frontiersin.org/about/contact

# IMPROVING MEDICAL DIAGNOSIS IN RARE DISEASES

Topic Editors:
**Natália Duarte Linhares,** Genuity Science, Ireland
**Alfredo Brusco,** University of Turin, Italy
**Kathleen M. Gorman,** Temple Street Children's University Hospital, Ireland

**Citation:** Linhares, N. D., Brusco, A., Gorman, K. M., eds. (2022). Improving Medical Diagnosis in Rare Diseases. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-83250-248-8

# Table of Contents

frontiers | Frontiers in Genetics

TYPE Editorial
PUBLISHED 06 September 2022
DOI 10.3389/fgene.2022.974129

Check for updates

OPEN ACCESS

EDITED AND REVIEWED BY
Erica E. Davis,
Stanley Manne Children's Research Institute, United States

*CORRESPONDENCE
Natália Duarte Linhares,
natalia.linhares@genuitysci.com

SPECIALTY SECTION
This article was submitted to Genetics of Common and Rare Diseases, a section of the journal Frontiers in Genetics

RECEIVED 20 June 2022
ACCEPTED 15 August 2022
PUBLISHED 06 September 2022

CITATION
Linhares ND, Gorman KM and Brusco A (2022), Editorial: Improving medical diagnosis in rare diseases. *Front. Genet.* 13:974129. doi: 10.3389/fgene.2022.974129

COPYRIGHT
© 2022 Linhares, Gorman and Brusco. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# Editorial: Improving medical diagnosis in rare diseases

Natália Duarte Linhares[1]*, Kathleen M Gorman[2,3] and Alfredo Brusco[4,5]

[1]Genuity Science, Dublin, Ireland, [2]Department of Neurology and Clinical Neurophysiology, Children's Health Ireland at Temple St., Dublin, Ireland, [3]School of Medicine and Medical Science, University College Dublin, Dublin, Ireland, [4]Department of Medical Sciences, University of Turin, Turin, Italy, [5]Medical Genetics Unit, Città della Salute e della Scienza University Hospital, Turin, Italy

KEYWORDS

**Rare disease (RD), whole-exome sequencing (WES), whole-genome sequencing (WGS), next-generation sequencing (NGS), diagnosis**

Editorial on the Research Topic
Improving medical diagnosis in rare diseases

## Introduction

A rare disease (RD) is defined as a disorder affecting fewer than 1 in 2,000 people (European Union, 2000). Although each is rare, they cumulatively involve greater than 300 million people worldwide (Nguengang Wakap et al., 2020). There are approximately 7,000 RDs reported by the Online Mendelian Inheritance in Man (OMIM) database, and 6,172 unique RDs by Orphanet, 70% of whom have childhood onset (Amberger et al., 2015; Nguengang Wakap et al., 2020).

One of the many challenges that persons living with an RD and their families face is the search for a diagnosis. Given their rarity, large number of disease entities and heterogeneous manifestation, patients often remain undiagnosed or misdiagnosed for years. The correct diagnosis has enormous implications for the patient and their family; it offers insight into the disease cause, natural history, and prognosis. It can facilitate contact with support groups and provides families with informed genetic counselling. Moreover, it can increase the understanding of the disease pathogenesis and thus improve therapeutic strategies.

More than 70% of RDs are genetic in origin (Nguengang Wakap et al., 2020). In the last decade, next-generation sequencing (NGS) technologies such as whole-exome sequencing (WES) and whole-genome sequencing (WGS) have enabled the identification of disease-causing variants in RDs. The diagnostic yields of NGS in RDs varies (24–68%) depending on the technology employed, the disease group studied, the patient age groups, clinical indications, family structures and variant types analysed (Clark et al., 2018). Genome test interpretation and reporting represents one of the challenges to laboratories seeking to implement or maximize the diagnostic yield of NGS,

and guidelines have just been published in order to address this gap (Austin-Tse et al., 2022). Even though NGS affords us a better opportunity to elucidate the genetic causes of RDs, a proportion of individuals remain undiagnosed. Therefore, within this Research Topic, we created a collection of articles demonstrating emerging approaches to facilitate genetic diagnosis and discussing the impact of identifying a genetic cause in RDs. The research topic was published in September 2020 and manuscripts were received until February 2022. It received 72 manuscripts, and 23 submissions were accepted after a rigorous and constructive reviewing process.

## Studies showing applications of NGS as a frontline test in healthcare settings to accelerate rare disease diagnosis

Time-to-diagnosis is an important metric to consider in children with RDs, especially to neonates admitted to intensive care settings. Liu et al. focus on the clinical application of WES in critically ill patients in the pediatric intensive care units; the median age of children enrolled was 10.5 months (range, 1 month to 14.8 years). Whole-exome sequencing was performed in 169 critically ill children, and approximately a quarter of them were diagnosed with a monogenic disorder. They show that WES played a key role in family decision-making and clinical treatment and may improve the prognosis of some children and reduce the economic burden on families and the society.

## Bioinformatic and molecular analysis approaches and results from genomics centers or diagnosis laboratories

Santos et al. describe a new protocol for the procedure of preconception screening of consanguineous couples by WES. They considered all genes listed in the Clinical Genomics Database as causally related to autosomal recessive (AR) diseases and analysed 39 consanguineous couples using bioinformatic filters to identify pathogenic variants that were present on the same gene in both members of the couple. In 21 couples (53.8%), they ascertained sharing of heterozygosity for at least one variant considered pathogenic for an AR disease. Once the specific pathogenic variant was identified, it became possible for the couple to undergo prenatal diagnosis or, if desired, preimplantation genetic diagnosis.

Botta et al. studied a total of 570 individuals with a clinical suspect of Myotonic dystrophy type 2 (DM2), which is an autosomal dominant multisystemic disorder caused by a (CCTG)n repeat expansion in intron 1 of *CNBP*. The DM2 locus was analysed by a combination of short-range PCR (SR-PCR), tetraplet-primed PCR (TP-PCR), long-range PCR (LR-PCR), and Sanger sequencing of *CNBP* alleles. DM2 molecular diagnosis was confirmed in 187 samples analysed (32.8%) and was mainly associated with the presence of myotonia in patients.

Mellone et al. show the utility of using a targeted NGS gene panel including 221 genes in providing molecular diagnosis as a second tier-test by detecting clinically relevant variants in 71 out of 338 (21%) patients with Neurodevelopmental Disorders (NDDs). Their findings show that this NGS panel represents a powerful and affordable clinical tool, significantly increasing the diagnostic yield in patients with different forms of NDDs in a cost- and time-effective manner.

The results of Santos et al., Botta et al. and Mellone et al. demonstrate their methods have diagnostic power, and may provide useful information and potential clinical benefits for other genomic centers offering preconception screening by WES, DM2 genetic testing, or customized NGS panels for NDD, respectively.

## Applications of methodologies to promote early genetic diagnosis

Due to advances in molecular diagnosis, individuals with a RD may even be diagnosed *in utero*. Xue et al. identified variants in *TALDO1* in a fetus with Transaldolase Deficiency, supporting the application of WES in prenatal diagnosis, and further supporting that effective postpartum treatments could improve prognosis.

The early molecular diagnosis is particularly important for treatable disorders, like Congenital adrenal hyperplasia (CAH). Tolba et al. show that Multiplex ligation-dependent probe amplification (MLPA) of *CYP21A2* may improve the diagnosis and management of missed cases with atypical CAH presentations. They analysed 112 unrelated Egyptian children with CAH and 79.5% of the patients were diagnosed within the first month of life.

Liu et al. report the clinical features of eight fetuses with pathogenic variants in *CPLANE1*, *TMEM67*, *NPHP4*, and *DYNC2H1* identified by WES analysis, expanding the prenatal clinical manifestations of ciliopathies. The fetuses showed prenatal diagnostic features including occipital encephalocele, polydactyly and polycystic kidneys.

Qiao et al. describe a prenatal case with Cornelia de Lange syndrome (CdLS) caused by a novel heterozygous synonymous *NIPBL* variant. RT-PCR confirmed that the variant affects splicing. Their study expands the mutation spectrum of *NIPBL* and shows the importance of in-depth analysis of apparent synonymous mutations in clinical diagnoses.

He et al. show that prenatal WES can improve the genetic diagnostic yield in anomalous fetuses with normal results both at karyotype and chromosomal microarray. They performed WES in 94 fetuses and identified a diagnostic genetic variant in 37

(39%). They also obtained fetal phenotypes by post-mortem examinations for terminated pregnancies, providing genotype-phenotype correlations and more precise interpretations of the genetic results.

Prenatal risk assessment of carriers of heterozygous X-linked deletions is a great challenge due to variable phenotype expression induced by X chromosome inactivation (XCI). Zhao et al. present four pedigrees with distinct pathogenic X-linked deletions larger than 1Mb, which were associated with completely skewed XCI in all female carriers. They also performed the first prenatal XCI pattern analysis in a female fetus carrier of heterozygous *PCDH19*-deletion to make risk prediction.

Hou et al. present a girl with Wiskott-Aldrich syndrome (WAS) due to a heterozygous *WAS* variant. WAS is a rare X-linked recessive immunodeficiency disorder; affected males show symptoms while females carrying a pathogenic *WAS* variant are usually asymptomatic. In the described girl, the complete inactivation of normal X-chromosome led to the dominant symptoms. This study illustrates the importance of performing in-depth molecular assays to confirm the diagnosis of symptomatic female carriers with X-linked disorders.

## Studies exploring incomplete penetrance and variable expressivity in rare disorders, and discussing genotype-phenotype correlations

With the expanded use of WES or WGS analysis, families with rare disorders with incomplete penetrance and variable expressivity are increasingly being reported, showing the importance of performing molecular analysis to facilitate the diagnosis of syndromes with heterogeneous phenotypes.

Yim et al. report a family with an incomplete presentation of Andersen-Tawil syndrome (ATS) diagnosed through WES; the proband had all classic symptom triad, while her father who also carried the same novel *KCNJ2* variant had only two symptoms and no cardiac manifestations.

Linhares et al. describe three patients with mild presentation of Arthrogryposis, renal dysfunction, and cholestasis syndrome (ARCS) diagnosed by WES, all of them carrying the same novel *VPS33B* variant. The authors raise awareness of the existence of a mild clinical picture of ARCS with prolonged survival, and they propose that molecular analysis of *VPS33B* and *VIPAS39* should be considered in patients with normal gamma-glutamyl transferase cholestasis.

Che et al. report a three-generation pedigree including eight family members with mild thrombocytopenia due to a novel heterozygous CYCS variant. There are only 4 pedigrees with non-syndromic thrombocytopenia reported in the literature, and the disease has been characterized with mild thrombocytopenia, normal platelet size and morphology, and no increased bleeding tendency.

Wen et al. report a novel *SOD1* missense variant (p.R116S) causing Rapid Progressive Familial Amyotrophic lateral sclerosis (ALS). Their study expands the list of ALS causing variants, and suggests a possible pathogenic determinant related to R116 mutant SOD1 neurotoxicity.

Xie et al. review pathogenic *COQ4* variants in patients with Primary Coenzyme Q10 Deficiency-7 (COQ10D7), a rare mitochondrial disorder. Their study provides a fundamental reference for the sub-classification of COQ10D7 and contribute to the knowledge of the pathogenesis, clinical diagnosis and prognosis of this disease, and to the development of possible interventions.

Peng et al. delineate the molecular features, clinical presentation, and aspects of the immunological phenotype of 6 paediatric patients with EBV-infection and N-linked glycosylation defect (XMEN) disease caused by novel *MAGT1* variants. They conclude that when evaluating patients with transaminase elevation, chronic EBV infection and EBV-associated lymphoproliferative disease, the possibility of XMEN should be considered in addition to isolated liver diseases.

Feng et al. characterize the phenotypic features in five families with Branchiootorenal syndrome (BOR) and branchiooto syndrome (BOS) due to novel variants in *EYA1* and *SIX1*, respectively; all cases exhibited a degree of phenotypic variability between or within families. The authors also demonstrate that cochlear implantation for auditory rehabilitation is a feasible option in some BOR/BOS patients.

Kim et al. report the first case of SHORT syndrome-related transient neonatal diabetes mellitus with insulin resistance due to a pathogenic *PIK3P1* variant. The patient exhibited three of the five characteristics in the SHORT syndrome acronym (short stature, deep set eyes, inguinal hernias), as well as the typical facial gestalt.

## Applications of bioinformatic methods to increase medical diagnosis considering repeat expansions, structural variants, mosaic variants and uniparental disomy

The diagnostic yield of NGS is expected to increase with the development of novel bioinformatics methods that could facilitate the detection of repeat expansions, structural variants, mosaic variants, uniparental disomy and disease-causing variants in non-coding regions.

Li et al. describe an 8-year-old boy with a novel homozygous *MYF5* variant due to paternal uniparental disomy (UPD) of chromosome 12. Their results re-emphasized the clinical importance of UPD analysis in case the results are inconsistent with recessive inheritance.

Li et al. show that WES could detect partial exonic deletion mutations even involving intronic sequences by copy number variants (CNVs) and breakpoint analysis. They described the case of a 10-year-old boy diagnosed with Duchenne muscular dystrophy (DMD) due to a *de novo* deletion spanning intron 50 and exon 51. The deletion was identified by WES and confirmed by Sanger sequencing and long-read WGS.

The accurate analysis of short tandem repeats is a critical aspect to define precise prognosis and transmission risk in more than 40 rare genetic diseases, including Fragile X syndrome (FXS). However, molecular diagnosis of such genetic features is traditionally challenging and incompletely achieved. Grosso et al. demonstrate that the use of indirect sequence capture (Xdrop technology) coupled to Nanopore and Illumina sequencing is an efficient approach to characterize repeat lengths and to identify Single Nucleotide Variants (SNVs) and small insertions/deletions (indels) in *FMR1*, the gene responsible of FXS.

Li et al. show the value of combining WES with low-coverage WGS for better characterization of structural rearrangements in RDs. They described an 18-year-old man with Nagashima-type palmoplantar keratoderma (NPPK) and diffuse white matter abnormalities in the brain. He presented with a mosaic heterozygous *SERPINB7* stop-gain variant inherited from the mother and a *de novo SERPINB7* exonic deletion identified via trio-based WES explaining NPPK. Copy number variant analysis of the WES data indicated an additional downstream heterozygous gross deletion of 18q22.3-q23. Furthermore, low-coverage WGS confirmed the 18q22.3-q23 deletion and additionally detected a mosaic 18q21.33-q22.3 deletion, together explaining the proband's MRI findings.

## Conclusions and perspectives

It is evident that the applications of NGS can facilitate the diagnosis of RD patients, yet the translational gap between NGS-based genetic testing and clinical implementation remains, mainly due to a scarcity of infrastructure and trained specialized staff. Better public health policies are necessary to fill this gap and support the implementation of care pathways for people living with a RD (Castro et al., 2017; Tumiene and Graessner, 2021). On the 16 December 2021, the United Nations (UN) formally adopted the first-ever resolution on "Addressing the challenges of persons living with a rare disease and their families" (United Nations, 2021). The Resolution was adopted by consensus with the support from all 193 UN Member States of the General Assembly. The Member States are encouraged to progressively cover RD patients with safe, effective, and affordable essential medicines, diagnostics, and health technologies. The Resolution also encourages Member States to increase research support by strengthening international collaboration and the sharing of data. Consequently, new hope arises with the perspectives offered by this new UN resolution in the field of RDs.

## Author contributions

NL wrote the manuscript, and all authors critically reviewed the article and approved it.

## Funding

MUR–Ministero dell'Università e della Ricerca, PRIN 2020 (20203P8C3X).

## Acknowledgments

We acknowledge the contribution of all authors, reviewers, and editors that have contributed to the realization of the Research Topic. Additionally, with our utmost gratitude and profound respect, we want to dedicate this article collection to all patients with Rare Disorders and their families.

## Conflict of interest

Author NL was employed by the company Genuity Science.

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.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

# References


Amberger, J. S., Bocchini, C. A., Schiettecatte, F., Scott, A. F., and Hamosh, A. (2015). OMIM.org: Online Mendelian Inheritance in Man (OMIM®), an online catalog of human genes and genetic disorders. *Nucleic Acids Res.* 43, D789–D798. doi:10.1093/nar/gku1205

Austin-Tse, C. A., Jobanputra, V., Perry, D. L., Bick, D., Taft, R. J., Venner, E., et al. (2022). Best practices for the interpretation and reporting of clinical whole genome sequencing. *npj Genom. Med.* 7, 27. doi:10.1038/s41525-022-00295-z

Castro, R., Senecat, J., de Chalendar, M., Vajda, I., Dan, D., Boncz, B., et al. (2017). "Bridging the gap between health and social care for rare diseases: Key issues and innovative solutions," in *Rare diseases epidemiology: Update and overview*. Editors M. Posada de la Paz, D. Taruscio, and S. C. Groft (Cham: Springer International Publishing), 605–627. doi:10.1007/978-3-319-67144-4_32

Clark, M. M., Stark, Z., Farnaes, L., Tan, T. Y., White, S. M., Dimmock, D., et al. (2018). Meta-analysis of the diagnostic and clinical utility of genome and exome sequencing and chromosomal microarray in children with suspected genetic diseases. *NPJ Genom. Med.* 3, 16. doi:10.1038/s41525-018-0053-8

European Union (2000). Regulation (EC) No 141/2000 of the European parliament and of the council of 16 december 1999 on orphan medicinal products. Available at: http://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2000:018:0001:0005:EN:PDF [Accessed February 7, 2022].

Nguengang Wakap, S., Lambert, D. M., Olry, A., Rodwell, C., Gueydan, C., Lanneau, V., et al. (2020). Estimating cumulative point prevalence of rare diseases: Analysis of the Orphanet database. *Eur. J. Hum. Genet.* 28, 165–173. doi:10.1038/s41431-019-0508-0

Tumiene, B., and Graessner, H. (2021). Rare disease care pathways in the EU: From odysseys and labyrinths towards highways. *J. Community Genet.* 12, 231–239. doi:10.1007/s12687-021-00520-9

United Nations (2021). Resolution 76/132 adopted by the General Assembly on 16 December 2021. Addressing the challenges of persons living with a rare disease and their families. Available at: https://www.rarediseasesinternational.org/wp-content/uploads/2022/01/Final-UN-Text-UN-Resolution-on-Persons-Living-with-a-Rare-Disease-and-their-Families.pdf [Accessed February 7, 2022].

frontiers in Genetics

CASE REPORT
published: 09 February 2021
doi: 10.3389/fgene.2021.628890

# Case Report: Prenatal Whole-Exome Sequencing to Identify a Novel Heterozygous Synonymous Variant in NIPBL in a Fetus With Cornelia de Lange Syndrome

*Fengchang Qiao†, Cuiping Zhang†, Yan Wang, Gang Liu, Binbin Shao, Ping Hu* and Zhengfeng Xu**

*Department of Prenatal Diagnosis, Nanjing Maternity and Child Health Care Hospital, Women's Hospital of Nanjing Medical University, Nanjing, China*

OPEN ACCESS

**Edited by:**
*Natália Duarte Linhares, Genuity Science, Ireland*

**Reviewed by:**
*Palma Finelli, University of Milan, Italy*
*Ilaria Parenti, University Hospital Essen, Germany*

***Correspondence:**
*Zhengfeng Xu*
*zhengfeng_xu_nj@163.com*
*Ping Hu*
*njfybjyhuping@163.com*

*†These authors have contributed equally to this work*

**Specialty section:**
*This article was submitted to Genetics of Common and Rare Diseases, a section of the journal Frontiers in Genetics*

**Received:** *13 November 2020*
**Accepted:** *19 January 2021*
**Published:** *09 February 2021*

**Citation:**
*Qiao F, Zhang C, Wang Y, Liu G, Shao B, Hu P and Xu Z (2021) Case Report: Prenatal Whole-Exome Sequencing to Identify a Novel Heterozygous Synonymous Variant in NIPBL in a Fetus With Cornelia de Lange Syndrome. Front. Genet. 12:628890. doi: 10.3389/fgene.2021.628890*

Cornelia de Lange syndrome (CdLS) is a genetically heterogeneous disorder characterized by a wide spectrum of abnormalities, including craniofacial dysmorphism, upper limb anomalies, pre- and post-natal growth restrictions, hirsutism and intellectual disability. Approximately 60% of cases are caused by *NIPBL* variants. Herein we report on a prenatal case presented with bilateral upper-extremity malformations and cardiac defects. Whole-exome sequencing (WES) was performed on the fetus–parental trio and a *de novo* heterozygous synonymous variant in *NIPBL* [chr5:37020979; NM_133433.4: c.5328G>A, p. (Gln1776=)] was identified. Reverse transcriptase–polymerase chain reaction (RT–PCR) was conducted to evaluate the potential splicing effect of this variant, which confirmed that the variant caused a deletion of exon 27 (103 bp) by disrupting the splice-donor site and changed the reading frame with the insertion of at least three stop codons. Our finding not only expands the mutation spectrum of *NIPBL* gene but also establishes the crucial role of WES in searching for underlying genetic variants. In addition, our research raises the important issue that synonymous mutations may be potential pathogenic variants and should not be neglected in clinical diagnoses.

**Keywords: Cornelia de Lange syndrome, whole-exome sequencing, NIPBL, splicing mutation, prenatal diagnosis**

## INTRODUCTION

Cornelia de Lange syndrome (CdLS, MIM #122470, #300590, #610759, #614701, and #300882), a congenital developmental disorder inherited in an autosomal dominant (*NIPBL*, *SMC3*, and *RAD21*) or X-linked (*SMC1A* and *HDAC8*) manner, was initially discovered by Brachmann (1916) and was further summarized by de Lange (1933) (Jang et al., 2015; Li et al., 2020). CdLS is complicated by a wide spectrum of abnormalities, including craniofacial dysmorphism, upper limb anomalies, pre- and post-natal growth restrictions, hirsutism, and intellectual disability (Lalatta et al., 2007). Facial dysmorphisms include micrognathia, synophrys, long eyelashes, depressed nasal bridge, low-set ears, small head, small and widely spaced teeth, long philtrum and thin lips (Ireland and Burn, 1993). CdLS occurs sporadically with a prevalence of 1 in 10,000 to 1 in 30,000 live births (Kline et al., 2018); however, the actual number is far more than that as some mild cases with atypical symptoms are not clinically diagnosed.

CdLS is typically diagnosed after birth based on some abnormal phenotypes (Kinjo et al., 2019). Although several cases have been reported before birth, prenatal diagnosis is still a challenge for clinicians and geneticists, with only 23% of all cases being detected prenatally (Avagliano et al., 2017). It has been reported that maternal serum pregnancy-associated placental protein-A (PAPP-A) produced by the placenta in the first and second trimester of pregnancy is an early indicator of CdLS, but it is not specific as only a very low level can support this diagnosis (Kinjo et al., 2019). At present, prenatal diagnosis of CdLS primarily relies on sonographic findings during the second and third trimester of pregnancy, such as characteristic facial features and congenital malformations. However, some such prenatal phenotypes that are detected by ultrasound are not unique to CdLS. Undoubtedly, the complexity and variability of CdLS phenotypes increase the difficulty of prenatal diagnosis. Therefore, considerable clinical experience and professional knowledge about the dysmorphic features of CdLS are still needed for clinicians.

With the rapid development of molecular diagnostic technology, the variants of relative pathogenic genes can be confirmed by molecular testing, which is the key to diagnosis. To date, it has been acknowledged that pathogenic variants of five genes encoding the structural components (*SMC1A*, *SMC3*, and *RAD21*) or function-related factors (*NIPBL* and *HDAC8*) of the cohesin complex are associated with CdLS. In particular, the heterozygous mutations of *NIPBL* account for ∼60–70% of all CdLS cases (Teresa-Rodrigo et al., 2014; Kline et al., 2018). *NIPBL* localized on 5p13.2 facilitates cohesin loading on chromatin, whereas the cohesin complex is involved in sister chromatid cohesion and ensures accurate chromosome segregation, recombination-mediated DNA repair, and genomic stability (Baquero-Montoya et al., 2014; Teresa-Rodrigo et al., 2014). It has been reported that the variants in *NIPBL* also manifest as severe CdLS phenotypes with a high frequency of limb anomalies (Selicorni et al., 2007). Additionally, individuals with a somatic heterozygous pathogenic mutation in *NIPBL* appear to be fully penetrant (Deardorff et al., 1993). Nevertheless, for 20–30% of affected individuals, molecular alterations cannot be found in the aforementioned genes (Pie et al., 2016; Wagner et al., 2019).

Whole-exome sequencing (WES), especially for the fetus–parental trio, has proven to be a feasible and effective approach for the identification of some rare disorders that are suggestive of genetic etiology but cannot be diagnosed using conventional testing methods. In two large prospective cohort studies, Petrovski and Lord and their respective teams identified an additional 8.5 and 10% diagnostic genetic variants, respectively, through WES for cases that tested negative in karyotype and chromosomal microarray analyses (CMA), which further confirmed the feasibility and potential value of prenatal WES in detecting genetic diseases (Lord et al., 2019; Petrovski et al., 2019). Currently, WES is a first-line detection method for hereditary diseases. In our study, we detected a *de novo* heterozygous variant of *NIPBL* in a fetus with bilateral upper-extremity malformations and cardiac defects via the application of WES, which provides a theoretical basis and guidance for this family in reproductive genetic counseling and prenatal diagnosis.

# MATERIALS AND METHODS

## Ethics Statement

This research was approved by the Ethics Committee of the Nanjing Maternity and Child Health Care Hospital and adhered to the Declaration of Helsinki. Samples and information were collected from the parents after written informed consent was obtained. In addition, they agreed to the publication of the manuscript.

## Case Report

A 26-year-old pregnant Chinese woman gravida 2 para 0, who was referred to the Center of Prenatal Diagnosis at Nanjing Maternity and Child Health Care Hospital opted for genetic tests because of an ultrasonic abnormality. The woman and her husband were healthy and non-consanguineous, with no history of infection and/or exposure to teratogens. However, their family history was significant for a previous pregnancy that was terminated because of skeletal and cardiac dysplasia.

At 24 weeks of gestation, a prenatal ultrasound examination in our hospital indicated skeletal dysplasia of the bilateral upper limbs and congenital heart malformation. The bilateral humerus was short, ∼1.93 cm, and only the ulna was angled with the humerus. Additionally, the fetus had congenital cardiac developmental abnormalities, and the structures of the ventricular outflow tract were unclear because of the small gestational age. After genetic counseling by the clinic's geneticist, the couple ultimately decided to terminate the pregnancy at $27^{+3}$ weeks because of the fetal malformations and uncertainty of prognosis. However, they did not consent to a pathological autopsy. To assess the risk of recurrence, they requested genetic testing to clarify the potential cause of the disease. Because of the negative results of CMA test, we applied WES for the proband and his healthy parents (**Figure 1A**) to search for potential variants. The detailed examinations of pregnant woman are listed in **Supplementary Table 1**.

## Whole-Exome Sequencing

Genomic DNA was extracted from the proband's tissue sample after the pregnancy was terminated and from the parental peripheral blood using an Automated Nucleic Acid Extractor (RBCBioscience) following standard procedures for WES analysis. First, genomic DNA was sheared and amplified to generate a fragment library with a read length of 150 bp that was captured by the IDT The xGen Exome Research Panel v1.0 according to the manufacturer's protocols. Sequencing was subsequently carried out on the Illumina NovaSeq 6000 (Illumina, Inc.) platform, and the resultant reads were mapped against GRCh37/hg19 assembly. Single-nucleotide variants (SNVs) and small indels (<50 bp) were identified using the Genome Analysis Toolkit 3.4 (GATK). Novel mutations were filtered against the 1000 Genomes database (1000 genomes release phase 3, http://www.1000genomes.org/), dbSNP database (http://www.ncbi.nlm.nih.gov/projects/

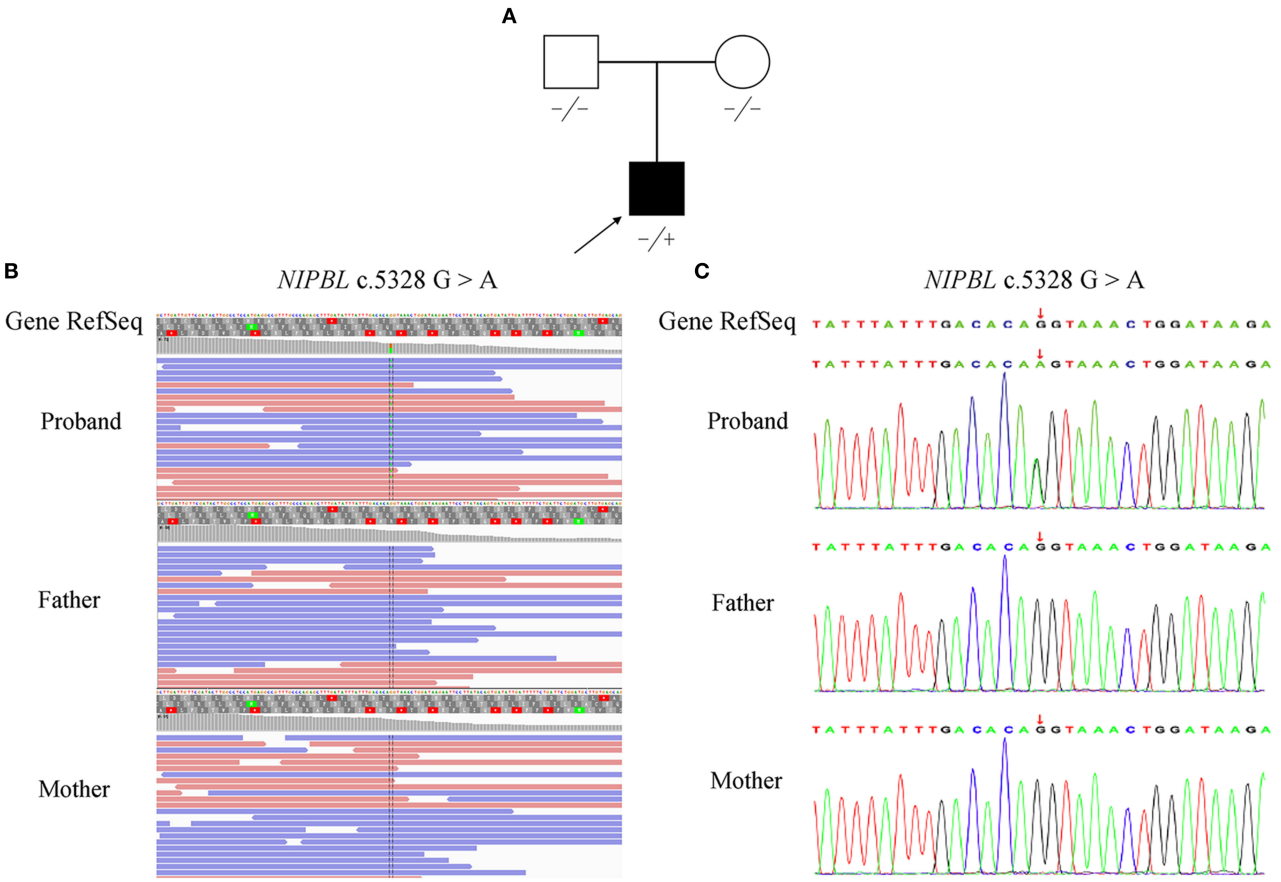



**FIGURE 1 |** Pedigree and mutation analysis of the family. **(A)** Pedigree of the family with segregation of the identified *NIPBL* mutation. The square represents the male, and circle represents the female. A filled symbol indicates the proband. **(B)** Visualization of the mutation in *NIPBL* is shown with an integrative genomics viewer. The variant (c.5328G>A) was heterozygous in the proband. **(C)** Confirmation of the mutation (c.5328G>A), which was not found in his parents, was conducted by Sanger sequencing.

SNP/snp_summary.cgi), and the genome aggregation database (gnomad.broadinstitute.org). Finally, polymerase chain reaction (PCR) was performed to amplify the affected fragment of *NIPBL* using specific primers (forward primer: 5′-TACACCCGGCCG GAGAACCTAAAA-3′ and reverse primer: 5′-ATGCATCG AGCTGAAAACCGAAAA-3′), and the purified PCR products were applied to Sanger sequencing to affirm the mutation.

## RNA Transcript Analysis by RT–PCR

MaxEntScan (Wappenschmidt et al., 2012; Afzal et al., 2019) and dbscSNV (Jian et al., 2014) were used to predict whether the variant affects splice site. To detect aberrant transcripts, total RNA was isolated from fetal skin and control cells using TRIzol™ Reagent (Invitrogen; Thermo Fisher Scientific) according to the manufacturer's protocols. Complementary DNA (cDNA) was synthesized from the total RNA using a PrimeScript™ RT reagent Kit with gDNA Eraser (TAKARA), which was further amplified to obtain PCR products using specific primers (forward primer: 5′-AGGAACACATCATGCAAAGGA-3′ and reverse primer: 5′-ACAGCAACAACCTCAGACAA-3′). PCR products were ultimately separated on a 2% agarose gel and validated by Sanger sequencing.

## RESULTS

We first examined the CMA for the proband, but no chromosomal abnormalities or submicroscopic chromosomal imbalances were detected at the whole genome level. Then, we performed sequencing and subsequent alignment of the whole exome and obtained 77.82 million reads with a read length of 150 bp and average sequencing depth of 99.59× and 98.93% of the whole exome target region covered at ≥20×. There were 32,561 single-nucleotide polymorphisms (SNPs), including 3,432 non-synonymous SNPs and 1,725 synonymous SNPs in the coding sequence, as well as 35 SNPs in the splice sites. In addition, there were 244 indels in the coding sequence.

Following the prioritized filtration strategy (Roy et al., 2018), we selected nine candidate variants, including seven missense mutations and two frameshift mutations (**Supplementary Table 2**). Notably, a plausible point mutation

in *NIPBL* [chr5:37020979; NM_133433.4: c.5328G>A, p.(Gln1776=)] may explain the features described above, considering the inheritance model of the disease and phenotypes of the gene. By bioinformatics, the variant is predicted to affect splicing according to MaxEntScan and dbscSNV (**Supplementary Table 3**). Next, we performed RNA transcript analysis using specific RT-PCR to explore the implication of this aberrant variant on the splicing of *NIPBL*. The resultant PCR products of the control RNA showed that a single amplified cDNA fragment of 300 bp was expected of a normal transcript, whereas the proband showed two bands comprising an aberrant band (197 bp) that was smaller than that of the control cells (**Figure 2A**). Sanger sequencing cDNA data analysis revealed that the aberrant transcript was affected by the variant, thus leading to complete skipping of exon 27 (103 bp) (**Figure 2B**). Although the variant was located at the last nucleotide of exon 27 and not within the canonical splice sites, the variant affects splicing generating an aberrant transcript with complete absence of exon 27 characterized by altered protein reading frame and premature stop codons (ACMG variant evidence PVS1). Using Sanger sequencing, the variant was further verified as a *de novo* heterozygous synonymous mutation in the proband, which means that it was not identified in either of the non-consanguineous parents (ACMG variant evidence PS2) (**Figures 1B,C**). Furthermore, this variant was absent from the 1000 Genomes, ExAC, or dbSNP databases (ACMG variant evidence PM2). Moreover, the probability of loss-of-function intolerance (pLI) is 1 in the ExAC Browser, which suggests that *NIPBL* is highly intolerant to heterozygous loss-of-function variants. Considered together, the variant (c.5328G>A) was classified as a pathogenic mutation according to the ACMG guidelines (Richards et al., 2015).

In addition, the remaining eight candidate variants were excluded because of insufficient evidence. *COL11A2*, *DCSH1*, *POC1A*, and *WDR60* follow the model of autosomal recessive inheritance, but the proband with heterozygosity variants was not affected. Although the variant of *RAB40AL* was assessed as likely pathogenic, its syndrome phenotype was not consistent with the ultrasonic phenotype of the fetus. Furthermore, the mutations of *IFIH1*, *WHSC1,* and *TBX4* were evaluated as variants of uncertain significance (**Supplementary Table 2**).

## DISCUSSION

Here, we report a fetus with CdLS whose diagnosis was prompted by abnormal sonographic findings of upper extremity anomalies and congenital heart defects. The cause of the disease was a previously undescribed *de novo* heterozygous synonymous variant in *NIPBL* (c.5328G>A) that was identified by WES. The variant located in the exon–intron border disrupted the normal splicing process as predicted by bioinformatics. By further validation using RT-PCR and Sanger sequencing, we detected the deletion of exon 27 (103 bp, 1,776–1,810 aa) at the mRNA level in only the proband.

*NIPBL* encodes the Nipped-B-like protein (delangin), which plays a pivotal role in loading the cohesin complex onto DNA and is crucial for sister chromatid cohesion and regulation of other proteins (Gause et al., 2008; Bot et al., 2017). To our knowledge, *NIPBL* is expressed in developing limbs and subsequently in cartilage primordia of the ulna and various hand bones, as well as in the developing heart, especially in the atrial and ventricular myocardium during the embryo stage (Ben-Asher and Lancet, 2004; Tonkin et al., 2004; Kang et al., 2018). In the case presented here, a prenatal ultrasound suggested that the bilateral humerus of the fetus was shorter than expected at standard gestational age, and the metacarpal bone could not be clearly identified. Furthermore, the fetus showed congenital cardiac developmental abnormalities. Combining the bias of *NIPBL* expression regions and the phenotypic characteristics of the fetus, we inferred that the mutation of *NIPBL* is correlated with the aberrant fetal phenotypes.

It is universally acknowledged that synonymous mutations are often neglected as silent variants, which may be because of a lack of alteration in the composition of amino acids caused by the degeneracy of the genetic code. To narrow down candidate mutations, synonymous mutations are usually eliminated first as benign variants in clinical analysis. However, with the advancement of research, mounting evidence illustrates the pathogenic mechanism of synonymous mutation by affecting the translation rate, mRNA stability, protein folding, and splicing accuracy in human diseases (Cartegni et al., 2002; Nackley et al., 2006; Sauna and Kimchi-Sarfaty, 2011). Previous studies have demonstrated that synonymous mutations could affect the splicing accuracy. According to a study by Faa et al. (2010), a synonymous mutation (c.2811G>T) in *CFTR* of an Italian patient with cystic fibrosis created a new $5'$ splice site within exon 15, which generated a transcript lacking 76 amino acid residues. The aberrant protein may affect the structure and regulation of the chloride ($Cl^-$) channel (Faa et al., 2010). Furthermore, Kim et al. (2013) also reported a novel synonymous mutation (c.1803G>A) in *SLC26A4* that resulted in complete skipping of exon 16 by affecting the splicing accuracy in two members of a Korean family suffering from hearing impairment (Kim et al., 2013). All these described cases reveal that synonymous mutations should be evaluated separately in clinical diagnoses as they usually affect splicing. As described in our case, the absence of exon 27 was caused by a synonymous mutation, which was suspected to affect the splice-donor site of *NIPBL*. To some degree, exon 27 skipping may impact the patient's phenotypes.

To our knowledge, the most common effect of splicing site changes is skipping of the next exon. In our case, this novel variant c.5328G>A occurred at the terminal nucleotide of exon 27 and generated an aberrant transcript with an exon 27 deletion by affecting the splicing accuracy. Previously, Teresa-Rodrigo et al. (2014) discovered that splice site mutations account for ∼17% of *NIPBL* variants. The variant (c.5328+1G>A) in *NIPBL* that they reported was at a different in genomic location from that reported in our case (c.5328G>A), but coincidentally generated an identical splicing effect with the skipping of exon 27. In addition, the patient seemed to have the same severe structural anomalies

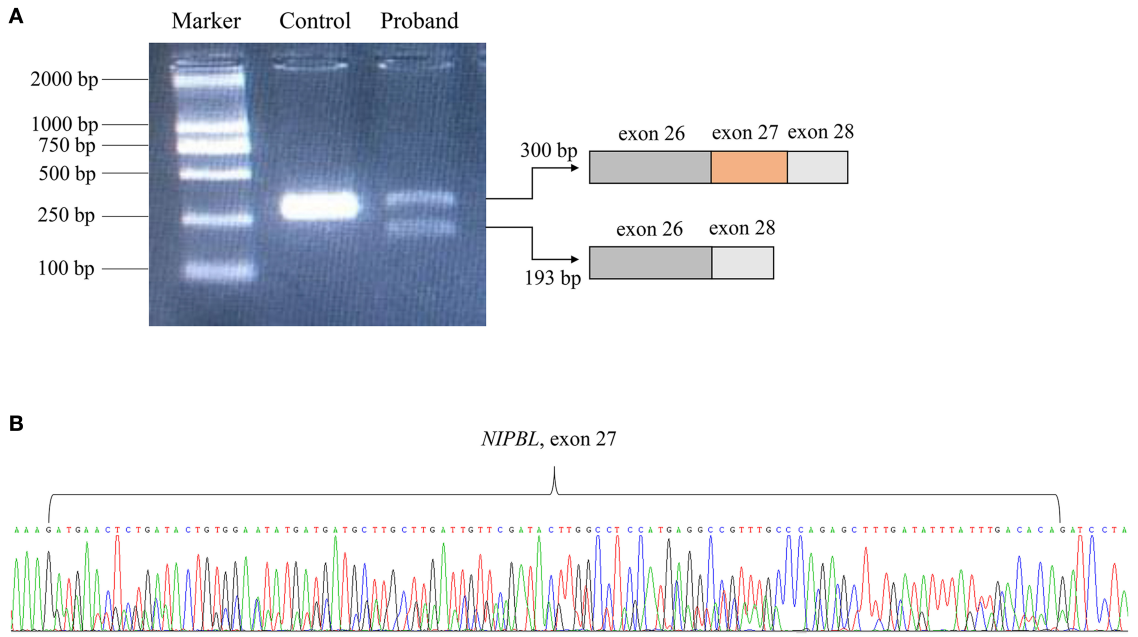



**FIGURE 2 |** RNA transcript analysis using specific RT-PCR. **(A)** Gel image showing the PCR products of *NIPBL* cDNA fragments. **(B)** The deletion of exon 27 in *NIPBL* mRNA was confirmed by Sanger sequencing.

of the limbs as in our case, such as bilateral aplasia of the ulna and distal humerus (Teresa-Rodrigo et al., 2014). Exon 27 is located in an evolutionarily conserved protein domain called HEAT repeats (H1: 1,767–1,805 aa, H2: 1,843–1,881 aa, H3: 1,945–1,984 aa, H4: 2,227–2,267 aa, H5: 2,313–2,351 aa), which is an essential component of the *NIPBL* gene product (delangin) and function in protein–protein interaction motifs (Neuwald and Hirano, 2000; Mannini et al., 2013; Li et al., 2020). On the one hand, absence of exon 27 might disrupt the integrity of the HEAT repeats. A previous study showed that the C-terminal region containing HEAT repeats and the heterochromatin protein 1 (HP1) interacting motif could recruit *NIPBL* to the sites of DNA damage, thus indicating the HEAT repeats are essential to repair DNA damage (Oka et al., 2011). In another research, Gause et al. (2008) proposed that some of the HEAT repeats are crucial for gene expression and sister chromatid cohesion. On the other hand, exon 27 skipping disrupted the reading frame and inserted premature stop codons that may have resulted in degradation of the corresponding transcript by nonsense-mediated mRNA decay. Consequently, the putative pathogenic mechanism in this case would be haploinsufficiency. Further exploration is needed to attain a detailed understanding of the effects of this variant at the transcript and protein level.

Although the mother of the proband experienced the adverse outcome of terminated pregnancy because of skeletal and cardiac dysplasia in her first pregnancy, we found a *de novo* explicit mutation in *COL1A1* [chr17:48264402; NM_000088.4: c.3505G>A, p. (Gly1169Ser)] in the first fetus by trio WES. Furthermore, the variant was already reported as pathogenic in ClinVar: SCV000695351.1 (Pathogenic*-Osteogenesis imperfecta), SCV000883622.1 (Pathogenic*-not provided), SCV000955221.1 (Pathogenic*-Osteogenesis imperfecta type I), SCV000574580.1 (Pathogenic*-Osteogenesis imperfecta type I). Therefore, the different causes of two adverse outcomes contribute to us ruling out the possibility of gonadal mosaicism. For future pregnancy, prenatal diagnosis is necessary for the fetus, especially to detect the above-mentioned mutations in these two genes.

Overall, this study presented the case of a fetus with CdLS with a *de novo* heterozygous mutation (c.5328G>A) in *NIPBL*. The synonymous mutation affects the normal splicing process and produces a frameshift that results in premature stop codons. Thus, the variant identified in our case is valuable for comprehending possible pathogenic mechanisms of *NIPBL*, which not only broadens the mutation spectrum of *NIPBL* but also contributes to supportive genetic counseling and timely management for pregnancies with a structural abnormality. Finally, we cannot ignore novel synonymous variants as silent mutations in clinical analysis; sometimes these variants elucidate the potential pathogenic mechanism of a disease.

## DATA AVAILABILITY STATEMENT

All datasets for this study are included in the article/**Supplementary Tables**.

## ETHICS STATEMENT

This research was approved by the Ethics Committee of the Nanjing Maternity and Child Health Care Hospital and adhered to the Declaration of Helsinki. Samples and information were collected from fetal parents after written informed consent was obtained and they also agreed to the publication of the manuscript.

## AUTHOR CONTRIBUTIONS

FQ and CZ performed the experiments, wrote the manuscript, and analyzed the WES data. YW and BS collected the clinical information and provided the genetics counseling. GL conducted the bioinformatics analysis. PH and ZX designed the study and revised the manuscript. All authors read and approved the final manuscript.

## FUNDING

This work was supported by the National Key R&D Program of China (No. 2018YFC1002402), National Natural Science Foundation of China (81602300, 81770236, 81801445, 81801373), Natural Science Foundation of Jiangsu Province (BK20181121), and Project on Maternal and Child Health Talents of Jiangsu Province (FRC201791).

## ACKNOWLEDGMENTS

We were thankful for the members of the group who participated and assisted us in the study.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.628890/full#supplementary-material

## REFERENCES

Afzal, R., Firasat, S., Kaul, H., Ahmed, B., Siddiqui, S. N., Zafar, S. N., et al. (2019). Mutational analysis of the CYP1B1 gene in Pakistani primary congenital glaucoma patients: Identification of four known and a novel causative variant at the 3′ splice acceptor site of intron 2. *Congenit. Anom. (Kyoto)* 59, 152–161. doi: 10.1111/cga.12312

Avagliano, L., Bulfamante, G. P., and Massa, V. (2017). Cornelia de Lange syndrome: to diagnose or not to diagnose *in utero*? *Birth Defects Res.* 109, 771–777. doi: 10.1002/bdr2.1045

Baquero-Montoya, C., Gil-Rodriguez, M. C., Hernandez-Marcos, M., Teresa-Rodrigo, M. E., Vicente-Gabas, A., Bernal, M. L., et al. (2014). Severe ipsilateral musculoskeletal involvement in a Cornelia de Lange patient with a novel NIPBL mutation. *Eur. J. Med. Genet.* 57, 503–509. doi: 10.1016/j.ejmg.2014.05.006

Ben-Asher, E., and Lancet, D. (2004). NIPBL gene responsible for Cornelia de Lange syndrome, a severe developmental disorder. *Isr. Med. Assoc. J.* 6, 571–572.

Bot, C., Pfeiffer, A., Giordano, F., Manjeera, D. E., Dantuma, N. P., and Strom, L. (2017). Independent mechanisms recruit the cohesin loader protein NIPBL to sites of DNA damage. *J. Cell Sci.* 130, 1134–1146. doi: 10.1242/jcs.197236

Brachmann, W. (1916). Ein fall von symmetrischer monodaktylie durch Ulnadefekt, mit symmetrischer flughautbildung in den ellenbeugen, sowie anderen abnormitaten (zwerghaftogkeit, halsrippen, behaarung). *Jarb. Kinder Phys. Erzie*. 84, 225–235.

Cartegni, L., Chew, S. L., and Krainer, A. R. (2002). Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nat. Rev. Genet.* 3, 285–298. doi: 10.1038/nrg775

de Lange, C. (1933). Sur un type nouveau de degenerescence (typus Amstelodamensis). *Arch. Med. Enfants* 36, 713–719.

Deardorff, M. A., Noon, S. E., and Krantz, I. D. (1993). "Cornelia de Lange syndrome," 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: U.S. National Library of Medicine).

Faa, V., Coiana, A., Incani, F., Costantino, L., Cao, A., and Rosatelli, M. C. (2010). A synonymous mutation in the CFTR gene causes aberrant splicing in an italian patient affected by a mild form of cystic fibrosis. *J. Mol. Diagn.* 12, 380–383. doi: 10.2353/jmoldx.2010.090126

Gause, M., Webber, H. A., Misulovin, Z., Haller, G., Rollins, R. A., Eissenberg, J. C., et al. (2008). Functional links between Drosophila Nipped-B and cohesin in somatic and meiotic cells. *Chromosoma* 117, 51–66. doi: 10.1007/s00412-007-0125-5

Ireland, M., and Burn, J. (1993). Cornelia de Lange syndrome–photo essay. *Clin. Dysmorphol.* 2, 151–160.

Jang, M. A., Lee, C. W., Kim, J. K., and Ki, C. S. (2015). Novel pathogenic variant (c.3178G>A) in the SMC1A gene in a family with Cornelia de Lange syndrome identified by exome sequencing. *Ann. Lab Med.* 35, 639–642. doi: 10.3343/alm.2015.35.6.639

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

Kang, M. J., Ahn, S. M., and Hwang, I. T. (2018). A novel frameshift mutation (c.5387_5388insTT) in NIPBL in Cornelia de Lange syndrome with severe phenotype. *Ann. Clin. Lab Sci.* 48, 106–109.

Kim, Y., Kim, H. R., Kim, J., Shin, J. W., Park, H. J., Choi, J. Y., et al. (2013). A novel synonymous mutation causing complete skipping of exon 16 in the SLC26A4 gene in a Korean family with hearing loss. *Biochem. Biophys. Res. Commun.* 430, 1147–1150. doi: 10.1016/j.bbrc.2012.12.022

Kinjo, T., Mekaru, K., Nakada, M., Nitta, H., Masamoto, H., and Aoki, Y. (2019). A Case of Cornelia de Lange syndrome: difficulty in prenatal diagnosis. *Case Rep. Obstet. Gynecol.* 2019:4530491. doi: 10.1155/2019/4530491

Kline, A. D., Moss, J. F., Selicorni, A., Bisgaard, A. M., Deardorff, M. A., Gillett, P. M., et al. (2018). Diagnosis and management of Cornelia de Lange syndrome: first international consensus statement. *Nat. Rev. Genet.* 19, 649–666. doi: 10.1038/s41576-018-0031-0

Lalatta, F., Russo, S., Gentilin, B., Spaccini, L., Boschetto, C., Cavalleri, F., et al. (2007). Prenatal/neonatal pathology in two cases of Cornelia de Lange syndrome harboring novel mutations of NIPBL. *Genet. Med.* 9, 188–194. doi: 10.1097/gim.0b013e31803183dd

Li, S., Miao, H., Yang, H., Wang, L., Gong, F., Chen, S., et al. (2020). A report of 2 cases of Cornelia de Lange syndrome (CdLS) and an analysis of clinical and genetic characteristics in a Chinese CdLS cohort. *Mol. Genet. Genomic Med.* 8:e1066. doi: 10.1002/mgg3.1066

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

Mannini, L., Cucco, F., Quarantotti, V., Krantz, I. D., and Musio, A. (2013). Mutation spectrum and genotype-phenotype correlation in Cornelia de Lange syndrome. *Hum. Mutat.* 34, 1589–1596. doi: 10.1002/humu.22430

Nackley, A. G., Shabalina, S. A., Tchivileva, I. E., Satterfield, K., Korchynskyi, O., Makarov, S. S., et al. (2006). Human catechol-O-methyltransferase haplotypes modulate protein expression by altering mRNA secondary structure. *Science* 314, 1930–1933. doi: 10.1126/science.1131262

Neuwald, A. F., and Hirano, T. (2000). HEAT repeats associated with condensins, cohesins, and other complexes involved in chromosome-related functions. *Genome Res.* 10, 1445–1452. doi: 10.1101/gr.147400

Oka, Y., Suzuki, K., Yamauchi, M., Mitsutake, N., and Yamashita, S. (2011). Recruitment of the cohesin loading factor NIPBL to DNA double-strand breaks depends on MDC1, RNF168 and HP1gamma in human cells. *Biochem. Biophys. Res. Commun.* 411, 762–767. doi: 10.1016/j.bbrc.2011.07.021

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

Pie, J., Puisac, B., Hernandez-Marcos, M., Teresa-Rodrigo, M. E., Gil-Rodriguez, M., Baquero-Montoya, C., et al. (2016). Special cases in Cornelia de Lange syndrome: the Spanish experience. *Am. J. Med. Genet. C Semin. Med. Genet.* 172, 198–205. doi: 10.1002/ajmg.c.31501

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

Roy, S., Coldren, C., Karunamurthy, A., Kip, N. S., Klee, E. W., Lincoln, S. E., et al. (2018). Standards and guidelines for validating next-generation sequencing bioinformatics pipelines: a joint recommendation of the Association for Molecular Pathology and the College of American Pathologists. *J. Mol. Diagn.* 20, 4–27. doi: 10.1016/j.jmoldx.2017.11.003

Sauna, Z. E., and Kimchi-Sarfaty, C. (2011). Understanding the contribution of synonymous mutations to human disease. *Nat. Rev. Genet.* 12, 683–691. doi: 10.1038/nrg3051

Selicorni, A., Russo, S., Gervasini, C., Castronovo, P., Milani, D., Cavalleri, F., et al. (2007). Clinical score of 62 Italian patients with Cornelia de Lange syndrome and correlations with the presence and type of NIPBL mutation. *Clin. Genet.* 72, 98–108. doi: 10.1111/j.1399-0004.2007.00832.x

Teresa-Rodrigo, M. E., Eckhold, J., Puisac, B., Dalski, A., Gil-Rodriguez, M. C., Braunholz, D., et al. (2014). Functional characterization of NIPBL physiological splice variants and eight splicing mutations in patients with Cornelia de Lange syndrome. *Int. J. Mol. Sci.* 15, 10350–10364. doi: 10.3390/ijms150610350

Tonkin, E. T., Wang, T. J., Lisgo, S., Bamshad, M. J., and Strachan, T. (2004). NIPBL, encoding a homolog of fungal Scc2-type sister chromatid cohesion proteins and fly Nipped-B, is mutated in Cornelia de Lange syndrome. *Nat. Genet.* 36, 636–641. doi: 10.1038/ng1363

Wagner, V. F., Hillman, P. R., Britt, A. D., Ray, J. W., and Farach, L. S. (2019). A *de novo* HDAC2 variant in a patient with features consistent with Cornelia de Lange syndrome phenotype. *Am. J. Med. Genet. A.* 179, 852–856. doi: 10.1002/ajmg.a.61101

Wappenschmidt, B., Becker, A. A., Hauke, J., Weber, U., Engert, S., Kohler, J., et al. (2012). Analysis of 30 putative BRCA1 splicing mutations in hereditary breast and ovarian cancer families identifies exonic splice site mutations that escape *in silico* prediction. *PLoS ONE.* 7:e50800. doi: 10.1371/journal.pone.0050800

**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.

*Copyright © 2021 Qiao, Zhang, Wang, Liu, Shao, Hu and Xu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*

frontiers in Genetics

ORIGINAL RESEARCH
published: 21 June 2021
doi: 10.3389/fgene.2021.668094

# A 14-Year Italian Experience in DM2 Genetic Testing: Frequency and Distribution of Normal and Premutated *CNBP* Alleles

*Annalisa Botta[1]*, Virginia Veronica Visconti[1], Luana Fontana[1], Paola Bisceglia[2,3], Mario Bengala[2], Roberto Massa[4], Ilaria Bagni[2], Rosanna Cardani[5], Federica Sangiuolo[1,2], Giovanni Meola[6,7], Giovanni Antonini[8], Antonio Petrucci[9], Elena Pegoraro[10], Maria Rosaria D'Apice[2] and Giuseppe Novelli[1,2]*

[1] Medical Genetics Section, Department of Biomedicine and Prevention, University of Rome Tor Vergata, Rome, Italy, [2] Laboratory of Medical Genetics, Tor Vergata Hospital, Rome, Italy, [3] Research Laboratory, Complex Structure of Geriatrics, Fondazione IRCCS Casa Sollievo della Sofferenza, San Giovanni Rotondo, Italy, [4] Neuromuscular Disease Unit, Department of Systems Medicine, University of Rome Tor Vergata, Rome, Italy, [5] BioCor Biobank, UOC SMEL-1 of Clinical Pathology, IRCCS-Policlinico San Donato, Milan, Italy, [6] Department of Biomedical Sciences for Health, University of Milan, Milan, Italy, [7] Department of Neurorehabilitation Sciences, Casa di Cura del Policlinico, Milan, Italy, [8] Neuromuscular and Rare Disease Center, Department of Neuroscience, Mental Health and Sensory Organs (NESMOS), Sant'Andrea Hospital, Sapienza University of Rome, Rome, Italy, [9] Center for Neuromuscular and Neurological Rare Diseases, S. Camillo Forlanini Hospital, Rome, Italy, [10] Department of Neuroscience, University of Padua, Padua, Italy

OPEN ACCESS

**Edited by:**
*Alfredo Brusco, University of Turin, Italy*

**Reviewed by:**
*Benedetta Izzi, Istituto Neurologico Mediterraneo Neuromed (IRCCS), Italy*
*Maria Do Carmo Costa, University of Michigan, United States*

***Correspondence:**
*Annalisa Botta*
*botta@med.uniroma2.it*

**Specialty section:**
*This article was submitted to Genetics of Common and Rare Diseases, a section of the journal Frontiers in Genetics*

**Received:** *15 February 2021*
**Accepted:** *05 May 2021*
**Published:** *21 June 2021*

**Citation:**
*Botta A, Visconti VV, Fontana L, Bisceglia P, Bengala M, Massa R, Bagni I, Cardani R, Sangiuolo F, Meola G, Antonini G, Petrucci A, Pegoraro E, D'Apice MR and Novelli G (2021) A 14-Year Italian Experience in DM2 Genetic Testing: Frequency and Distribution of Normal and Premutated CNBP Alleles. Front. Genet. 12:668094. doi: 10.3389/fgene.2021.668094*

Myotonic dystrophy type 2 (DM2) is a multisystemic disorder caused by a $(CCTG)_n$ in intron 1 of the *CNBP* gene. The CCTG repeat tract is part of a complex $(TG)_v(TCTG)_w(CCTG)_x(NCTG)_y(CCTG)_z$ motif generally interrupted in *CNBP* healthy range alleles. Here we report our 14-year experience of DM2 postnatal genetic testing in a total of 570 individuals. The DM2 locus has been analyzed by a combination of SR-PCR, TP-PCR, LR-PCR, and Sanger sequencing of *CNBP* alleles. DM2 molecular diagnosis has been confirmed in 187/570 samples analyzed (32.8%) and is mainly associated with the presence of myotonia in patients. This set of *CNBP* alleles showed unimodal distribution with 25 different alleles ranging from 108 to 168 bp, in accordance with previous studies on European populations. The most frequent *CNBP* alleles consisted of 138, 134, 140, and 136 bps with an overall locus heterozygosity of 90%. Sequencing of 103 unexpanded *CNBP* alleles in DM2-positive patients revealed that $(CCTG)_5(NCTG)_3(CCTG)_7$ and $(CCTG)_6(NCTG)_3(CCTG)_7$ are the most common interruption motifs. We also characterized five *CNBP* premutated alleles with $(CCTG)_n$ repetitions from $n = 36$ to $n = 53$. However, the molecular and clinical consequences in our cohort of samples are not unequivocal. Data that emerged from this study are representative of the Italian population and are useful tools for National and European centers offering DM2 genetic testing and counseling.

**Keywords: myotonic dystrophy type 2, *CNBP* premutations, alleles distribution, DM2 genetic testing, penetrance**

## INTRODUCTION

Myotonic dystrophy type 2 (DM2; MIM#602668) is an autosomal dominant multisystemic disorder caused by a $(CCTG)_n$ repeat expansion in intron 1 of *CNBP* gene (previously *ZNF9*) on chromosome 3q21.3 (Liquori et al., 2001). DM2 is characterized by progressive proximal muscle weakness, myotonia, myalgia, calf hypertrophy, and multiorgan involvement with cataract, cardiac conduction defects, and endocrine disorders (Meola and Cardani, 2015; Montagnese et al., 2017). The clinical phenotype typically arises in the adulthood, and to date, no congenital form of the disease has been reported (Meola et al., 2017). The true prevalence of DM2 is still uncertain since DM2 is almost under diagnosed; nevertheless, clinical experience indicates that DM2 is roughly fivefold less common than myotonic dystrophy type 1 (DM1; MIM# # 160900), at least in the Mediterranean area. DM2 mutations have been identified mainly in North Eastern Europe, where the DM2 expansion only occurs on a specific chromosomal haplotype, suggesting the occurrence of a predisposing mutation in a common ancestral founder (Bachinski et al., 2003; Liquori et al., 2003; Coenen et al., 2011; Meola and Cardani, 2017). The CCTG repeat tract is part of a complex $(TG)_v(TCTG)_w(CCTG)_x$ motif, which is generally interrupted in healthy range alleles by one or more GCTG, TCTG, or ACTG (NCTG) motifs resulting in more stability (Radvanszky et al., 2013). Consequently, the combined repeat tract in those healthy range alleles can be described by five distinct repetitive motifs $(TG)_v(TCTG)_w(CCTG)_x(NCTG)_y(CCTG)_z$ (Bachinski et al., 2009). Normal alleles have <25 copies of the CCTG repeat tract, whereas expanded alleles have as many as 75–11,000 copies (Thornton, 2014). Alleles between 27 and 74 CCTG units (called gray area) are very rare, and their clinical significance is still unclear and subject of discussion in the scientific community. In DM2 patients, the interval of CCTG repeat units is extremely wide, ranging from 75 to over 11,000 units with a mean of 5,000 (Turner and Hilton-Jones, 2010). Unstable expanded alleles have an uninterrupted CCTG tract which favors unusual secondary structure more susceptible to strand slippage and unequal crossing over during cell division (Bachinski et al., 2009; Guo and Lam, 2016). The CCTG repeat size increases with age with a high level of somatic mosaicism; however, different from DM1, no significant correlation has been observed between CCTG repeat size and age of onset or other measures of disease severity (e.g., muscle weakness and age of cataract extraction) (Day et al., 2003). More interestingly, on transmission to the next generation, *CNBP* repeat length sometimes tends to decrease dramatically, without significant differences determined by the parental origin of the mutation (Day et al., 2003). The pathogenic mechanism of DM2 is an RNA gain of function toxicity leading to the formation of CCUG-containing ribonuclear foci and spliceopathy mediated mainly by the sequestration of MBNL1 protein (Cardani et al., 2009; Meola and Cardani, 2015). The molecular diagnostic protocol of DM2 is a challenging approach because of the CCTG expansion and the high level of somatic mosaicism. The initial step is the short-range PCR (SR-PCR), which allows excluding the DM2 diagnosis in the case of two alleles in the normal range. If only one allele is visible, CCTG expansion can be detected either with long-range PCR (LR-PCR) or tetraplet-primed PCR (TP-PCR) leading to about 99% of detection rate (Kamsteeg et al., 2012). The interpretation of the molecular data can be complicated by the polymorphic variability of *CNBP* normal alleles and by the length of the CCTG expansion, which can be only determined with classical Southern blot on digested genomic DNA. The molecular characterization and the distribution of the complex repeated motif and allele frequency of *CNBP* alleles in the normal population has been already reported only in German, Slovak, and American populations (Bachinski et al., 2009; Radvanszky et al., 2013; Mahyera et al., 2018). In the present work, we report our 14-year experience (from 2007 to 2020) of DM2 postnatal genetic testing in a total of 570 individuals from the Italian population. Main clinical indications in DM2 positive genetic testing, allele length, frequency, and distribution of *CNBP* normal alleles and detailed composition of the complex $(TG)_v(TCTG)_w(CCTG)_x(NCTG)_y(CCTG)_z$ are given. Furthermore, we also described at the clinical and molecular levels five individuals carrying *CNBP* unstable premutated alleles with the intent to provide useful information in DM2 genetic counseling.

## MATERIALS AND METHODS

### Individuals Included in This Study

The enrollment of the subjects described in this study was approved by the Institutional Review Board of Policlinico Tor Vergata (document no. 232/19). All individuals were referred to our center for DM2 genetic testing, and informed consents have been obtained from each participant. All individuals were of Caucasian origin from Italy. Muscle biopsies used for this study have been collected in Policlinico San Donato after receiving written informed consent from the patients (study authorized by the Institutional Ethics Committee ASL MI2-Melegnano via VIII Giugno, Milan).

### *CNBP* Fragment Analysis Data

DNA was extracted from peripheral blood using the EZ1 Advanced XL Robotic workstation for automated purification of nucleic acids (QIAGEN, Germany). The molecular characterization of the DM2 mutation has been carried out using a combination of short-range PCR (SR-PCR), tetraplet-primed PCR (TP-PCR) and long-range PCR (LR-PCR), as described (Botta et al., 2006; Kamsteeg et al., 2012). SR-PCR enabled us to assess the length of the complex motif, and up to 40 CCTG repeats can be amplified using this method. DNA region was amplified with the following PCR pair primers: CL3N58D-F (5′-GGCCTTATAACCATGCAAA-3′) and CL3N58D-R (5′-CCTAGGGGACAAAGTGAG-3′). Amplification was carried out in 30 μl volume, containing: 50 ng genomic DNA, PCR reaction buffer (5×), 25 mM $MgCl_2$, 1.25 mM dNTPs, 70 pmol of each primer, and 1 U of TaKaRa Taq polymerase (Takara Bio). The PCR cycle program consists of the following amplification conditions: 94°C for 3 min, 30 cycles of denaturation at 94°C for 1 min, annealing at 58°C for 40 s, and extension at 72°C

for 40 s. A final elongation was carried out at 72°C for 5 min. Finally, capillary electrophoresis was carried out in order to determine the length of the *CNBP* unexpanded alleles in DM2-negative individuals. The statistical analysis of the frequency and distribution of the different 756 *CNBP* normal allele lengths characterized was done with MS Excel (Office Suite 2016).

## Sequencing of *CNBP* Normal and Premutated Alleles

Sanger sequencing was performed to characterize *CNBP* repeat motif compositions of non-DM2 alleles. To simplify the analysis, PCR products were amplified from affected offspring who inherited the normal haplotype of interest, as only a single allele within the normal range would be amplified. SR-PCR products were purified and sequencing directly using BigDye Terminator v1.1 Cycle Sequencing Kit and visualized by capillary electrophoresis on an Applied Biosystems 3130xl Genetic Analyzer. Sequencing conditions were the following: 96°C for 1 min, 25 cycles of denaturation at 96°C for 10 s, annealing at 50°C for 5 s and extension at 60°C for 4 min. Data were analyzed with the software Sequencing Analysis v5.2 (Applied Biosystem).

## Muscle Histopathology and RNA FISH/MBNL1 Immunofluorescence Analyses

Human muscle biopsies from biceps brachii muscle were taken under sterile conditions. Muscle samples were trimmed of blood vessels, fat, and connective tissues and fresh-frozen in isopentane cooled in liquid nitrogen. Histopathological analysis was performed on serial sections (8 μm) processed for routine histological or histochemical stainings. A standard myofibrillar ATPase staining protocol was used after preincubation at pH 4.3, 4.6, and 10.4. The most typical alterations, such as nuclear clump fibers (i.e., aggregates of myonuclei with a thin rim of cytoplasm), nuclear centralization, and fiber size variability, with type II atrophy, were evaluated on serial muscle sections. Fluorescence *in situ* hybridization was performed on muscle frozen sections using a $(CAGG)_5$ probe as previously reported by Cardani et al. (2009) to verify the presence of ribonuclear inclusions and their colocalization with MBNL1 protein.

## Study of Alternative Splicing

Frozen muscle samples were practiced for the extraction of total RNA using TRIzol reagent (Gibco BRL, Gaithersburg, MD, United States), and 1 μg of RNA was reverse transcribed according to the cDNA protocol of the High Capacity cDNA Archive kit (Applied Biosystems, Foster City, CA, United States). Splicing pattern profile of the *IR*, *CLCN1*, and *MBNL1* genes was carried out as described (Nakamori et al., 2013). Total PCR products, obtained within the linear range of amplification, were electrophoresed on 2.5% agarose gel. Quantitative analysis of the amplified products was performed using SYBR Green II stained gels (Perkin-Elmer Life Science, Boston, MA, United States) scanned on a fluorimager 595 (Amersham Biosciences, Buckinghamshire, United Kingdom). The intensity of each band and the fraction of abnormally (or pathologically) spliced (AS) isoforms (AS-isoforms/total) were quantified by densitometry using ImageQuant software. Control of the RT-PCR reaction was based on the expression level of the glucose phosphate isomerase housekeeping gene (*GPI*) and all amplifications have been carried out in triplicate using independent cDNA samples.

# RESULTS

## Sociodemographic Characteristics of DM2 Patients and Major Clinical Indications for DM2 Genetic Testing

The Medical Genetics Unit at Tor Vergata Hospital is one of the main reference diagnostic labs for myotonic dystrophies in Italy. From 2007 to 2020, a total of 570 DM2 molecular analyses have been performed (286 females and 284 males) (**Table 1**). The main clinical indications for DM2 genetic testing were myotonia (28%), familiarity for the disease (24%), asthenia (20%), hyperCKaemia (18%), muscle weakness (17%), myalgia (15%), and cataract (6%). After DM2 genetic testing, only 187 individuals (33%) showed the presence of a $(CCTG)_n$ expansion in the *CNBP* gene. Among 187 DM2 confirmed patients, 96 were males (51%) and 91 were females (49%), with a mean disease age onset of $49 \pm 16.9$ years without a significant difference between sexes ($p = 0.2$) (**Figure 1**). The low percentage of DM2-positive test is in accordance with a data reported by Mahyera et al. (2018) and demonstrates that most clinicians still do not easily recognize the cardinal features of the DM2 disease. A positive family history was referred by 70/187 DM2-confirmed patients (37%); the familiarity was considered "positive" if patients reported of family members diagnosed with DM2. The confirmation of the DM2 mutation is mainly associated with the presence of myotonia (35%), hyperCKaemia (19%), asthenia (18%), muscle weakness (15%), and cataract (11%) in our cohort. On the contrary, cardiac, and endocrine dysfunctions are not specifically associated with

**TABLE 1 |** DM2 genetic tests requested to our lab in the period between 2007 and 2020.

| Year | Cohort (*n* = 570) | Males (*n* = 284) | Females (*n* = 286) | DM2 positive (*n* = 187) |
|---|---|---|---|---|
| 2020 | 30 | 14 | 16 | 7 |
| 2019 | 52 | 27 | 25 | 22 |
| 2018 | 56 | 22 | 34 | 18 |
| 2017 | 57 | 34 | 23 | 11 |
| 2016 | 63 | 27 | 36 | 23 |
| 2015 | 49 | 19 | 30 | 14 |
| 2014 | 40 | 20 | 20 | 14 |
| 2013 | 20 | 7 | 13 | 8 |
| 2012 | 39 | 21 | 18 | 17 |
| 2011 | 42 | 23 | 19 | 13 |
| 2010 | 35 | 19 | 16 | 12 |
| 2009 | 34 | 21 | 13 | 11 |
| 2008 | 40 | 22 | 18 | 14 |
| 2007 | 13 | 8 | 5 | 3 |

the DM2 phenotype being present in less than 5% of DM2-positive tests (**Figure 2**). However, it is important to point out that, given the heterogeneity of submitters from all over the country, some clinical information may be missed or incomplete. In DM2 families where the transmission of the expansion could be determined ($n = 62$), no differences have been found between the parental origin of the DM2 mutation.

## Frequency and Distribution of *CNBP* Healthy Range Alleles

To define the polymorphic spectrum of the $(TG)_v(TCTG)_w(CCTG)_x(NCTG)_y(CCTG)_z$ repeated motif in the Italian population, we analyzed the distribution of *CNBP* alleles in non-DM2 chromosomes. Fragment analysis by capillary electrophoresis was used to calculate the length of unexpanded *CNBP* normal alleles in 378 DM2-negative Italian individuals. The combined repeat tract length was plotted against the allele frequency. In this cohort of samples, a unimodal distribution of 25 different *CNBP* alleles length is shown, ranging from 108 to 168 bps (**Figure 3**). The most frequent alleles in the Italian population consisted of 138 (21%), 134 (13%), 140 (11%), and 136 (10%), and these data are in accordance with the data reported in the Slovak and German populations (Radvanszky et al., 2013; Mahyera et al., 2018). The combined use of SR-PCR and LR-PCR in DM2-negative probands revealed a homozygosity frequency for the *CNBP* locus of 90%, according to published data (Liquori et al., 2001; Mahyera et al., 2018).

## Structure and Composition of the $(TG)_v(TCTG)_W(CCTG)_X(NCTG)_y(CCTG)_z$ Repeated Motif in *CNBP* Normal Alleles

To determine the composition of the complex repeated tract in the *CNBP* gene, we sequenced *CNBP* normal alleles in 103 non-DM2-chromosomes: 87 derived from DM2-positive patients with full expansions and 16 from DM2-negative probands. Five additional DNA samples belonging to patients with mild muscular phenotype requesting DM2 genetic testing have also been analyzed. At least one sample in each allele length class has been sequenced; for the most frequent alleles, multiple samples have been characterized (138 bp, $n = 20$; 134 bp, $n = 14$; 140 bp, $n = 9$; and 136 bp, $n = 14$). The composition of the *CNBP* healthy range alleles is summarized in **Figure 4**. Results allow to distinguish four allelic classes: (1) short uninterrupted alleles with $(CCTG)_{10-12}$ (2) short interrupted alleles with $(CCTG)_{9-15}$ containing three interruption motifs, (3) long interrupted alleles with $(CCTG)_{13-17}$ containing five interruption motifs, and 4) long uninterrupted alleles with $(CCTG)_{22-55}$ including five DM2 premutations. Short uninterrupted alleles with $(CCTG)_{10-12}$ are very rare alleles representing only 2% of normal *CNBP* chromosomes in our population, in accordance with published data (Radvanszky et al., 2013). *CNBP* interrupted alleles were divided into long and short classes depending on the length of the interruption located within the $(CCTG)_n$ array: short alleles with $(CCTG)_{9-15}$ interrupted by three tetraplets $(NCTG)_3$: T/GCTG-CCTG-TCTG and longer alleles containing $(CCTG)_{13-15}$ interrupted by domains containing five tetraplets $(NCTG)_5$: T/GCTG-CCTG-TCTG-CCTG-TCTG repetitions. Interestingly, the most frequent *CNBP* alleles have a $(CCTG)_x(NCTG)_3(CCTG)_z$ structure (92.9%), whereas only 7.1% contains $(CCTG)_x(NCTG)_5(CCTG)_z$ motifs. The longest 168-bp allele contains no interruptions of the (CCTG) tract ($y = 0$) with a $(TG)_{18}(TCTG)_{10}(CCTG)_{23}$ structure and is identical to the longest *CNBP* allele reported in German non-DM2-chromosomes. The shortest 108-bp allele has been also sequenced resulting with the following repeated tract $(TG)_{16}(TCTG)_7(CCTG)_5(NCTG)_3(CCTG)_4$. Analysis of the repeated motifs within the interruption tract shows that the most common combinations are represented by $(CCTG)_5(NCTG)_3(CCTG)_7$ and $(CCTG)_6(NCTG)_3(CCTG)_7$, with a frequency of 62.1% and 20.4%, respectively. Furthermore, by analyzing the whole repeated stretch $(TG)_v(TCTG)_W(CCTG)_X(NCTG)_y(CCTG)_z$, we noticed that both $(TG)_{21}(TCTG)_9(CCTG)_5(NCTG)_3(CCTG)_7$ and $(TG)_{17}(TCTG)_9(CCTG)_5(NCTG)_3(CCTG)_7$ motif combinations (138 p and 130 bp alleles, respectively) were present in 8.7% of the sequenced alleles. In contrast to the length of the whole $(TG)_v(TCTG)_w(CCTG)_x(NCTG)_y(CCTG)_z$ motif, the $(CCTG)_x$ tract alone was found to be less polymorphic, with 16 different alleles identified in our cohort of samples. On the other hand, the $(TG)_v$ and the $(TCTG)_w$ tracts upstream of the CCTG array are highly polymorphic within the range of v and w tract, showing hypervariability of the $(TG)_v$ and $(TCTG)_w$ tracts with $v = 14–26$ and $w = 7–11$. Given the extreme variability in the sizes of the (TG) and (TCTG) different subunits, the precise length of the pathogenic CCTG unit within this repeat can only be determined by DNA sequencing when DM2 genetic testing is performed. Interestingly, we identified the presence of an uninterrupted $(CCTG)_n$ tract below 26 repetitions, which is considered the upper threshold for stable *CNBP* alleles, in 4% of non-DM2 chromosomes Overall, we can consider our data representative of the genetic variability associated with the DM2 locus in the Italian population. The general structure of the *CNBP* repeated motif in non-DM2 chromosomes is the following: $(TG)_{14-26}(TCTG)_{7-11}(CCTG)_{5-9}(NCTG)_{3-5}(CCTG)_{4-8}$.

## *CNBP* Premutated Alleles Are Unstable and Have Uncertain Functional and Clinical Consequences

During our diagnostic procedure, we characterized also five DM2-negative patients carrying *CNBP* premutated alleles from individuals referred to our centers for DM2 genetic testing (**Figure 5A**). Patient A1 is a 35-year-old male who came to our attention because he suffered from myalgia since the age of 16, and no other clinical signs have been reported. EMG did not reveal myotonic discharges with either myopathic or neurogenic changes in vastus lateral muscle. The CK level was mildly elevated with values ranging between 300 and 550 U/L. SR-PCR and Sanger sequencing showed that he is a carrier of a $(CCTG)_{48}$ *CNBP* uninterrupted allele. The asymptomatic sister (A2), 46 years of age, carries a $(CCTG)_{51-53}$ *CNBP* uninterrupted allele, and both premutations were inherited from the 66-year-old asymptomatic mother (A3) harboring

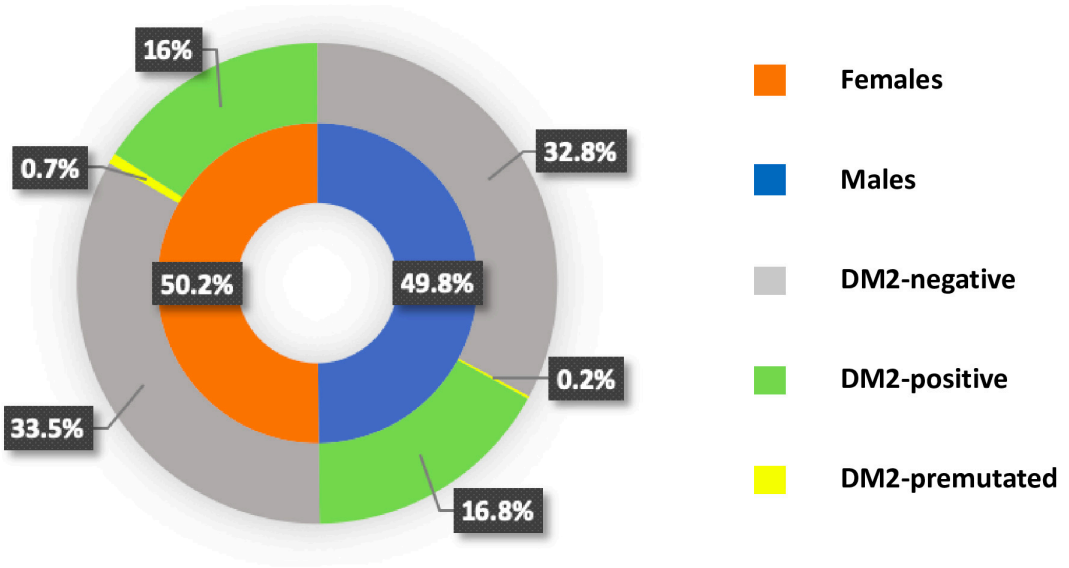



**FIGURE 1 |** Pie chart showing the frequency of myotonic dystrophy type 2 (DM2)-positive, DM2-negative, and DM2-premutated with sex distribution in each category.

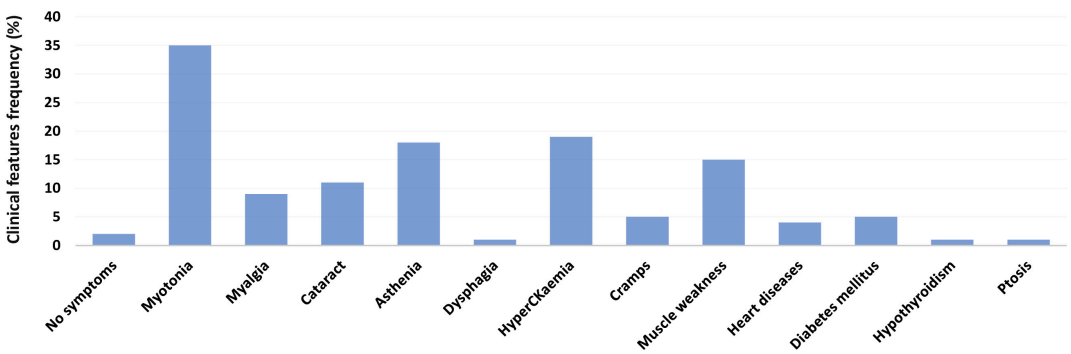



**FIGURE 2 |** Main clinical features of genetically confirmed DM2 patients ($n$ = 187).

$(CCTG)_{53-55}$ repetitions. Patient B is a 51-year-old woman, carrying a $(CCTG)_{36}$ *CNBP* uninterrupted allele, whose clinical signs reflect progressive proximal leg weakness and muscle pain, iperCKemia (>200 U/L), no clinical or EMG myotonia, insulin resistance, and hypovitaminosis D. Her father died of respiratory insufficiency at 69 years old and was reported to be affected by a not defined muscular dystrophy. Finally, patient C is a 72-year-old woman, carrying a $(CCTG)_{37}$ *CNBP* uninterrupted allele, who has been suffering from myalgia since the age of 9. She has cardiomyopathy and is a carrier of PM. In her family history, premature deaths are also reported. In addition, she showed adipose involution of the great dentate, no EMG myotonia, hypothyroidism, obstructive sleep apnea, normal CK level, and a slight increase in lactates. Unfortunately, biceps brachii muscle biopsies were available for histological and molecular analyses only for patients A1 and B. Muscle biopsy from patient A1 showed only a slight fiber size variability both of types I and II fibers (**Figure 5B**). Analysis of muscle tissue from patient B revealed an increase in fiber size variability, nuclear clumps, central nuclei, and types I and II fiber atrophy (**Figure 5B**). RNA FISH and MBNL1 immunofluorescence analyses showed no nuclear accumulation of *CNBP* mutant RNA of MBNL1 protein in myofibers from both patients (**Figure 5C**). Similar to healthy individuals, no alterations in alternative splicing of muscle blind-like 1 (*MBNL1*), insulin receptor (*IR*), and chloride voltage-gated channel 1 (*CLCN1*) genes have been evidenced by RT-PCR analysis performed on RNA extracted from muscle biopsies (**Figure 5D**).

## DISCUSSION

In this study, we report our 14-year experience in DM2 genetic testing as one of the main Italian reference centers for the molecular diagnosis of myotonic dystrophies. The prevalence of DM2 likely varies by population: DM2 is rare in the Japanese

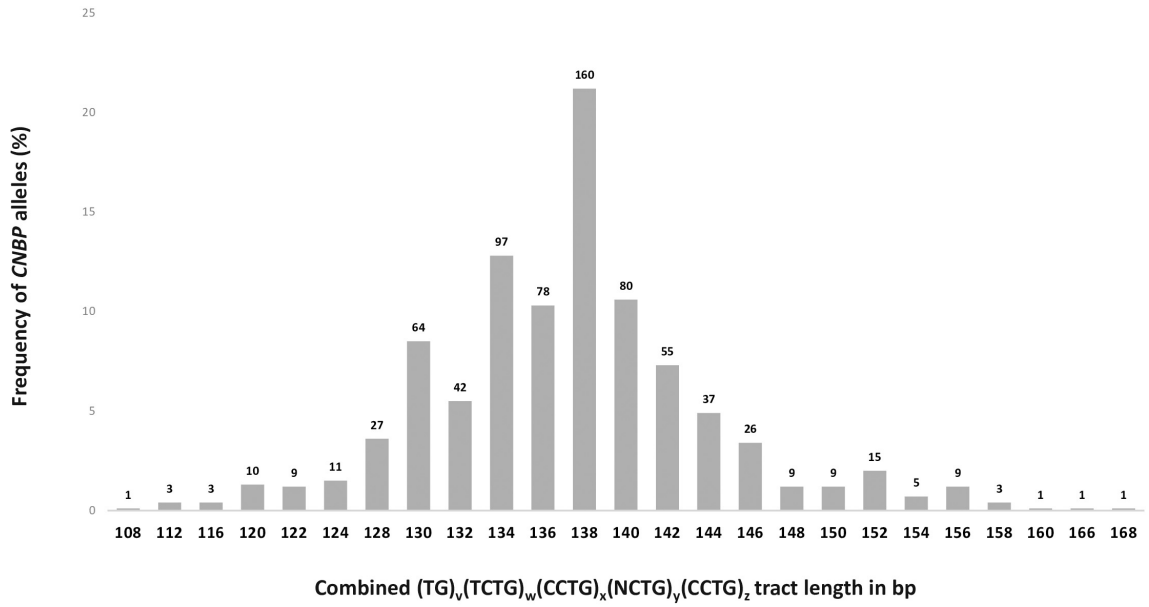



**FIGURE 3 |** Frequency and distribution of *CNBP* healthy range alleles in DM2-negative Italian individuals (*n* = 378). Absolute number of alleles has been reported above the bars in the graph (total number of alleles = 756).

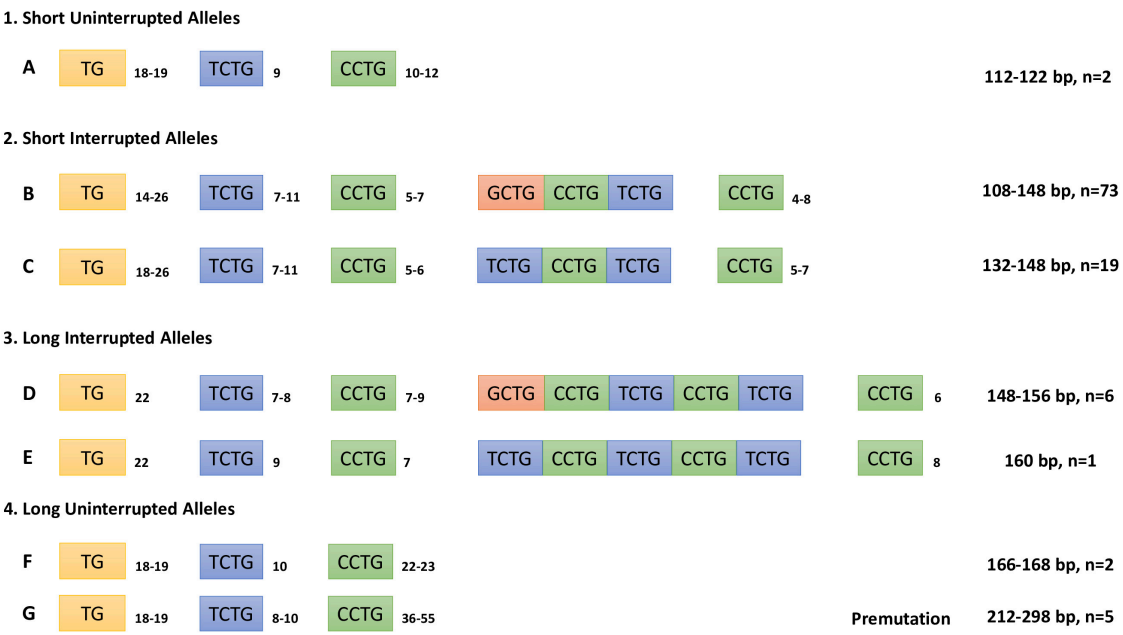



**FIGURE 4 |** Composition of the combined repeat tract $(TG)_v(TCTG)_w(CCTG)_x(NCTG)_y(CCTG)_z$ of *CNBP* non-expanded alleles determined by Sanger sequencing analysis. Repeat tract length in bp and number of alleles characterized are also indicated.

population (Matsuura et al., 2012), and in the United States, clinical experience suggests that DM2 is roughly fivefold less common than DM1 (Thornton, 2014; Meola and Cardani, 2015). Data about the prevalence of DM2 in Europe are limited except for the German and Finland populations, where it may be like that reported for DM1 (Suominen et al., 2011; Mahyera et al., 2018). The only study performed in Italy, so far, includes the Rome province with an estimated prevalence of DM2, 10% that of DM1 (Vanacore et al., 2016). Unfortunately, we cannot calculate with accuracy the true prevalence of the DM2 disease in Italy because our data do not include the totality of DM2 patients overall the country. Nevertheless, since

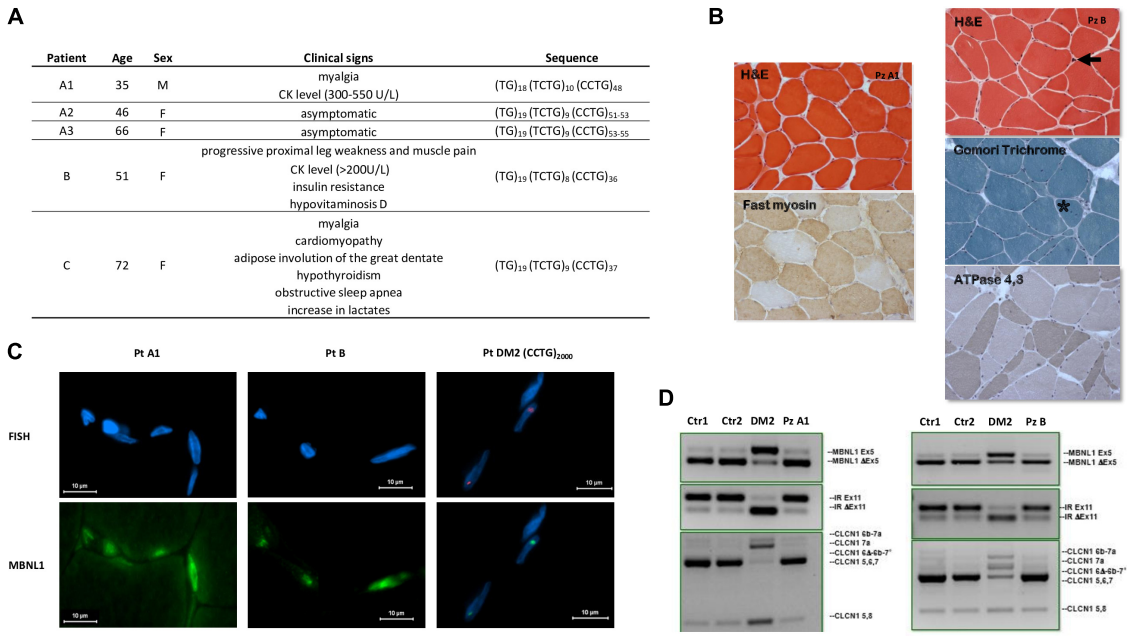



**FIGURE 5 |** Patients with *CNBP* premutated alleles. **(A)** Clinical characteristics of DM2 premutated patients and sequence of the uninterrupted *CNBP* alleles. **(B)** Left: Patient A1, Biceps brachii muscle biopsy performed at the age of 16 shows only a slight fiber size variability both of type I and II fibers (200×). Right: Patient B, Biceps brachii muscle biopsy performed at the age of 51 shows an increase of fiber size variability, nuclear clumps (arrow), central nuclei (asterisk) and type I (200×). **(C)** FISH + MBNL1 immunofluorescence of patient A1 and B shows no nuclear accumulation of *CNBP* mutant RNA or of MBNL1 as instead observable in a DM2 patient with $(CCTG)_{2000}$ repetitions. **(D)** Similarly to healthy individuals (Ctr1 and Ctr2), no alterations of alternative splicing of *MBNL1*, *IR*, and *CLCN1* genes have been evidenced by RT-PCR analysis compared with a DM2 patient, used as positive control.

we are one of the main national diagnostic labs for myotonic dystrophies, we can estimate that the DM2 tests performed represent about 50% of all DM2 analyses in Italy. This allows us to calculate an estimated prevalence of DM2 in the Italian population that is about 10% that of DM1, in accordance with our previous published data (Vanacore et al., 2016). Moreover, additional important considerations are possible. The first is that we observed a very low percentage of DM2-positive test (33%) compared with the number of samples sent to our lab with a clinical suspicion of DM2. This positivity rate is also lower than what we have experienced in DM1 genetic testing (45% from our unpublished data) and suggests that clinical indications for DM2 molecular diagnosis are generally less defined than in DM1. This conclusion fully agrees with what was reported by Mahyera et al. (2018) for one of the main DM diagnostic labs in Germany and demonstrates how most clinicians still do not easily recognize the cardinal features of the DM2 disease probably due to its rather aspecific and late onset. Consequently, the average delay between the onset of symptoms and the genetic confirmation of DM2 is 14 years (Hilbert et al., 2013). Patients with DM2 usually come to medical attention because of myotonia, prominent in distal limb muscles, however, this symptom is present in less than 50% of subjects. Other symptoms such as cardiac and endocrine dysfunctions are not specifically associated with the DM2 phenotype being present in less than 5% of DM2-positive tests. Recent studies indicate that, unlike DM1, either *CLCN1* or *SCN4A* acts as genetic modifiers in DM2 patients since mutations/polymorphisms in these genes may contribute to exaggerated phenotype with more severe muscle stiffness and myotonia (Suominen et al., 2008; Cardani et al., 2012; Bugiardini et al., 2015; Peddareddygari et al., 2016; Binda et al., 2018). This suggests that DM2 patients with co-segregating *CLCN1* or *SCN4A* genetic variants could be more easily identified than DM2 patients without the modifier alleles, and consequently that the majority of patients remain undiagnosed even in clinical centers with considerable experience with DM2. Among 190 DM2 confirmed patients, we did not observe a different distribution of the disease symptoms and age of onset between males and females. Gender differences have been identified with a tendency toward the worsening of symptoms in females (Montagnese et al., 2017). However, we should point out that, different from the study of Montagnese et al., our work is not focused to provide a detailed clinical characterization of DM2 patients. Most of the clinical information has been collected retrospectively over time from different Italian centers with no follow-up of patients; it is therefore possible that some clinical data may be incomplete, underestimated, or missed at all. For this reason, a gender and age influence on the DM2 phenotype cannot be ruled out on the basis of our data. Comparison with the complex $(TG)_v(TCTG)_w(CCTG)_x(NCTG)_y(CCTG)_z$ repeat tract length in the Italian population vs. the German and Slovak populations showed that the distribution and the size of the *CNBP* repeat tract is conserved among different European populations (Radvanszky et al., 2013; Mahyera et al., 2018). Normal *CNBP* alleles from

DM2-negative probands showed a unimodal distribution with allele sizes ranging from 108 to 168 bps, including very rare alleles bigger than 158 bp in the larger extreme end of the distribution. The range of *CNBP* normal alleles was 118–156 bp in a cohort of 95 proband in the Slovak population and 102–166 in the German population, where a cohort of 739 probands, closer in size to the one reported here, has been analyzed. The four most frequent *CNBP* allele among Italians consisted of 138 (21%), 134 (13%), 136 (11%), and 140 (10%) bps. Frequencies of the 138-, 134-, and 140-bp alleles are identical to the German population, whereas the 142-bp allele, which was found to be the third most frequent allele in Germany and the second most frequent allele in Slovenia, was at fifth place among the Italians (11 and 13% vs. 7% in Italians). However, it should be pointed out that these studies analyzed different study cohorts consisting of normal individuals from the general population (Radvanszky et al., 2013) or DM2-negative probands showing DM2-like symptoms (Mahyera et al., 2018 and our study). In normal-sized alleles from DM2 patients, the CCTG repeat tract is usually found to be interrupted by one or two tetraplets $(NCTG)_n$, which modulate the genetic stability of the locus. However, as the difference between normal alleles (repeats to 26 CCTG units) and disease-associated alleles (75 units and more, with a mean of 5,000) is usually evident, exact sizing is not routinely performed. The variability in the sizes of the different subunits of the complex repeat, $(TG)_v(TCTG)_w(CCTG)_x$, in the *CNBP* gene complicates the determination of the precise number of the CCTG unit within this repeat, which can only be determined by DNA sequencing. Our characterization of *CNBP* normal alleles confirms the highly polymorphic nature of the $(TG)_v(TCTG)_w$ repeated tract compared with the whole $(TG)_v(TCTG)_w(CCTG)_x$ motif. Indeed, a great variability of these repeat units was found, between $(TG)_{14-26}$ and $(TCTG)_{7-11}$, which is not proportional to the increase in allelic size. Interestingly, sequencing analysis allowed us to identify two short uninterrupted alleles containing, respectively, $(CCTG)_{10}$ and $(CCTG)_{12}$ repetitions. These rare alleles found in our study cohort, represent only 2% of normal *CNBP* chromosomes, in accordance with previous analysis (Radvanszky et al., 2013). The existence of *CNBP* alleles with uninterrupted CCTG repeated tract has been reported for the first time by Liquori et al. (2003) who described a 136-bp allele containing $(CCTG)_{20}$, thus, considering the 20 CCTG uninterrupted repeated as the shortest possible DM2 premutation allele. A few years later, Bachinski et al. (2009) identified 4 of 44 alleles with uninterrupted $(CCTG)_{24-32}$. More recently, in the Slovak population, Radvanszky described six additional *CNBP* uninterrupted alleles containing up to 70 repeats in a large cohort of about 500 individuals. Interestingly, short alleles containing $(CCTG)_{12-14}$ among the uninterrupted alleles have been reported, according to our data (Radvanszky et al., 2013). Finally, Mahyera et al. (2018) detected the presence of a 168-bp healthy allele containing $(CCTG)_{23}$ in a total of 34 alleles in DM2-positive patients. There are at least two main still unsolved questions regarding *CNBP* uninterrupted alleles below the "70 CCTG threshold": their genetic instability (either somatic or germinal) and their clinical significance. The international guidelines for the DM2 molecular diagnosis consider non-pathogenic *CNBP* alleles as those having up to $(CCTG)_{26}$ with one or more interruptions (Bachinski et al., 2009; Kamsteeg et al., 2012), whereas disease-associated alleles contain 75 units and more uninterrupted CCTG units. No information is given on the range and penetrance of premutated *CNBP* alleles comprised of between 26 and 75 CCTG units. What do we know about these premutated alleles from literature? The paper from Bachinski and colleagues demonstrated that uninterrupted $(CCTG)_{22-33}$ alleles are unstable and represent a premutation allele pool for DM2 full mutations. Mahyera et al. (2018) also reported two patients harboring $(CCTG)_{55}$ and $(CCTG)_{30}$ *CNBP* alleles with a DM2-like phenotype and (mild myopathy), respectively. (CCTG)70 and (CCTG) > 70 *CNBP* alleles were found in individuals belonging to families with genetically proven Huntington disease (HD); unfortunately, no detailed clinical information was available from these patients (Radvanszky et al., 2013). Our study enlarges the cohort of *CNBP* premutated individuals described so far. During our diagnostic procedure, we characterized five *CNBP* premutated alleles ranging from 36 to 55 uninterrupted CCTG repetitions. The analysis of our two familial cases confirmed that *CNBP* alleles containing more than 50 uninterrupted CCTG are unstable, as previously described. However, the clinical consequences are uncertain leading to only mild myopathy and hyperCKaemia in one patient, whereas his mother is asymptomatic by the age of 66. On the other side, one patient (patient C) carrying $(CCTG)_{37}$ repetitions shows a more complex phenotype with myalgia, hyperthyroidism, and familiar cardiomyopathy. However, no additional genetic tests have been performed to exclude the presence of other genetic causes associated with the observed clinical phenotype. Histological examinations of muscle tissues, available only for two of these premutated individuals, showed mild muscular pathogenic signs with atrophy of both types I and II fibers, which have not been specifically associated with DM2 where high rates of atrophic and hypertrophic type 2 fibers were observed (Pisani et al., 2008). Moreover, no molecular hallmarks of DM (*IR*, *MBNL1*, *CLCN1* splicing alterations, and CCUG-containing foci accumulation) have been detected in muscle tissues. Our data argue in favor of an incomplete penetrance for uninterrupted *CNBP* alleles in the range of $(CCTG)_{36-55}$ repetitions. Taken together, our published studies indicated that there is no valid "cut-off" to distinguish *CNBP* alleles in the clinical range from those in the non-pathological range. Rather, there is an overlap of repeat sizes, some of which will be associated with mild disease expression, while others will not, depending upon still unknown modifying factors whose identification may represent a fruitful approach to defining therapeutic strategies in DM2. To conclude, our findings extend current knowledge concerning the polymorphic spectrum of the $(TG)_v$ and $(TCTG)_w$ tract variability and the occurrence of intergenerational segregation of *CNBP* alleles with uninterrupted CCTG tract. Our data argues in favor of a relative instability of *CNBP* premutated alleles, even if their clinical relevance is still a matter of debate. In this context, the molecular and clinical characterization of additional individuals would be desirable to provide improved genetic testing and counseling for DM2 patients.

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethical Committee PTV Tor Vergata Hospital study # 232/19. The patients/participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

AB conceived and designed the study. AB, VVV, MRD'A, and GN drafted and critically revised the manuscript. VVV, LF, and PB collected the data. VVV, LF, IB, PB, MRD'A, and FS handled the genetic testing and molecular characterization of DM2 locus. MB contributed to genetic counselling. RM, GM, GA, AP, and EP contributed to patient's collection and clinical evaluation. RC and GM contributed to muscle biopsies collection and analysis. All authors contributed to the article and approved the submitted version.

## FUNDING

This research was funded by the "Beyond Borders" project from "Tor Vergata" University of Rome #E89C20000500005. GM was supported by the Fondazione Malattie Miotoniche (FMM), Milan, Italy.

## ACKNOWLEDGMENTS

We thank the patients and their families for participating. We acknowledge Silvia Russo and Raniero Iraci for the technical assistance with genotyping of the DM2 patients. EP is a member of the European Reference Network for Neuromuscular Diseases – Project ID No. 739543. We acknowledge the Telethon Genetic BioBank (GTB12001D) and the Eurobiobank Network.

## REFERENCES

Bachinski, L. L., Czernuszewicz, T., Ramagli, L. S., Suominen, T., Shriver, M. D., Udd, B., et al. (2009). Premutation allele pool in myotonic dystrophy type 2. *Neurology* 72, 490–497. doi: 10.1212/01.wnl.0000333665.01888.33

Bachinski, L. L., Udd, B., Meola, G., Sansone, V., Bassez, G., Eymard, B., et al. (2003). Confirmation of the type 2 myotonic dystrophy (CCTG)n expansion mutation in patients with proximal myotonic myopathy/proximal myotonic dystrophy of different European origins: a single shared haplotype indicates an ancestral founder effect. *Am. J. Hum. Genet.* 73, 835–848. doi: 10.1086/378566

Binda, A., Renna, L. V., Bosè, F., Brigonzi, E., Botta, A., Valaperta, R., et al. (2018). SCN4A as modifier gene in patients with myotonic dystrophy type 2. *Sci. Rep.* 8:11058. doi: 10.1038/s41598-018-29302-z

Botta, A., Bonifazi, E., Vallo, L., Gennarelli, M., Garrè, C., Salehi, L., et al. (2006). Italian guidelines for molecular analysis in myotonic dystrophies. *Acta Myol.* 25, 23–33.

Bugiardini, E., Rivolta, I., Binda, A., Soriano Caminero, A., Cirillo, F., Cinti, A., et al. (2015). SCN4A mutation as modifying factor of myotonic dystrophy type 2 phenotype. *Neuromuscul. Disord.* 25, 301–307. doi: 10.1016/j.nmd.2015.01.006

Cardani, R., Baldassa, S., Botta, A., Rinaldi, F., Novelli, G., Mancinelli, E., et al. (2009). Ribonuclear inclusions and MBNL1 nuclear sequestration do not affect myoblast differentiation but alter gene splicing in myotonic dystrophy type 2. *Neuromuscul. Disord.* 19, 335–343. doi: 10.1016/j.nmd.2009.03.002

Cardani, R., Giagnacovo, M., Botta, A., Rinaldi, F., Morgante, A., Udd, B., et al. (2012). Co-segregation of DM2 with a recessive CLCN1 mutation in juvenile onset of myotonic dystrophy type 2. *J. Neurol.* 259, 2090–2099. doi: 10.1007/s00415-012-6462-1

Coenen, M. J. H., Tieleman, A. A., Schijvenaars, M. M. V. A. P., Leferink, M., Ranum, L. P. W., Scheffer, H., et al. (2011). Dutch myotonic dystrophy type 2 patients and a North-African DM2 family carry the common European founder haplotype. *Eur. J. Hum. Genet.* 19, 567–570. doi: 10.1038/ejhg.2010.233

Day, J. W., Ricker, K., Jacobsen, J. F., Rasmussen, L. J., Dick, K. A., Kress, W., et al. (2003). Myotonic dystrophy type 2: molecular, diagnostic and clinical spectrum. *Neurology* 60, 657–664.

Guo, P., and Lam, S. L. (2016). Short conceptual overview unusual structures of CCTG repeats and their participation in repeat expansion. *Biomol. Concepts* 7, 331–340. doi: 10.1515/bmc-2016-0024

Hilbert, J. E., Ashizawa, T., Day, J. W., Luebbe, E. A., Martens, W. B., McDermott, M. P., et al. (2013). Diagnostic odyssey of patients with myotonic dystrophy. *J. Neurol.* 260, 2497–2504. doi: 10.1007/s00415-013-6993-0

Kamsteeg, E. J., Kress, W., Catalli, C., Hertz, J. M., Witsch-Baumgartner, M., Buckley, M. F., et al. (2012). Best practice guidelines and recommendations on the molecular diagnosis of myotonic dystrophy types 1 and 2. *Eur. J. Hum. Genet.* 20, 1203–1208. doi: 10.1038/ejhg.2012.108

Liquori, C. L., Ikeda, Y., Weatherspoon, M., Ricker, K., Schoser, B. G. H., Dalton, J. C., et al. (2003). Myotonic dystrophy type 2: human founder haplotype and evolutionary conservation of the repeat tract. *Am. J. Hum. Genet.* 73, 849–862. doi: 10.1086/378720

Liquori, C. L., Ricker, K., Moseley, M. L., Jacobsen, J. F., Kress, W., Naylor, S. L., et al. (2001). Myotonic dystrophy Type 2 caused by a CCTG expansion in intron 1 of ZNF9. *Science* 293, 864–867. doi: 10.1126/science.1062125

Mahyera, A. S., Schneider, T., Halliger-Keller, B., Schrooten, K., Hörner, E.-M., Rost, S., et al. (2018). Distribution and structure of DM2 repeat tract alleles in the German population. *Front. Neurol.* 9:463. doi: 10.3389/fneur.2018.00463

Matsuura, T., Minami, N., Arahata, H., Ohno, K., Abe, K., Hayashi, Y. K., et al. (2012). Myotonic dystrophy type 2 is rare in the Japanese population. *J. Hum. Genet.* 57, 219–220. doi: 10.1038/jhg.2011.152

Meola, G., Biasini, F., Valaperta, R., Costa, E., and Cardani, R. (2017). Biomolecular diagnosis of myotonic dystrophy type 2: a challenging approach. *J. Neurol.* 264, 1705–1714. doi: 10.1007/s00415-017-8504-1

Meola, G., and Cardani, R. (2015). Myotonic dystrophy type 2: an update on clinical aspects, genetic and pathomolecular mechanism. *J. Neuromuscul. Dis.* 2, S59–S71. doi: 10.3233/JND-150088

Meola, G., and Cardani, R. (2017). Myotonic dystrophy type 2 and modifier genes: an update on clinical and pathomolecular aspects. *Neurol. Sci.* 38, 535–546. doi: 10.1007/s10072-016-2805-5

Montagnese, F., Mondello, S., Wenninger, S., Kress, W., and Schoser, B. (2017). Assessing the influence of age and gender on the phenotype of myotonic dystrophy type 2. *J. Neurol.* 264, 2472–2480. doi: 10.1007/s00415-017-8653-2

Nakamori, M., Sobczak, K., Puwanant, A., Welle, S., Eichinger, K., Pandya, S., et al. (2013). Splicing biomarkers of disease severity in myotonic dystrophy. *Ann. Neurol.* 74, 862–872. doi: 10.1002/ana.23992

Peddareddygari, L. R., Grewal, A. S., and Grewal, R. P. (2016). Focal seizures in a patient with myotonic disorder type 2 co-segregating with a chloride voltage-gated channel 1 gene mutation: a case report. *J. Med. Case Rep.* 10:167.

Pisani, V., Panico, M. B., Terracciano, C., Bonifazi, E., Meola, G., Novelli, G., et al. (2008). Preferential central nucleation of type 2 myofibers is an invariable feature of myotonic dystrophy type 2. *Muscle Nerve* 38, 1405–1411. doi: 10.1002/mus.21122

Radvanszky, J., Surovy, M., Polak, E., and Kadasi, L. (2013). Uninterrupted CCTG tracts in the myotonic dystrophy type 2 associated locus. *Neuromuscul. Disord.* 23, 591–598. doi: 10.1016/j.nmd.2013.02.013

Suominen, T., Bachinski, L. L., Auvinen, S., Hackman, P., Baggerly, K. A., Angelini, C., et al. (2011). Population frequency of myotonic dystrophy: higher than expected frequency of myotonic dystrophy type 2 (DM2) mutation in Finland. *Eur. J. Hum. Genet.* 19, 776–782. doi: 10.1038/ejhg.2011.23

Suominen, T., Schoser, B., Raheem, O., Auvinen, S., Walter, M., Krahe, R., et al. (2008). High frequency of co-segregating CLCN1 mutations among myotonic dystrophy type 2 patients from Finland and Germany. *J. Neurol.* 255, 1731–1736. doi: 10.1007/s00415-008-0010-z

Thornton, C. A. (2014). Myotonic dystrophy. *Neurol. Clin.* 32, 705–719. doi: 10.1016/j.ncl.2014.04.011

Turner, C., and Hilton-Jones, D. (2010). The myotonic dystrophies: diagnosis and management. *J. Neurol. Neurosurg. Psychiatry* 81, 358–367.

Vanacore, N., Rastelli, E., Antonini, G., Bianchi, M. L. E., Botta, A., Bucci, E., et al. (2016). An age-standardized prevalence estimate and a sex and age distribution of myotonic dystrophy types 1 and 2 in the Rome province, Italy. *Neuroepidemiology* 46, 191–197. doi: 10.1159/000444018

**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.

The handling editor declared a past co-authorship with one of the authors RM.

*Copyright © 2021 Botta, Visconti, Fontana, Bisceglia, Bengala, Massa, Bagni, Cardani, Sangiuolo, Meola, Antonini, Petrucci, Pegoraro, D'Apice and Novelli. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*

frontiers in Pediatrics

CASE REPORT
published: 23 June 2021
doi: 10.3389/fped.2021.650920

# Transient Neonatal Diabetes Mellitus in SHORT Syndrome: A Case Report

*Shin-Hee Kim[1], Minsung Kim[1], Jisook Yim[2,3], Myungshin Kim[2,3] and Dae-Hyun Jang[4]**

[1] Department of Pediatrics, Incheon St. Mary's Hospital, College of Medicine, The Catholic University of Korea, Seoul, South Korea, [2] Department of Laboratory Medicine, College of Medicine, The Catholic University of Korea, Seoul, South Korea, [3] Catholic Genetic Laboratory Center, Seoul St. Mary's Hospital, College of Medicine, The Catholic University of Korea, Seoul, South Korea, [4] Department of Rehabilitation Medicine, Incheon St. Mary's Hospital, College of Medicine, The Catholic University of Korea, Seoul, South Korea

SHORT syndrome is a rare autosomal dominant disorder characterized by multiple congenital defects and is historically defined by its acronym: short stature, hyperextensibility of joints and/or inguinal hernia, ocular depression, Rieger anomaly, and teething delay. Herein, we report a male infant with SHORT syndrome who presented with transient neonatal diabetes mellitus (TNDM) with insulin resistance. The proband was born at 38 weeks of gestation but displayed facial dysmorphic features. Intrauterine growth restriction (IUGR) was detected on a prenatal ultrasonography test. His birth weight was 1.8 kg (<3rd percentile), length 44 cm (<3rd percentile), and head circumference 31 cm (<3rd percentile). The patient's blood glucose level started to increase at 5 days of age (218–263 mg/dl) and remained high at 20 days of age (205–260 mg/dl). He was treated with subcutaneous insulin and the blood glucose level gradually stabilized. Blood glucose level was stabilized over time without insulin treatment at 6 weeks of age. Clinical exome sequencing showed a heterozygous pathogenic variant, NM_181523.3:c.1945C>T (p.Arg649Trp) in exon 15 of the phosphoinositide-3-kinase regulatory subunit 1 (*PIK3R1*) known as the causative gene for SHORT syndrome. Examination of the patient at 10 months of age revealed no hyperglycemic episode and glycated hemoglobin level was 5.2%. To the best of our knowledge, this is the first case of TNDM in SHORT syndrome due to a pathogenic variant of *PIK3R1*. We believe that our case can aid in expanding the phenotypes of SHORT syndrome.

**Keywords: SHORT syndrome, PIK3R1 gene, transient neonatal diabetes mellitus, insulin resistance, dysmorphic feature**

OPEN ACCESS

***Edited by:***
*Kathleen M. Gorman, Temple Street Children's University Hospital, Ireland*

***Reviewed by:***
*Brian Hon-Yin Chung, The University of Hong Kong, SAR China*
*Emanuele Micaglio, IRCCS Policlinico San Donato, Italy*

***Correspondence:**
*Dae-Hyun Jang*
*dhjangmd@naver.com*

***Specialty section:***
*This article was submitted to Genetics of Common and Rare Diseases, a section of the journal Frontiers in Pediatrics*

***Received:*** *08 January 2021*
***Accepted:*** *24 May 2021*
***Published:*** *23 June 2021*

***Citation:***
*Kim S-H, Kim M, Yim J, Kim M and Jang D-H (2021) Transient Neonatal Diabetes Mellitus in SHORT Syndrome: A Case Report. Front. Pediatr. 9:650920. doi: 10.3389/fped.2021.650920*

## INTRODUCTION

SHORT syndrome is a rare genetic disorder with features that include short stature, hyperextensibility of joints and/or inguinal hernia, ocular depression, Rieger anomaly, and teething delay (1). Additional clinical features include intrauterine growth restriction (IUGR), partial lipodystrophy, inguinal hernia, sensorineural hearing loss, and facial characteristics such as triangular face, small chin, low set posteriorly rotated ears (2, 3). SHORT syndrome was first described in 1975 and to date, fewer than 50 cases have been reported in the literature (2, 3). SHORT syndrome is caused by a heterozygous loss of function in a variant of phosphoinositide-3-kinase regulatory subunit 1 (*PIK3R1*) gene. *PIK3R1* codes for the regulatory subunits of the phosphatidyl inositol-3 kinase of class IA (PI3K) and is involved in the activation of the AKT/mTOR pathway

to ensure proper growth and cell proliferation. A pathogenic variant of *PIK3R1* in SHORT syndrome disrupts not only IGF-1R signaling but also insulin signaling and thereby predispose the individual to insulin resistance and diabetes mellitus. It reduce its insulin-stimulated activity that leads to lower AKT and mTOR phosphorylation (4–6). Reduced insulin signaling may also limit intrauterine growth, usually resulting in very small infants at birth (< 3rd percentile) (7).

In the present study, we report an infant with transient neonatal diabetes mellitus (TNDM) with insulin resistance who was finally diagnosed with SHORT syndrome due to a pathogenic variant in *PIK3R1*.Written informed consent was obtained from the parents for the publication of this case report.

## Case Report

The patient was a newborn male and the only child of a healthy non-consanguineous Korean couple with a non-contributory family history. The height of his father and mother was 170 cm (−0.70 SD score) and 160 cm (−0.04 SD score), respectively. They had no dysmorphic features. The mother had regular antenatal check-up and did not have any history of medical and obstetric problems during pregnancy. He was born at 38 weeks of gestation but displayed features of IUGR during pregnancy. His birth weight was 1.8 kg (<3rd percentile), length 44 cm (<3rd percentile), and head circumference 31 cm (<3rd percentile) according to the Korean reference for birth weight based on gestational age and sex. The initial blood glucose level was 70 mg/dl. The baby was exclusively breastfed starting on day 3 and was in generally good condition. However, blood glucose level was between 218 and 263 mg/dl at 5 day of age. At the age of 20 day, his blood glucose level was still high (205–260 mg/dl), and the infant was referred to the endocrine clinic for persistent hyperglycemia assessment. On physical examination, several dysmorphic features (triangular-shaped face, prominent forehead, ocular depression, lipodystrophy at the lumbar region) and inguinal hernia were present. The systolic and diastolic blood pressure measurements were 74 and 42 mmHg, respectively. The serum c-peptide and insulin levels were 2.83 ng/ml (normal: 1.0–3.5) and 120 μU/ml (normal: 2.8–13.5), respectively. Baseline chemistry including serum blood urea nitrogen was 15.3 mg/dl (normal: 7.0–20.0), creatinine 0.9 mg/dl (normal: 0.6–1.2), aspartate aminotransferase 38 U/L (normal: 14–40), and alanine aminotransferase 16 U/L (normal: 9–45), as well as complete blood count profile were within normal range. Urinalysis showed no glucose or ketones. There was no sign of ketoacidosis and the patient had no type 1 diabetes autoantibodies (antibodies against glutamic acid decarboxylase, islet cell, islet antigen-2, and insulin). The liver and pancreas ultrasonography revealed no structural abnormality. Echocardiography at the age of 1 month confirmed mild pulmonary stenosis and ASD secundum (2 mm) which did not require surgical intervention. Neonatal diabetes mellitus (NDM) was suspected on the basis of hyperglycemia occurring within the first month of life that lasted for >2 weeks and required insulin therapy. At age of 25 day, clinical exome sequencing was performed to identify the genetic cause of NDM.

To monitor the glycemic level, his blood glucose was measured at the beginning of each feeding session. The patient was treated with subcutaneous insulin, and blood glucose level gradually stabilized. The blood glucose levels ranged from 110–250 mg/dl during the next 10 days. An adequate glucose level was achieved at 6 weeks of age without insulin treatment. His body weight was 4.4 kg (<3rd percentile) and his length was 61.6 cm (<3rd percentile) at 10 months of age. The patient experienced no hyperglycemic episode and the glycated hemoglobin was 5.0% and insulin level 2.8 μU/ml. At 10 months of age, the patient had no teeth erupted in the oral cavity.

## Genetic Testing

Genomic DNA was extracted from peripheral blood samples using the QIAsymphony DSP DNA mini kit (Qiagen, Hilden, Germany). Sequencing was performed using a TruSight One sequencing panel (Illumina Inc., San Diego, CA, USA) consisting of 4,813 genes associated with known Mendelian genetic disorders on a NextSeq 550 (Illumina, CA, USA). Variants were assessed using the ACMG variant classification guidelines (8). Subsequently, a heterozygous missense variant, NM_181523.3: c.1945C>T (p.Arg649Trp) in exon 15 of the *PIK3R1* gene was identified (**Figure 1**), which was previously reported as the pathogenic variant of SHORT syndrome (4). This variant was validated using conventional Sanger sequencing. The clinical exome sequencing revealed no other pathogenic variants in genes associated with NDM (*KCNJ11*, *ABCC8*, *INS*, *GATA6*, *EIF2AK3*, *FOXP3*), mitochondrial diabetes (*POLG*, *POLG2*, *OPA1*, *RRM2B*), and maturity onset diabetes of the young (*HNF1A*, *GCK*, *HNF4A*, *ABCC8*, *BLK*, *HNF1B*, *INS*, *KCNJ11*, *KFL11*, *PAX4*, and *PDX1*). Segregation analysis could not be performed due to the unavailability of parental samples.

# DISCUSSION

Herein, we report a case of TNDM in SHORT syndrome caused by a pathogenic variant in *PIK3R1*. SHORT syndrome is inherited in an autosomal dominant manner and caused by a heterozygous loss of function variant of *PIK3R1* located on chromosome 5 (5q13.1) (2). To date, fewer than 50 cases have been reported in the literature (3). Approximately 10 different pathogenic variants in *PIK3R1* have been found to cause SHORT syndrome and the c.1945C>T (p.Arg649Trp) variant was identified most frequently (2, 9). Approximately half of the patients exhibit at least four of the five characteristics that comprise its acronym. Short stature, ocular depression, and teething delay are the most common findings, with hyperextensibility of joints, and Rieger anomaly being less common (2). Our patient exhibited two of the five characteristics in the SHORT syndrome acronym (inguinal hernia, and ocular depression). His length was less than 3rd percentile and he had no teeth erupted at last visit. However, further follow-up is need to accurately the presence of short stature and teething delay because he was only 10 months old.

*PIK3R1* mutation in SHORT syndrome appears to disrupt the insulin signaling pathway and thereby predispose the patient to insulin resistance and diabetes mellitus (4–6). Of the 17 previously reported patients with c.1945C>T mutations, 9 (53%) developed diabetes mellitus and/or insulin resistance (4–6, 10, 11). The authors suggested 1945C>T (p.Arg649Trp) led to

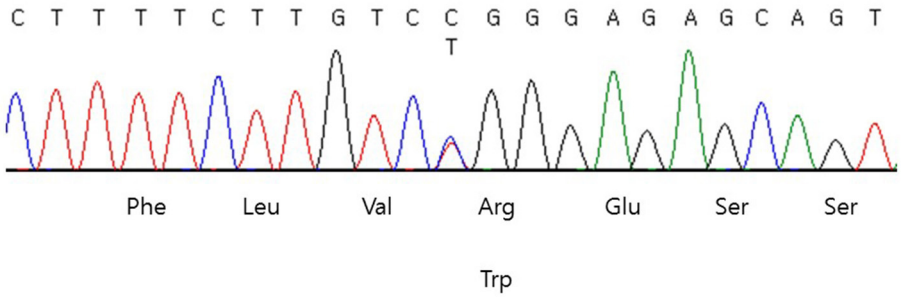



**FIGURE 1 |** DNA sequencing chromatogram of the patient. Heterozygous missense variant, NM_181523.3:c.1945C>T (p.Arg649Trp) in exon 15 of the *PIK3R1* was identified.

impaired interaction between p85a and IRS-1 and reduced AKT-mediated insulin signaling in fibroblasts from affected subjects (6). Insulin resistance/diabetes mellitus has also been diagnosed in patients with the other following pathogenic variants: c.1929_1933delTGGCA, c.1615_1617delATT, c.1465G>A, c.1943dupT, and c.1892G>A (5, 12).

Diabetes mellitus is common in adults with SHORT syndrome but infrequently observed in adolescents/young children, indicating that time (and/or other susceptibility factors) may be required for insulin resistance to develop (2). Our patient experienced remission of diabetes at the age of 6 week, and he is currently 10 months of age, euglycemic, and not taking any insulin or oral hypoglycemic agents. However, TNDM frequently recur in childhood, during puberty, or later in adulthood. Busiah et al. reported that remission of TNDM occurred in 51% (89/174) of patients at a median age of 17 weeks and recurrences occurred in 26% (23/89) of patients with TNDM at a median age of 12.7 years (13). When reviewing the literature to date, TNDM in SHORT syndrome has not been reported. In our patient, there were no symptoms even though the glucose level was high, insulin requirements were low, and NDM duration was short. In SHORT syndrome, DM may be present in newborns and overlooked. Therefore, we recommend that clinicians should suspect SHORT syndrome for newborns with NDM, IUGR, and facial gestalt. Genetic diagnosis of SHORT syndrome is important for an appropriate treatment and genetic counseling in affected families. Insulin resistance/diabetes mellitus, Rieger anomaly/glaucoma, dental anomalies, and hearing loss can often be treated by appropriate medical specialists if present. Clinical exome sequencing allowed us to simultaneously evaluate many genes associated with early onset diabetes mellitus. Uniparental disomy of chromosome 6 (UPD 6) is one of the important genetic causes of NDM, but it could not be evaluated by clinical exome sequencing.

IUGR is commonly observed in cases of TNDM because insulin acts as a fetal growth hormone. Consequently, insulin insufficiency coupled with the failure of transplacental insulin delivery causes newborns with TNDM to be born small for their gestational age (14). Reduced insulin signaling may also limit intrauterine growth because individuals with SHORT syndrome are usually very small at birth (<3rd percentile) and varying degrees of short stature are usually present throughout childhood (2). Despite the short stature of the patients, growth hormone therapy is not recommended for SHORT syndrome patients because of the potential to promote the development of diabetes mellitus and poor responses. Conversely, in a recent study, pubertal development and/or advanced age rather than growth hormone therapy was shown to play a critical role in the development of insulin resistant diabetes mellitus (IRDM) (11). The usefulness of an SGLT2 inhibitor and metformin in the management of IRDM in SHORT syndrome was reported (10, 11). However, in a recent study, metformin treatment was shown to paradoxically lead to a potential deterioration of insulin resistance and development of glucose intolerance in SHORT syndrome patients (15). Further studies are warranted to confirm the clinical benefits of oral hypoglycemic agents for the treatment of genetic syndromes of insulin resistance.

To the best of our knowledge, this is the first case report of SHORT syndrome with TNDM. Our findings provide important information that TNDM with insulin resistance can be a phenotype of SHORT syndrome. Familiarization and discernment of complicated genetic syndromes can enable early implementation of appropriate management and proper care of the patient.

## DATA AVAILABILITY STATEMENT

The data presented in this study are available on request from the corresponding author.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Incheon St. Mary's Hospital Institutional Review Board. 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 minor(s)' legal guardian/next of kin for the publication of any potentially identifiable images or data included in this article. Informed consent was obtained from the patient's parents for the publication of this case report.

## AUTHOR CONTRIBUTIONS

S-HK, D-HJ, and MyK wrote the manuscript. JY and MyK analyzed the genetic study. All the authors contributed to the manuscript.

## REFERENCES


1. Gorlin RJ, Cervenka J, Moller K, Horrobin M, Witkop CJ. Malformation syndromes. A selected miscellany. *Birth Defects Orig Artic Ser.* (1975) 11:39–50.
2. Avila M, Dyment DA, Sagen JV, St-Onge J, Moog U, Chung BHY, et al. Clinical reappraisal of SHORT syndrome with *PIK3R1* mutations: toward recommendation for molecular testing and management. *Clin Genet.* (2016) 89:501–6. doi: 10.1111/cge.12688
3. Innes AM, Dyment DA. SHORT syndrome. In: Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJ, G. Mirzaa G, et al., editors. *GeneReviews®*. Seattle, WA: University of Washington (2021).
4. Dyment DA, Smith AC, Alcantara D, Schwartzentruber JA, Basel-Vanagaite L, Curry CJ, et al. Mutations in *PIK3R1* Cause SHORT Syndrome. *Am J Hum Genet.* (2013) 93:158–66. doi: 10.1016/j.ajhg.2013.06.005
5. Thauvin-Robinet C, Auclair M, Duplomb L, Caron-Debarle M, Avila M, St-Onge J, et al. *PIK3R1* Mutations cause syndromic insulin resistance with lipoatrophy. *Am J Hum Genet.* (2013) 93:141–9. doi: 10.1016/j.ajhg.2013.05.019
6. Chudasama KK, Winnay J, Johansson S, Claudi T, König R, Haldorsen I, et al. SHORT syndrome with partial lipodystrophy due to impaired phosphatidylinositol 3 kinase signaling. *Am J Hum Genet.* (2013) 93:150–7. doi: 10.1016/j.ajhg.2013.05.023
7. Forbes BE, Blyth AJ, Wit JM. Disorders of IGFs and IGF-1R signaling pathways. *Mol Cell Endocrinol.* (2020) 518:111035. doi: 10.1016/j.mce.2020.111035
8. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med.* (2015) 17:405–24. doi: 10.1038/gim.2015.30
9. Zhang Y, Ji B, Li J, Li Y, Zhang M, Ban B. SHORT syndrome in two Chinese girls: a case report and review of the literature. *Mol Genet Genomic Med.* (2020) 8:e1385. doi: 10.1002/mgg3.1385
10. Hamaguchi T, Hirota Y, Takeuchi T, Nakagawa Y, Matsuoka A, Matsumoto M, et al. Treatment of a case of severe insulin resistance as a result of a *PIK3R1* mutation with a sodium-glucose cotransporter 2 inhibitor. *J Diabetes Investig.* (2018) 9:1224–7. doi: 10.1111/jdi.12825
11. Masunaga Y, Fujisawa Y, Muramatsu M, Ono H, Inoue T, Fukami M, et al. Insulin resistant diabetes mellitus in SHORT syndrome: case report and literature review. *Endocr J.* (2021) 68:111–7. doi: 10.1507/endocrj.EJ20-0291
12. Bárcena C, Quesada V, De Sandre-Giovannoli A, Puente DA, Fernández-Toral J, Sigaudy S, et al. Exome sequencing identifies a novel mutation in *PIK3R1* as the cause of SHORT syndrome. *BMC Med Genet.* (2014) 15:51. doi: 10.1186/1471-2350-15-51
13. Busiah K, Drunat S, Vaivre-Douret L, Bonnefond A, Simon A, Flechtner I, et al. Neuropsychological dysfunction and developmental defects associated with genetic changes in infants with neonatal diabetes mellitus: a prospective cohort study. *Lancet Diabetes Endocrinol.* (2013) 1:199–207. doi: 10.1016/S2213-8587(13)70059-7
14. Rubio-Cabezas O, Ellard S. Diabetes mellitus in neonates and infants: genetic heterogeneity, clinical approach to diagnosis, and therapeutic options. *Horm Res Paediatr.* (2013) 80:137–46. doi: 10.1159/000354219
15. Lewandowski KC, Dabrowska K, Brzozowska M, Kawalec J, Lewiński A. Metformin paradoxically worsens insulin resistance in SHORT syndrome. *Diabetol Metab Syndr.* (2019) 11:81. doi: 10.1186/s13098-019-0477-z


**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.

*Copyright © 2021 Kim, Kim, Yim, Kim and Jang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*

frontiers in Pediatrics

CASE REPORT
published: 08 July 2021
doi: 10.3389/fped.2021.691524

# Case Report: Wiskott-Aldrich Syndrome Caused by Extremely Skewed X-Chromosome Inactivation in a Chinese Girl

*Xuening Hou*[1†], *Jie Sun*[2†], *Chen Liu*[1] *and Jihong Hao*[1*]

[1] *Department of Clinical Laboratory, The Second Hospital of Hebei Medical University, Shijiazhuang, China,* [2] *Department of Gynecological Ultrasound, The Second Hospital of Hebei Medical University, Shijiazhuang, China*

OPEN ACCESS

**Edited by:**
*Kathleen M. Gorman, Temple Street Children's University Hospital, Ireland*

**Reviewed by:**
*Aline L. Petrin, The University of Iowa, United States*
*Enrique Medina-Acosta, State University of the North Fluminense Darcy Ribeiro, Brazil*

***Correspondence:**
*Jihong Hao*
*haojihong2003@163.com*

[†]*These authors have contributed equally to this work and share first authorship*

**Specialty section:**
*This article was submitted to Genetics of Common and Rare Diseases, a section of the journal Frontiers in Pediatrics*

**Received:** *06 April 2021*
**Accepted:** *15 June 2021*
**Published:** *08 July 2021*

**Citation:**
*Hou X, Sun J, Liu C and Hao J (2021) Case Report: Wiskott-Aldrich Syndrome Caused by Extremely Skewed X-Chromosome Inactivation in a Chinese Girl. Front. Pediatr. 9:691524. doi: 10.3389/fped.2021.691524*

Wiskott-Aldrich syndrome (WAS) is a rare X-linked immunodeficiency disorder caused by abnormal expression of Wiskott-Aldrich syndrome protein due to *WAS* gene mutation, which is generally characterized by microthrombocytopenia, eczema, recurrent infections, and high risk of autoimmune complications and hematological malignancies. Although affected males with WAS usually manifest severe symptoms, female carriers have no significant clinical manifestations. Here, we describe a Chinese girl diagnosed with WAS carrying a heterozygous missense mutation in exon 2 of the *WAS* gene. The patient presented with persistent thrombocytopenia with small platelets and decreased WAS protein detected by flow cytometry and western blot analysis. The methylation analysis of the *HUMARA* gene displayed an extremely skewed X-chromosome inactivation (SXCI) pattern, where the X-chromosomes bearing normal *WAS* gene were predominantly inactivated, leaving the mutant gene active. Hence, our results suggest that completely inactivating the unaffected paternal X-chromosomes may be the reason for such phenotype in this female patient. SXCI has important implications for genetic counseling of female carriers with a family history of WAS.

**Keywords: Wiskott-Aldrich syndrome, microthrombocytopenia, female carrier, heterozygous mutation, X-chromosome inactivation**

## INTRODUCTION

Wiskott-Aldrich syndrome (WAS, OMIM 301000) is a rare X-linked recessive immunodeficiency disorder characterized by thrombocytopenia with small platelets, eczema, and recurrent infections. The clinical phenotype has a wide spectrum, varying from mild thrombocytopenia to severe clinical manifestations, including life-threatening hemorrhages, immunodeficiency, atopy, autoimmunity, and malignancies (1). The *WAS* gene, whose mutations are associated with the clinical phenotype of WAS, is mapped to Xp11.23, and encodes for the Wiskott-Aldrich syndrome protein (WASp). The severity of WAS phenotype in patients correlates with specific mutations and the expression levels of WASp in the blood cells. Usually, affected males show symptoms at an early age, while females carrying the defective *WAS* gene are asymptomatic. Till date, fewer than 20 female cases have been reported with WAS worldwide (**Supplemental Table 1**). Most studies favor the hypothesis that the clinical phenotype of female carriers is attributed to skewed X-chromosome inactivation (XCI),

where the mutated X-chromosomes are preferentially selected to be active (2–6).

Herein, we describe a 9-year-old Chinese girl with WAS who presented with congenital thrombocytopenia and hemorrhagic events. The patient had a heterozygous mutation of the *WAS* gene on the maternally derived X-chromosome, associated with the paternal X-chromosome completely inactive in the peripheral blood cells.

## CASE REPORT

In 2019, a 9-year-old girl was admitted to the hospital because

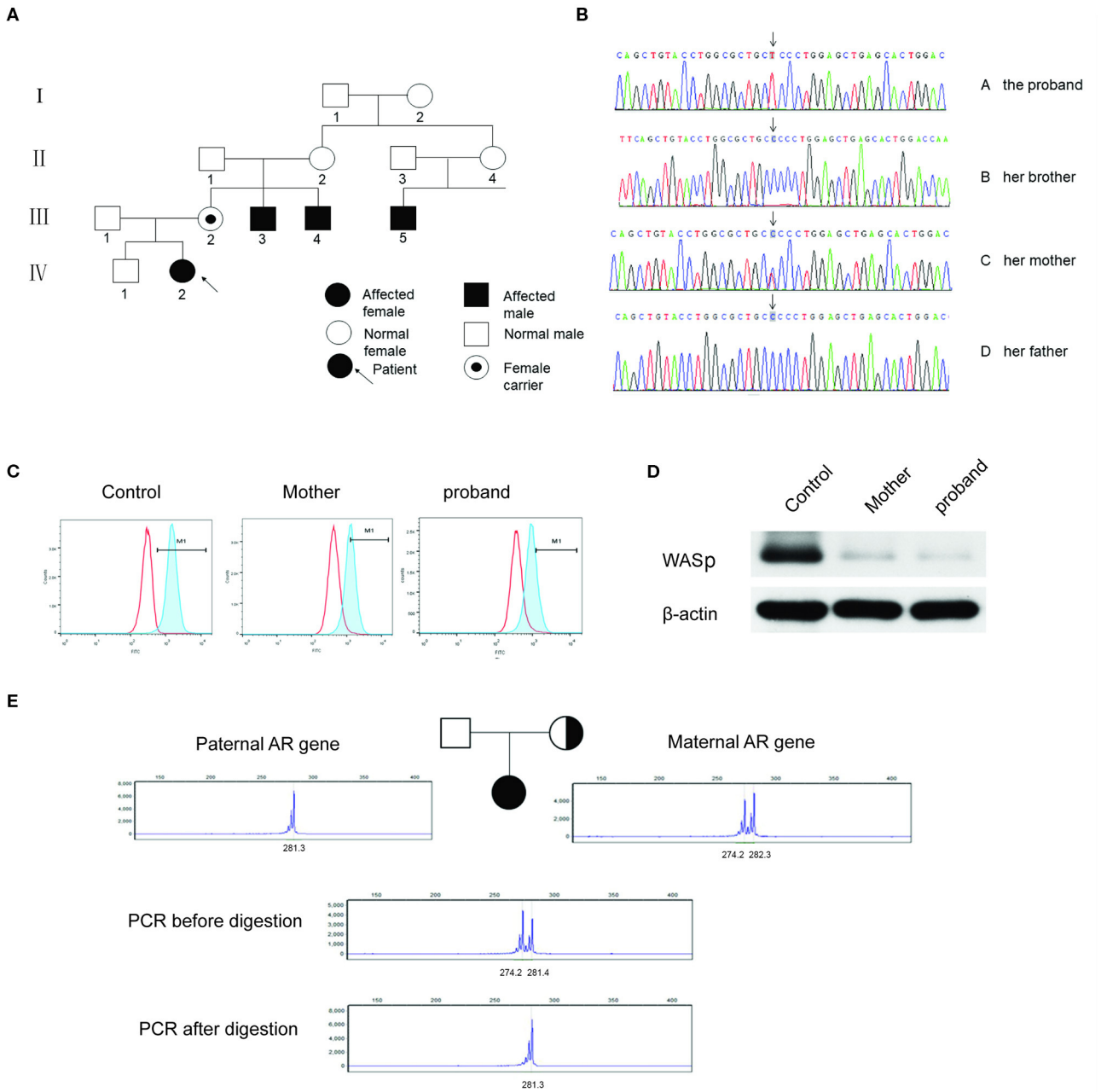



**FIGURE 1 | (A)** The pedigree of the patient's family. III.2: female carrier; III.3, III.4, and III.5: affected male; IV.2: female proband. **(B)** Sanger sequencing of the *WAS* gene of c.173C>T. The proband and her mother were heterozygous of the mutated gene, and her older brother and her father were normal. **(C)** WASp expression levels in the PBMCs by flow cytometry assay. The proportion of M1 area was 98.5% of normal control, 50.6% of her mother and 20.6% of the proband, respectively. **(D)** WASp expression levels were analyzed by western blotting. The gray mean value of the proband's mother is 10.8% of the control level and the gray mean value of proband is 5.4% of the level. **(E)** X-chromosome inactivation analysis of the patient and her parents. The experiments were performed in triplicate.

**TABLE 1 |** Immunologic and hematologic characteristic of the patient and her mother.

| Variable | Patient | The patient's mother | Normal range |
|---|---|---|---|
| **Platelets** | | | |
| Count ($\times 10^9$/L) | 19 | 270 | 125–350 |
| Mean Volume (fL) | 5.9 | 9.1 | 7.4–11.0 |
| **Serum immunoglobins (g/L)** | | | |
| IgG | 3.78 | 11.97 | 7.5–15.0 |
| IgM | 0.14 | 0.52 | 1.0–5.0 |
| IgA | 1.74 | 0.34 | 0.46–3.0 |
| **Lymphocyte phenotype (%)** | | | |
| Total T cells | 77.9 | 66.9 | 20–50 |
| CD3+CD4+ | 48.7 | 38.5 | 33–58 |
| CD3+CD8+ | 26.9 | 21.7 | 20–39 |
| CD3+CD4+/CD3+CD8+ | 1.81 | 1.77 | 0.7–2.5 |
| CD19+/CD20+ | 9.9 | 7.6 | 5–18 |
| CD3-CD56+ | 11.4 | 23.7 | 10–20 |

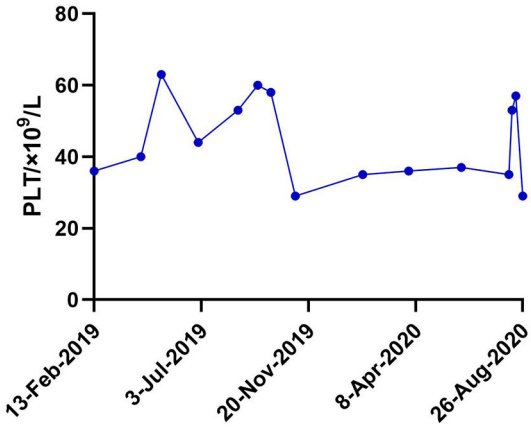



**FIGURE 2 |** Platelet fluctuations of the patient within one and a half years (13-Feb-2019–26-Aug-2020) after admission.

of severe colporrhagia for 3 days. She was initially admitted to the hospital due to thrombocytopenia at the age of 4 months and was diagnosed with immune thrombocytopenia (ITP). However, she failed to recover with hormone and human immunoglobulin injections, and her platelet count fluctuated between 20 $\times$ $10^9$/L and 30 $\times$ $10^9$/L. She then experienced frequent fever, followed by recurrent administration of anti-infective drugs due to upper respiratory tract infection. One and a half months ago, she was admitted to a local hospital for respiratory infection before being discharged with improved condition through anti-infection treatment.

Physical examination revealed petechiae on both eyelids and upper limbs, bulbous conjunctival hemorrhage and pharyngeal hyperemia, and tonsil I° swelling, although no eczema was noted. Peripheral blood cell count demonstrated 16.91 $\times$ $10^9$/L (3.5–9.5 $\times$ $10^9$/L) WBC, 139 g/L (120–140 g/L) hemoglobin, and 19 $\times$ $10^9$/L (125–350 $\times$ $10^9$/L) platelet with the low mean platelet volume (MPV) of 5.90 fL (7.4–11.0 fL). Small platelets were found in the peripheral blood smears.

Her serum IgG and IgM levels were lower than normal (3.78 and 0.14 g/L, respectively). Mycoplasma pneumoniae antibodies were positive, with the titer of 1:160, and without platelet-associated antibodies. Analysis of the lymphocyte subsets showed elevated total T cell count without any immunologic abnormalities. Bone marrow aspirate exhibited delayed maturation of megakaryocytes. Combined with her medical history and ineffective hormonal treatment, she was suspected of having inherited platelet disorders.

Upon investigating her family history, her father and elder brother showed no symptoms (**Figure 1A**), while her mother showed decreased IgM and IgA levels, with normal PLT. The detailed results are shown in **Table 1**. Her three maternal uncles presented thrombocytopenia for many years.

Sanger sequencing revealed a heterozygous missense mutation in exon 2 of the *WAS* gene, leading to a change from proline to leucine at position 58 in the patient and her mother (**Figure 1B**). No mutations were detected in her father or elder brother. Meanwhile, flow cytometry showed the level of WASp from the proband and in the cells from her mother at 20.6 and 50.4%, respectively (**Figure 1C**). The ratio of WASp was reduced by ∼5-fold in mononuclear cells from the proband compared to the normal control (98.5%). Western blot analysis revealed significantly decreased WASp levels in the patient and her mother compared to normal control (**Figure 1D**). Based on the above findings, the girl was confirmed to have WAS. Her platelet count fluctuated between 30$\times$$10^9$/L and 60$\times$$10^9$/L within one and a half years after admission (**Figure 2**). At present, the patient is receiving symptomatic treatment and awaiting hematopoietic stem cell transplantation.

Moreover, XCI analysis was performed using the human androgen receptor assay (HUMARA). Briefly, genomic DNA was extracted from peripheral blood leukocytes and used to amplify the highly polymorphic CAG repeat in the first exon of the *HUMARA* gene on Xq11-q12 (4). The degree of X-inactivation was calculated from the peak height. The results showed that the proband had only one paternal allele amplified after *HpaII* digestion (**Figure 1E**). In contrast, her mother showed two bands in the presence of *HpaII*; however, the rate of random inactivation indicated a mildly skewed pattern of XCI in mature blood cells (**Figure 1E**). The rate of XCI was 100 and 62.7% for the patient and her mother, respectively (**Table 2**).

## DISCUSSION

Herein, we present a Chinese girl with a congenital thrombocytopenia, hemorrhagic tendency, and recurrent infections with c.173C>T, a missense mutation of the *WAS* gene. Although missense mutations are most common in patients with WAS, this mutation is so rare that only three cases have been reported in the USA, China, and Brazil (7–9). Unlike the three male cases, our patient was a girl with this mutation. The c.173C>T mutation reported here is in the second exon that falls in the WH1 domain, which binds to the site of WASp interacting

**TABLE 2 |** The results of X-chromosome inactivation of the patient and her mother.

| | PCR fragment location (-) | PCR fragment location (+) | Peak 1 height (+) | Peak 2 height (+) | Peak 1 height (-) | Peak 2 height (-) | Rate of random inactivation | The proportional relation of random inactivation |
|---|---|---|---|---|---|---|---|---|
| IV2 | n1 274.2 | / | 0 | 1,301 | 580 | 770 | 100% | 100:0 |
| | n2 281.4 | 281.3 | | | | | | |
| III2 | n1 274.2 | 274.1 | 969 | 824 | 1,173 | 1,686 | 62.8% | 63:37 |
| | n2 282.3 | 282.1 | | | | | | |
| III1 | n1 281.3 | | | | | | | |

*-, before digestion by HpaII; +, after digestion by HpaII; n1/n2, the locations of two alleles; IV2 means female proband, and III2 means the proband's mother, a female carrier and III1 means the proband's father; the length of the corresponding allele peaks in the proband represents the paternal allele (281nt) and maternal allele (274nt). The degree of random inactivation: 50:50∼59:41: complete random X-chromosome inactivation; 65:35∼79:21: mild skewed X-chromosome inactivation; >80:20: serious skewed X-chromosome inactivation.*

protein (WIP), crucially influencing the stability of WASp. Moreover, this domain plays a critical role in a series of events, including intracellular signal transduction, protein folding, and transport (10).

The *WAS* gene located on the X-chromosome contributed to WAS, which is transmitted in compliance with an X-linked recessive inheritance feature. X-linked female carriers show no clinical manifestations due to the pattern of XCI, which may ameliorate the destructive effects of the X-linked mutations (11). Nevertheless, XCI maintenance mechanisms are remarkably diverse, which closely affects X-linked gene dosage and sex differences in X-linked diseases (12). Therefore, the pattern of XCI may aggravate the clinical manifestations of the disease. Our present study proposed that skewed XCI led to the clinical phenotype of this female patient with WAS. In some cases, this is possibly due to genetic defects in the X-inactivation transcriptional process, contributing to the unusually skewed XCI pattern (4). The imbalanced expression of the parental X-chromosomes seems to be common in the general population; however, the skewed XCI pattern increased the risk of recessive X-linked disorders in females (13).

The results of XCI analysis showed two peaks at different positions in the proband and her mother before digestion with *HpaII*, indicating that the alleles were derived from both paternal and maternal chromosomes. However, the proband had only one allele left after enzymatic digestion, demonstrating that the proband had preferential inactivation of the paternally derived wild-type allele. The selection of XCI was independent of the mutation originating from either parent. While the spontaneous heterozygous mutation of the *WAS* gene was on the paternally derived X-chromosome, the extremely skewed XCI preferentially selected the maternally derived wild-type X-chromosome to be inactivated (5). From a comprehensive perspective, the genetic basis of the clinical presentation of WAS in our proband is the fortuitous combined occurrence of a maternally inherited mutation and the extreme skewing of inactivation against the unaffected paternal chromosome.

The proportion of inactivation was 100% in the patient and 62.7% in her mother (**Table 2**), manifesting the overwhelmingly skewed XCI of the patient and slight inactivation of her mother. These results were consistent with the expression levels of WASp. At the cytogenetic level, it also explained the differences in clinical signs between the mother and daughter, who were both carriers of heterozygous mutation.

Based on WASp-deficient mice, the skewed inactivation of WAS carriers was attributed to homing deficits in fetal hematopoietic stem cells (14). In addition, the selected inactivation can be present at a later stage, even in mature populations of hematopoietic differentiation (15). Thus, the predominant selection against mutated alleles as well as skewed inactivation can occur throughout the progress of hematopoietic development, increasing the risk of WAS in females. Diverse patterns of XCI exist in female carriers, such as random, slightly skewed, or extremely skewed patterns (3, 16). As a heterozygous carrier, the patient's mother exhibited slightly skewed inactivation in our study. This illustrates that the XCI pattern in humans shows diversity between individuals without stochastic mechanisms. In addition, a previous study indicated that the blood cell population bearing the mutated *WAS* message was greater in patients with WAS than in their mothers (3). Hence, the degree of imbalanced XCI in WAS carriers may vary with age. A WAS carrier has been reported to present random XCI in granulocytes and non-random inactivation in T lymphocytes (15). In contrast, a female with X-linked thrombocytopenia was also reported with random inactivation in lymphocytes along with skewed XCI in buccal mucosal cells (17). Therefore, the XCI pattern also varies at the cellular level and across different cell types.

In summary, the cumulative clinical results and molecular assays illustrated that a female carrier with a heterozygous *WAS* mutation could express an X-linked WAS disease. To the best of our knowledge, the case reported here is the first Chinese female patient presenting WAS clinical symptoms with an extremely skewed XCI pattern, in which the mutated allele was preferentially active. XCI analysis is a relatively effective and rapid technique to analyze the pattern of XCI in female carriers, who are suspected of having an X-linked recessive disorder. Furthermore, this case reminds us of the fact that female carriers with a family history of WAS may also develop certain clinical symptoms.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Research Ethics Committee of the Second Hospital of Hebei Medical University. 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

All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission. XH drafted the manuscript. JS and CL contributed to the design, acquisition, and analysis of data. JH final approval of the version to be published and offered professional guidance.

## FUNDING

This study was supported by grants from the Natural Science Foundation of Hebei Province of China (H2020206349) and the Research Fund of the Second Hospital of Hebei Medical University of China (2h2019041).

## ACKNOWLEDGMENTS

We would like to thank all members of the study team, the patient and her family.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fped.2021.691524/full#supplementary-material

## REFERENCES

1. Candotti F. Clinical manifestations and pathophysiological mechanisms of the Wiskott-Aldrich syndrome. *J Clin Immunol.* (2018) 38:13–27. doi: 10.1007/s10875-017-0453-z
2. Parolini O, Ressmann G, Haas OA, Pawlowsky J, Gadner H, Knapp W, et al. X-linked Wiskott-Aldrich syndrome in a girl. *N Engl J Med.* (1998) 338:291–5. doi: 10.1056/NEJM199801293380504
3. Ariga T, Yamada M, Wada T, Saitoh S, Sakiyama Y. Detection of lymphocytes and granulocytes expressing the mutant WASP message in carriers of Wiskott-Aldrich syndrome. *Br J Haematol.* (1999) 104:893–900. doi: 10.1046/j.1365-2141.1999.01248.x
4. Andreu N, Pujol-Moix N, Martinez-Lostao L, Oset M, Muñiz-Diaz E, Estivill X, et al. Wiskott-Aldrich syndrome in a female with skewed X-chromosome inactivation. *Blood Cells Mol Dis.* (2003) 31:332–7. doi: 10.1016/s1079-9796(03)00168-2
5. Boonyawat B, Dhanraj S, Al Abbas F, Zlateska B, Grunenbaum E, Roifman CM, et al. Combined *de-novo* mutation and non-random X-chromosome inactivation causing Wiskott-Aldrich syndrome in a female with thrombocytopenia. *J Clin Immunol.* (2013) 33:1150–5. doi: 10.1007/s10875-013-9927-9
6. Takimoto T, Takada H, Ishimura M, Kirino M, Hata K, Ohara O, et al. Wiskott-Aldrich syndrome in a girl caused by heterozygous WASP mutation and extremely skewed X-chromosome inactivation: a novel association with maternal uniparental isodisomy 6. *Neonatology.* (2015) 107:185–90. doi: 10.1159/000370059
7. Kwan SP, Hagemann TL, Radtke BE, Blaese RM, Rosen FS. Identification of mutations in the Wiskott-Aldrich syndrome gene and characterization of a polymorphic dinucleotide repeat at DXS6940, adjacent to the disease gene. *Proc Natl Acad Sci USA.* (1995) 92:4706–10. doi: 10.1073/pnas.92.10.4706
8. Zhang J, Shi J, Li X, Shao Y, Liu C, Ge M, et al. The gene mutation analysis of a Wiskott-Aldrich syndrome family with normal mean platelet volume. *Zhonghua Xue Ye Xue Za Zhi.* (2015) 36:754–8. doi: 10.3760/cma.j.issn.0253-2727.2015.09.007
9. Medina SS, Siqueira LH, Colella MP, Yamaguti-Hayakawa GG, Duarte BKL, Dos Santos Vilela MM, et al. Intermittent low platelet counts hampering diagnosis of X-linked thrombocytopenia in children: report of two unrelated cases and a novel mutation in the gene coding for the Wiskott-Aldrich syndrome protein. *BMC Pediatr.* (2017) 17:151. doi: 10.1186/s12887-017-0897-6
10. Matalon O, Reicher B, Barda-Saad M. Wiskott-Aldrich syndrome protein–dynamic regulation of actin homeostasis: from activation through function and signal termination in T lymphocytes. *Immunol Rev.* (2013) 256:10–29. doi: 10.1111/imr.12112
11. Migeon BR. The role of X inactivation and cellular mosaicism in women's health and sex-specific diseases. *JAMA.* (2006) 295:1428–33. doi: 10.1001/jama.295.12.1428
12. Carrel L, Willard HF. X-inactivation profile reveals extensive variability in X-linked gene expression in females. *Nature.* (2005) 434:400–4. doi: 10.1038/nature03479
13. Shvetsova E, Sofronova A, Monajemi R, Gagalova K, Draisma HHM, White SJ, et al. Skewed X-inactivation is common in the general female population. *Eur J Hum Genet.* (2019) 27:455–65. doi: 10.1038/s41431-018-0291-3
14. Lacout C, Haddad E, Sabri S, Svinarchouk F, Garçon L, Capron C, et al. A defect in hematopoietic stem cell migration explains the nonrandom X-chromosome inactivation in carriers of Wiskott-Aldrich syndrome. *Blood.* (2003) 102:1282–9. doi: 10.1182/blood-2002-07-2099
15. Notarangelo LD, Parolini O, Porta F, Locatelli F, Lanfranchi A, Marconi M, et al. Analysis of X-chromosome inactivation and presumptive expression of the Wiskott-Aldrich syndrome (WAS) gene in hematopoietic cell lineages of a thrombocytopenic carrier female of WAS. *Hum Genet.* (1991) 88:237–41. doi: 10.1007/BF00206081
16. de Saint Basile G, Lagelouse RD, Lambert N, Schwarz K, Le Mareck B, Odent S, et al. Isolated X-linked thrombocytopenia in two unrelated families is associated with point mutations in the Wiskott-Aldrich syndrome protein gene. *J Pediatr.* (1996) 129:56–62. doi: 10.1016/s0022-3476(96)70190-7
17. Inoue H, Kurosawa H, Nonoyama S, Imai K, Kumazaki H, Matsunaga T, et al. X-linked thrombocytopenia in a girl. *Br J Haematol.* (2002) 118:1163–5. doi: 10.1046/j.1365-2141.2002.03740.x

**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.

*Copyright © 2021 Hou, Sun, Liu and Hao. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*

frontiers in Genetics

ORIGINAL RESEARCH
published: 22 July 2021
doi: 10.3389/fgene.2021.627204

# The Added Value of Whole-Exome Sequencing for Anomalous Fetuses With Detailed Prenatal Ultrasound and Postnatal Phenotype

*Miao He[1†], Liu Du[1†], Hongning Xie[1*], Lihe Zhang[1], Yujun Gu[1], Ting Lei[1], Ju Zheng[1] and Dan Chen[2]*

[1] *Department of Ultrasonic Medicine, Fetal Medical Center, The First Affiliated Hospital of Sun Yat-sen University, Guangzhou, China,* [2] *Guangzhou Kingmed Diagnostics Group, Guangzhou, China*

OPEN ACCESS

***Edited by:***
*Kathleen M. Gorman, Temple Street Children's University Hospital, Ireland*

***Reviewed by:***
*Antoni Borrell, University of Barcelona, Spain*
*Fan Jin, Zhejiang University, China*
*Khalda Amr, National Research Centre, Egypt*

****Correspondence:***
*Hongning Xie*
*xiehn@mail.sysu.edu.cn*

*† These authors have contributed equally to this work*

***Specialty section:***
*This article was submitted to Genetics of Common and Rare Diseases, a section of the journal Frontiers in Genetics*

***Received:*** *08 November 2020*
***Accepted:*** *10 June 2021*
***Published:*** *22 July 2021*

***Citation:***
*He M, Du L, Xie H, Zhang L, Gu Y, Lei T, Zheng J and Chen D (2021) The Added Value of Whole-Exome Sequencing for Anomalous Fetuses With Detailed Prenatal Ultrasound and Postnatal Phenotype. Front. Genet. 12:627204. doi: 10.3389/fgene.2021.627204*

**Objectives:** The objective of the study was to explore the added value of whole-exome sequencing (WES) in abnormal fetuses with detailed prenatal ultrasound and postnatal phenotype with normal karyotype and chromosomal microarray analysis (CMA).

**Methods:** Parents of fetuses with structural abnormalities by prenatal ultrasound who consented to provide fetal samples were prospectively recruited from January 2017 to December 2019. With aneuploidies or cases with copy number variations (CNVs) excluded, WES was performed for cases with normal karyotype and CMA results. Detailed prenatal ultrasound and postnatal imaging or pathology features were recommended for further interpretation of genetic variants.

**Results:** WES was performed for 94 eligible fetuses, DNA samples of which were extracted from 53 parent–fetus trios and 41 proband-only fetal tissues. A diagnostic genetic variant was identified in 37 (39.4%) of 94 fetuses, and 34 (64.2%) were detected in 53 trios, which was significantly greater than 3 (7.3%) in 41 proband-only cases ($p < 0.001$). In 34 trios with diagnostic genetic variants, 23 (67.6%) were *de novo* and 11 (32.4%) were inherited with two homozygous and nine heterozygous variants. Fourteen (14.9%) of 94 fetuses had a variant of uncertain significance (VUS). Among 94 cases, six affected pregnancies continued and 88 terminated, and 57 of 88 terminated cases underwent postmortem examinations. With accurate phenotypes demonstrated by prenatal ultrasound and postnatal autopsies, the clinical phenotypes were correlated in 33 (89.2%) of 37 cases with specific genotypes, with the highest matching ratio in skeletal diseases (20/33, 60.6%).

**Conclusion:** WES has added value in the genetic diagnosis of abnormal fetuses with normal karyotypes and CMA, particularly in skeletal diseases. Using WES in various anomalous fetuses can broaden the understanding of prenatal phenotypes and genetic variants.

**Keywords: whole-exome sequencing, anomalous fetuses, prenatal ultrasound, postnatal phenotype, genetic variants**

# INTRODUCTION

Congenital anomalies affect 2–4% of all infants and are responsible for 21% of perinatal deaths (Ely and Driscoll, 2019). Using ultrasound, these abnormalities can be identified in the first and second trimesters, ranging from fetal hydrops to major lethal disorders. Although the standard genetic investigation of fetal abnormalities, karyotype, and/or chromosomal microarray analysis (CMA) is applied for aneuploidies and copy number variations (CNVs), greater than 60% of pregnancies do not have a genomic diagnosis to guide future care and genetic counseling (Petrovski et al., 2019).

Prenatal whole-exome sequencing (WES) is a useful and valuable tool for the genetic diagnosis of fetal anomalies with a diagnostic yield between 6.2 and 81% (Carss et al., 2014; Alamillo et al., 2015; Drury et al., 2015; Yates et al., 2017; Chandler et al., 2018; Yadava and Ashkinadze, 2019). The underlying etiology has been identified in many fetal abnormalities by the use of WES, including anomalies of the kidney and urinary tract (13%) (Lei et al., 2017), skeletal disease (88.9%) (Han et al., 2020), and non-immune hydrops fetalis (29%) (Sparks et al., 2020); however, the limited prenatal phenotype information, difficulties in variant interpretation of secondary findings, and the time-consuming nature of this costly technology have put off the widespread clinical implementation of WES. To expand the prenatal phenotype–genotype spectrum and facilitate the utility of prenatal WES, we analyze the diagnostic yield of WES in fetuses with different kinds of abnormalities with detailed ultrasound and postnatal phenotypes.

# MATERIALS AND METHODS

## Participants and Sample Collection

From January 2017 to December 2019, parents of fetuses with structural anomalies detected by prenatal ultrasound were invited to participate in this study at The First Affiliated Hospital of Sun Yat-sen University. Prenatal ultrasound, including a detailed ultrasound scan and the assessment of fetal biometry, was performed by an expert with 30 years of experience in the obstetrics ultrasound field. In the cases of termination of pregnancy (TOP), postmortem examinations were performed as much as possible after obtaining permission of parents to access detailed and accurate clinical phenotypes. When autopsy findings or phenotypes differed from those of prenatal ultrasound, clinical manifestations were corrected to optimize the interpretation of genetic variants. In this study, singleton pregnancies were required, and cases with available fetal samples were included. A detailed pre-test counseling was offered for these parents. Cases with anomalies in the first trimester and fetuses with aneuploidies or CNVs were excluded. Fetuses with a known family history of genetic mutation or a known infection or exposure to a known teratogenic drug were excluded.

After prenatal ultrasound diagnosis and pre-test counseling, if parents wanted to continue the pregnancy, amniocentesis or cord blood sampling was collected for genetic testing according to the fetal gestational age at the time of invasive procedures. Karyotype and CMA were performed first, and then WES was carried out for cases with normal karyotypes and CMA by using the same sample. In the cases of fetal demise or pregnancy termination, fetal tissue samples were collected for testing after termination. Parents were encouraged to provide blood samples for subsequent testing to assist in the interpretation of genetic variants. Samples of the abnormal fetus, mother, and father all collected in one case, were termed the parent–fetus trios. Informed consent was obtained from all parents, and they were informed that only results related to clinical phenotypes detected by ultrasound and autopsy would be reported back, not secondary findings. The turnaround time was 6–8 weeks, including reports of trios. This study was approved by the Ethics Committee of The First Affiliated Hospital of Sun Yat-sen University.

## Whole-Exome Sequencing

Genomic DNA was extracted with a QIAamp DNA Blood Mini kit (Qiagen, Hilden, Germany) following the manufacturer's protocol. The extracted DNA was fragmented randomly and then purified using the magnetic particle method. DNA fragments were ligated with adaptors and captured by probes of the IDT XGen Exome Research Panel (IDT, Lowa, United States) targeting approximately 19,396 genes. The DNA libraries, after enrichment and purification, were sequenced on a NovaSeq 6000 sequencer according to the manufacturer's instructions (Illumina, San Diego, CA, United States). All reads were aligned to the reference human genome (UCSC hg19) by Burrows–Wheeler Aligner (BWA) (v.0.5.9-r16) (Li and Durbin, 2010). After data annotation using the PriVar toolkit (Zhang et al., 2013), the clinical significance of the variants was identified (Yang et al., 2013). Databases such as OMIM[1], GeneCards[2], ClinVar[3], and Human Gene Mutation Database[4] were used to determine pathogenicity. The detected variants were confirmed using PCR, and the products were subjected to direct sequencing on a 3500XL Genetic Analyzer (Applied Biosystems, Foster City, CA, United States) according to the manufacturer's instructions. Each variant was categorized by the performing laboratory as pathogenic, likely pathogenic, variants of uncertain significance (VUS), likely benign, or benign according to the American College of Medical Genetics and Genomics (ACMG) guidelines (Richards et al., 2015).

## Statistical Analysis

The number of pathogenic and likely pathogenic variants in the trios and probands and the number of diagnostic variants in phenotype classes were compared with the chi-square test. A value of $p < 0.05$ was used to define statistical significance, and

[1]http://www.omim.org

[2]https://www.genecards.org

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

[4]http://www.hgmd.org

the tests were two-tailed. The statistical analysis was performed with SPSS (version 25.0).

# RESULTS

## Demographic Characteristics

From January 2017 to December 2019, 102 couples with fetuses that had structural anomalies consented to participate in our study. Eight cases with pathogenic CNVs were excluded, and 94 eligible cases with normal karyotyping and CMA results underwent WES. The median maternal age was 31 years (range 22–42 years). Gestational age (GA) at ultrasound screening ranged from 16 to 37 weeks (median GA 25 + 1 weeks). Six couples continued the pregnancies, and 88 chose to terminate. Postmortem examinations were done for 57 (57/88, 64.8%) cases. The postmortem phenotypes of four cases with different types of genetic results are shown in **Figure 1**. The clinical outcomes of the remaining 31 cases were known through telephone tracts without postmortem examination.

## Phenotypes

Accurate phenotypic information was obtained from prenatal ultrasound and postnatal necropsy/postnatal imaging [fetal magnetic resonance imaging (MRI) and X-ray after termination]. Of the 57 autopsy cases, the diagnostic correction was only made in case 89 with unilateral anophthalmia, and unilateral microphthalmia was confirmed after postmortem examination. Case 80 with asymmetric brain dysplasia was detailed as lissencephaly of the right cerebral hemisphere by pathological analysis. The postnatal phenotypes of the remaining cases were consistent with prenatal ultrasound phenotypes. The 94 cases were categorized into 11 classes according to the anatomical systems involved (**Table 1**). There were 60 cases (60/94, 63.8%) that affected one anatomic system and 34 cases (34/94, 36.2%) with anomalies that affected two or more anatomic systems. The phenotypes included various kinds of abnormalities as shown in examples of specific findings in **Table 1**.

## Genetic Variants

In our prospective cohort, DNA samples were obtained from 76 fetal tissues, 12 amniocenteses, and 6 cord blood samples. The diagnostic yield of WES was 39.4% (37/94) with likely pathogenic (19/94, 20.2%) and pathogenic variants (18/94, 19.1%). VUS was found in 14 cases (14/94, 14.9%) with potential clinical usefulness. Of the 94 fetuses who underwent WES, 53 were trios, and 41 were proband-only cases. The diagnostic yield of pathogenic variants in trios (34/53, 64.2%) was significantly higher than that in probands (3/41, 7.3%) ($p < 0.001$). In 34 trios with diagnostic genetic variants, 23 (23/34, 67.6%) were *de novo*, and 11 (11/34, 32.4%) were inherited with two homozygous and nine heterozygous variants. Maternal carriers were found in four cases, paternal carriers were found in two cases, and carriers of both parents were found in five cases. The ultrasonic phenotype and diagnostic genetic variants of 37 cases are shown in **Table 2**. Of the 37 cases with diagnostic genetic variants, the clinical phenotypes were correlated with genotypes in 33 cases (89.2%) with the highest matching ratio in skeletal diseases (20/33, 60.6%), such as *FGFR3* in the case with short bones considering achondroplasia, *FGFR2* in the case with turricephaly, brachy-syndactyly of hands and feet considering Apert syndrome, and *COL1A1* or *COL1A2* in the case with short and angulated bones considering osteogenesis imperfecta. The remaining four cases had incidental findings that could expand the phenotype–genotype spectrum, including cardiac anomalies with *ANKRD11*, central nervous system anomalies with *FGFR2*, multisystem anomalies with *C5orf42*, and genitourinary anomalies with *ETFDH*.

In 94 cases, the diagnostic yield of cases with anomalies affecting one anatomic system was 40% (24/60), and that was 38.2% (13/34) in cases with anomalies involved in two or more anatomic systems. The diagnostic yield of cases with fetal abnormalities affecting one anatomic system was higher than that involved in two or more anatomic systems, while there was no significant difference ($p = 0.87$). Of the 60 cases affected with one anatomic system, 19 cases had skeletal abnormalities, with the highest diagnostic yield of 68.4% (13/19). Cardiac abnormalities

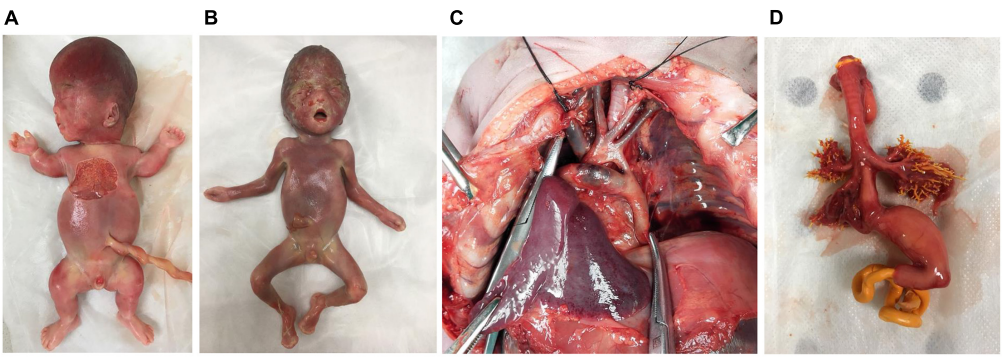



**FIGURE 1 |** The postmortem phenotypes of four cases. **(A)** Case 16. Thanatophoric dysplasia (type I), fetus with frontal bossing, thoracic hypoplasia, and shortened bone (<-6 to 8 SD) with pathogenic missense variant in *FGFR3*. **(B)** Case 29. Apert syndrome, fetus with depressed nasal bridge, turricephaly, and brachy-syndactyly of hands and feet with pathogenic missense variant in *FGFR2*. **(C)** Case 43. Coarctation of the aorta, fetus with coarctation of the aorta with variant of uncertain significance (*CITED2*). **(D)** Case 10. Esophago-tracheal fistula, fetus with persistent left superior vena cava, esophago-tracheal fistula and polyhydramnios without diagnostic genetic variant.

**TABLE 1 |** The frequency of fetal anomalies and examples of specific findings.

| Category of anomaly | Numbers of fetuses (*n* = 94) | Examples of specific findings |
|---|---|---|
| Skeletal system | 36 (38.3%) | Shortened and bowing bone, dysplasia of thoracic vertebra and ribs, hemivertebra, scoliosis, syndactyly |
| Cardiovascular system | 26 (27.7%) | Tetralogy of Fallot, interrupted aortic arch, transposition of the great arteries, pulmonary artery sling |
| Central nervous system | 21 (22.3%) | Agenesis of corpus callosum, Dandy–Walker malformation, pachygyria, cerebral dysplasia, hydrocephalus |
| Genitourinary system | 19 (20.2%) | Enlarged polycystic and echogenic kidneys, hypospadias, common cloacal deformity |
| Neuromuscular system | 14 (14.9%) | Clubfeet, arthrogryposis |
| Head | 12 (12.8%) | Micrognathia, bilateral anophthalmia, unilateral microphthalmia, depressed nasal bridge |
| Hydrops | 5 (5.3%) | Hydrops, pleural effusion |
| Gastrointestinal system | 4 (4.3%) | Esophago-tracheal fistula, mesenteric cyst, situs in vs. |
| Growth abnormality | 3 (3.2%) | Intrauterine growth restriction |
| Abdominal wall | 2 (2.1%) | Omphalocele, exstrophy of bladder |
| Respiratory system | 1 (1.1%) | Dysplasia of right lung |

*Fetuses with several anomalies involving different anatomic systems were counted several times.*

were detected in 12 cases. Isolated cardiac abnormalities, such as transposition of the great arteries (case 8), double outlet left ventricle, pulmonary valve dysplasia (case 23), and common arterial trunk (case 82), did not reveal pathogenic variants. Only one case (case 3) with a ventricular septal defect and interrupted aortic arch (type B) detected the pathogenic variant *ANKRD11*. There were 11 cases with genitourinary abnormalities. Two cases with polycystic and echogenic kidneys had variants in *ETFDH* (case 53) and *SALL1* (case 90), while another two cases with similar phenotypes had a biparentally inherited missense variant in *PKD1* of uncertain significance. Two of three hypospadias cases had pathogenic variants in *SRD5A2* (case 51) and *AR* (case 94). No pathogenic variant was detected in the severe case with common cloacal deformity. Six cases involved the central nervous system. Only one fetus with pachygyria had a non-sense variant in *ASPM* (case 70). The remaining cases, including bilateral ventriculomegaly and Dandy–Walker malformation (case 11), lissencephaly (case 39), and pachygyria (case 81), did not show detectable pathogenic variants. Of the six cases with head anomalies, three variants were found: two had abnormalities in the eyes with bilateral anophthalmia in *SOX2* (case 83) and persistent hyperplastic primary vitreous in *NDP* (case 92), and one had micrognathia in *BMP2* (case 36). Another three cases with microphthalmia or microtia did not show pathogenic variants. Four cases with arthrogryposis were screened and categorized as neuromuscular abnormalities, and the diagnostic yield was 50% (2/4) with variants in *KLHL40* (case 26) and *ZC4H2* (case 91). One hydrops fetalis and one case with an anomaly of the abdominal wall did not show a pathogenic variant.

## DISCUSSION

In our study, the diagnostic yield of WES in anomalous fetuses with normal karyotype and CMA results was 39.4%. A genetic diagnosis was nearly nine times more frequent in trios than in proband-only cases. In fetuses with abnormalities in two or more anatomic systems, the diagnostic yield of cases was 38.2% compared with 40% if a single anatomic system was involved. Among the cases involving one anatomic system, the skeletal and neuromuscular systems and head anomalies had higher diagnostic yields.

The strength of our study was that it included various kinds of fetal abnormalities and added or compared the postnatal phenotypes with prenatal phenotypes for the genetic interpretation. The diagnostic yield in our study was higher than that in other studies. Lord et al. (2019) found that 8.5% of pathogenic variants were detected by WES in unselected fetuses with structural abnormalities, including increased nuchal translucency (NT). Fu et al. (2018) and Petrovski et al. (2019), respectively, showed 24% and 10.3% diagnostic yields by WES in fetuses with structural abnormalities with normal karyotype or CMA results. In highly selected cases, such as terminated or miscarried fetuses (Sa et al., 2018) or fetuses with anomalies in the specific anatomic system (Han et al., 2020), the diagnostic rate of pathogenic variants was up to 53.3–88.9%. We hypothesized that the highly selected cases of skeletal dysplasia and fetuses with severe abnormalities led to the higher diagnostic yield of WES, as was shown in our prospective cohort.

Proband-only and trio sequencing were included in our research. Parents decided whether to provide the blood sample after genetic counseling according to their willingness. Both proband-only and trio groups included fetal anomalies in one and two or more anatomic systems, and a higher proportion of cases with abnormalities affecting multisystems was found in the proband-only group. There were 15 cases (15/41, 46.3%) with abnormalities in two or more anatomic systems among the proband-only group, while 19 cases (19/53, 35.8%) were in the trio group; however, the trios had a higher diagnostic yield. The proportion of skeletal abnormalities in the trio group led to a higher rate of genetic variants than that in the proband-only group.

The genetic diagnosis of congenital heart defects and central nervous system defects was challenging because of genetic and phenotypic heterogeneity, while the condition was different in the skeletal system. The fetuses with skeletal anomalies had the highest detection yield in WES. Skeletal dysplasia is a heterogeneous group of disorders consisting of 436 entities and 364 known genes (Bonafe et al., 2015). Osteogenesis imperfecta with short and angulated bones, thanatophoric dysplasia with extremely short bones, and achondroplasia with short bones are the most common prenatal diagnoses (Vora et al., 2020). Despite these disorders, Apert syndrome with turricephaly and brachysyndactyly of hands and feet was also commonly observed in our cohort. These kinds of diseases are associated with the *COL1A1* or *COL1A2*, *FGFR3*, and *FGFR2* variants; however, cases of short bones do not always find exact diagnostic genetic variants. Some cases with short bones (−2 to −3 SD) had variants of uncertain

**TABLE 2 |** Phenotypes and pathogenic variants of 37 cases by WES.

| Case ID | Prenatal imaging phenotype | Gene | Alternation | Consequence | Inheritance | ACMG classification | Origin (inherited/*de novo*) | Novel or previously reported (PMID) | Associated clinical condition |
|---|---|---|---|---|---|---|---|---|---|
| 2 | AMC, micrognathia, polyhydramnios | *MYH3* | c.2015G > A, p. Arg672His | Missense | AD; heterozygous | Pathogenic (PS3 + PM2 + PP2 + PP3 + PP4 + PP5) | *De novo* | AMC (25740846) | AMC |
| 3 | VSD, IAA (type B) | *ANKRD11* | c.576_579dupAGAC, p. Val194fs | Frameshift | AD; heterozygous | Likely pathogenic (PM2 + PM4 + PP3 + PP5) | *De novo* | Novel | KBG syndrome |
| 4 | Shortened and bowing bone (−2−−3 SD) | *COL1A2* | c.2565 + 1G > A, P.? | Splicing mutation | AD; heterozygous | Pathogenic (PVS1 + PS2 + PM4 + PP3 + PP4 + PP5) | *De novo* | Osteogenesis imperfecta (16705691) | Osteogenesis imperfecta |
| 9 | Low nasal bridge, micrognathia, enlarged left ventricular system, ARSA, clenched hands | *KAT6B* | c.5385C > A, p. Tyr1795* | Non-sense | AD; heterozygous | Likely pathogenic (PS2 + PM2 + PM4) | *De novo* | Novel | SBBYSS syndrome |
| 13 | sacral vertebra defects, fusion of vertebra and ribs | *MESP2* | c.718delG, p. Gly240fs | Frameshift | AR, homozygous | Likely pathogenic (PM2 + PM4 + PP3 + PP55) | Parental inherited | Novel | Spondylocostal dysostosis |
| 14 | VSD, IAA (type B), duplex kidney | *USP9X* | c.3502C > T, p. Arg1168* | Non-sense | X-linked; heterozygous | Likely pathogenic (PVS1 + PM2 + PM4) | – | Novel | Intellectual disability |
| 16 | Frontal bossing, thoracic hypoplasia, shortened bone (< −6SD−8SD) | *FGFR3* | c.1118A > G, p. Tyr373Cys | Missense | AD; heterozygous | Pathogenic (PS2 + PM2 + PP2 + PP3 + PP4 + PP5) | *De novo* | Thanatophoric dysplasia type I (8845844) | Thanatophoric dysplasia |
| 17 | Bowing femur bone | *COL1A2* | c.587G > T, p. Gly196Val | Missense | AD; heterozygous | Pathogenic (PS2 + PM2 + PP2 + PP3 + PP4 + PP5) | *De novo* | Osteogenesis imperfecta (23692737) | Osteogenesis imperfecta |
| 19 | Pachygyria, bilateral ventriculomegaly, spina bifida occulta, multiple vertebra hypoplasia, fusion of vertebral arch | *FGFR2* | c.1052C > G, p. Ser351Cys | Missense | AD; heterozygous | Likely pathogenic (PM1 + PM2 + PP3 + PP5) | *De novo* | Crouzon syndrome (8946174) | Crouzon syndrome |
| 20 | Shortened and bowing bone (−2SD−−3SD) | *COL1A1* | c.1751G > A, p. G584E | Missense | AD; heterozygous | Likely pathogenic (PS2 + PM2 + PP3) | *De novo* | Novel | Osteogenesis imperfecta/Ehlers–Danlos syndrome/Caffey disease |
| 22 | Hydrops, AMC | *HBA2* | c.377T > C, p. Leu126Pro | Missense | AD; heterozygous | Likely pathogenic (PM1 + PM2 + PP3 + PP5) | Paternal inherited | HbQS (7070526) | – |
| 24 | Shortened and bowing bone (−2SD−−3SD) | *COL1A1* | c.1669−2A > G | Missense | AD; heterozygous | Pathogenic (PS1 + PM2 + PP2 + PP3 + PP4 + PP5) | Maternal inherited | Osteogenesis imperfecta (16705691) | Osteogenesis imperfecta |
| 26 | AMC, polyhydramnios | *KLHL40; EMD* | c.1516A > C, p. T506P; c.631C > T, p. R211C | Missense; missense | AR; Homozygous; X-linked; hemizygous | Likely pathogenic (PM1 + PM2 + PP3 + PP5); VUS | Parental inherited; maternal inherited; | Nemaline myopathy type 8 (23746549) | Nemaline myopathy |
| 29 | Depressed nasal bridge, turricephaly, brachy-syndactyly of hands and feet | *FGFR2* | c.755C > G, p. Ser252Trp | Missense | AD; heterozygous | Pathogenic (PS1 + PM2 + PP2 + PP3 + PP4 + PP5) | *De novo* | Apert syndrome (7719344) | Apert syndrome |

*(Continued)*

**TABLE 2 |** Continued

| Case ID | Prenatal imaging phenotype | Gene | Alternation | Consequence | Inheritance | ACMG classification | Origin (inherited/*de novo*) | Novel or previously reported (PMID) | Associated clinical condition |
|---|---|---|---|---|---|---|---|---|---|
| 31 | Turricephaly, brachy-syndactyly of hands and feet, ARSA, polyhydramnios | *FGFR2; FLNB* | c.755C > G, p. Ser252Trp; c.1076C > A, p. Ala359Glu | Missense; Missense | AD; heterozygous | Pathogenic (PS1 + PM2 + PP2 + PP3 + PP4 + PP5); VUS | *De novo* | Apert syndrome (7719344); Novel | Apert syndrome |
| 35 | Dysplasia of thoracic vertebra and ribs, scoliosis | *COL10A1* | c.928delA, p. Arg310fs | Frameshift | AD; heterozygous | Likely pathogenic (PM2 + PM4 + PP3 + PP5) | Maternal inherited | Novel | Schmid metaphyseal chondrodysplasia |
| 36 | Micrognathia, polyhydramnios | *BMP2* | c.709C > T, p. Gln237* | Non-sense | AD; heterozygous | Likely pathogenic (PM2 + PM4 + PP3 + PP5) | – | Novel | SSFSC syndrome |
| 41 | Depressed nasal bridge, turricephaly, agenesis of corpus callosum, brachy-syndactyly of hands and feet | *FGFR2* | c.755C > G, p. Ser252Trp | Missense | AD; heterozygous | Pathogenic (PS1 + PM2 + PP2 + PP3 + PP4 + PP5) | *De novo* | Apert syndrome (7719344) | Apert syndrome |
| 42 | Several cardiac rhabdomyoma, cerebral calcifications and subependymal nodules | *TSC2* | c.1257 + 1G > T, p.? | Splicing mutation | AD; heterozygous | Pathogenic (PS1 + PM2 + PP2 + PP3 + PP4 + PP5) | *De novo* | Tuberous sclerosis (15798777) | Tuberous sclerosis |
| 45 | Depressed nasal bridge, turricephaly, ACC, brachy-syndactyly of hands and feet | *FGFR2* | c.755C > G, p. Ser252Trp | Missense | AD; heterozygous | Pathogenic (PS1 + PM2 + PP2 + PP3 + PP4 + PP5) | *De novo* | Apert syndrome (7719344) | Apert syndrome |
| 48 | Thoracic hypoplasia, shortened and bowing bone (< −4−−6 SD) | *COL1A2* | c.1360G > T p. Gly454Cys | Missense | AD, heterozygous | Likely pathogenic (PS2 + PM2 + PP3 + PP5) | *De novo* | Osteogenesis imperfecta (17078022) | Osteogenesis imperfecta |
| 50 | TOF, tracheoesophageal fistula, unilateral renal agenesis, hemivertebra | *C5orf42* | c.4804C > T, p. Arg1602* | Non-sense | AR, heterozygous | Likely pathogenic (PVS1 + PM2) | – | Joubert syndrome (27158779) | Joubert syndrome |
| 51 | Hypospadias | *SRD5A2* | c.680G > A, p. Arg227Gln; c.164T > C, p. leu55Pro | Missense | AR, heterozygous | Pathogenic (PS1 + PM2 + PP2 + PP3 + PP4 + PP5) | Parental inherited | Hypospadias (8768837, 19342739) | Hypospadias |
| 52 | Shortened bone (< −6−−8 SD) | *FGFR3* | c.2419T > A, p.*807Argext*101 | Terminator codon mutation | AD; heterozygous | Pathogenic (PVS1 + PS2 + PM2) | *De novo* | Thanatophoric dysplasia type I (7647778) | Thanatophoric dysplasia |
| 53 | Enlarged polycystic and echogenic kidneys | *ETFDH* | c.1281_1282delAA, p. I428Rfs*6; c.1305T > G, p.Y435* | Frameshift; Non-sense | AR; heterozygous | Likely pathogenic (PVS1 + PM2) | Parental inherited | Novel | |
| 57 | Depressed nasal bridge, opened mouth, abnormal posture of hands and feet | *ABCA12* | c.6858delT, p.F2286Lfs*6; c.3456G > A, p.S1152? | Frameshift; splicing mutation | AR, heterozygous | Pathogenic (PVS1 + PS3) | Parental inherited | Ichthyosis (22992804, 23528209) | Ichthyosis |

*(Continued)*

**TABLE 2 |** Continued

| Case ID | Prenatal imaging phenotype | Gene | Alternation | Consequence | Inheritance | ACMG classification | Origin (inherited/*de novo*) | Novel or previously reported (PMID) | Associated clinical condition |
|---|---|---|---|---|---|---|---|---|---|
| 61 | Shortened and bowing bone (−2−−3 SD) | *COL1A2* | c.1009G > A, p. G337S | Missense | AD; heterozygous | Likely Pathogenic (PM1 + PM2 + PM5 + PP3) | Maternal inherited | Osteogenesis imperfecta (8829649) | Osteogenesis imperfecta |
| 62 | Depressed nasal bridge, turricephaly, unilateral ventriculomegaly, brachy-syndactyly of hands and feet | *FGFR2* | c.755C > G, p. Ser252Trp | Missense | AD; heterozygous | Pathogenic (PS1 + PM2 + PP2 + PP3 + PP4 + PP5) | *De novo* | Apert syndrome (7719344) | Apert syndrome |
| 63 | Shortened and bowing bone (−2−−3 SD) | *COL1A1* | c.779G > A, p. G260D | Missense | AD, heterozygous | Likely pathogenic (PM1 + PM2 + PP3 + PP5) | Paternal inherited | Osteogenesis imperfecta (25741868) | Osteogenesis imperfecta |
| 70 | Pachygyria, FGR | *ASPM* | c.9032G > A, p.W3011*; c.9862G > T, p.E3288* | Non-sense | AR, heterozygous | Likely pathogenic (PM2 + PM4 + PP3 + PP5) | Parental inherited | Microcephalus (15806441, 19332161, 19770472) | Microcephalus |
| 78 | Depressed nasal bridge, turricephaly, dysplasia of septi pellucidi and cerebellar vermis, brachy-syndactyly of hands and feet | *FGFR2* | c.755C > G, p. Ser252Trp | Missense | AD; heterozygous | Pathogenic (PS1 + PM2 + PP2 + PP3 + PP4 + PP5) | *De novo* | Apert syndrome (7719344) | Apert syndrome |
| 79 | Shortened bone (−2−−3 SD), polyhydramnios | *FGFR3* | c.1138G > A, p. Gly380Arg | Missense | AD; heterozygous | Pathogenic (PS2 + PS3 + PM2 + PP1 + PP3) | *De novo* | Achondroplasia (8078586) | Achondroplasia |
| 83 | Bilateral anophthalmia | *SOX2* | c.310G > T, p. Glu104* | Non-sense | AD; heterozygous | Pathogenic (PS2 + PM2 + PP2 + PP3 + PP4 + PP5) | *De novo* | Microphthalmia (18385794) | Microphthalmia |
| 90 | Unilateral polycystic kidney and ureterocele in the bladder | *SALL1* | c.598C > T, p. L200F | Missense | AD, Heterozygous | Likely pathogenic (PS2 + PM2 + PP3) | *De novo* | Townes–Brocks syndrome type I (20520617) | Townes–Brocks syndrome |
| 91 | AMC | *ZC4H2* | c.162delT, p. Leu55fs | Frameshift | X-linked inherited, heterozygous | Pathogenic (PVS1 + PS2 + PM2 + PP2+ PP3 + PP4 + PP5) | *De novo* | Novel | Wieacker–Wolff syndrome |
| 92 | Persistent hyperplastic primary vitreous | *NDP* | c.240_243del, p. Phe81fs | Frameshift | X-linked inherited, hemizygous | Pathogenic (PVS1 + PM2 + PM4 + PP3 + PP5) | *De novo* | Novel | Norrie disease |
| 94 | Hypospadias | *AR** | c.170_172dupTGC, p. Leu57dup | Microduplication mutation | X-linked inherited | Likely pathogenic (PS3 + PM6 + PP5) | *De novo* | Hypospadias (25500996) | Hypospadias |

*ACMG, American College of Medical Genetics and Genomics; AMC, arthrogryposis multiplex congenital; ARSA, aberrant right subclavian artery; ACC, agenesis of corpus callosum; AD, autosomal dominant; AR, autosomal recessive; FGR, fetal growth restriction; IAA, interrupted aortic arch; TOF, Tetralogy of Fallot; WES, whole-exome sequencing; VSD, ventricular septal defect.*
**AR is the name of the gene.*

significance (VUS), such as *COL11A2* of uncertain significance in case 6. The mother terminated two pregnancies with skeletal anomalies and eventually found a VUS that was useless for the guidance of the next pregnancy.

Except for the skeletal system, the neuromuscular system and head anomalies had higher diagnostic yields. In the neuromuscular system, eight fetuses with arthrogryposis multiplex congenita (AMC) of different genotypes were found. Four diagnostic genetic variants were detected in *HBA2*, *KLHL40*, *MYH3*, and *ZC4H2*. Two had VUS in *FBN3* and *TTN*. As previously reported (Hall et al., 2019), over 400 known conditions are associated with AMC. The etiology of arthrogryposis is complicated and may be related to the decrease in fetal movement caused by the abnormal central nervous system, muscle and connective tissues, lack of exercise space, maternal disease, environmental factors, and viral infection (Haliloglu and Topaloglu, 2013). Only 30% of AMC cases can find clear pathogenic genes (Skaria et al., 2019). In our study, 50% of genetic variants were found, which suggested that WES would be helpful for prenatal AMC cases. In addition, of the six cases with one anatomic system affected by head anomalies, two cases with abnormalities in the eyes and one with mandibular anomaly detected genetic variants. The major single-gene *SOX2* variant was found in a fetus with bilateral anophthalmia, which accounted for 10–15% of all anophthalmia and microphthalmia cases (Ragge et al., 2007). Although *SOX2* in anophthalmia, persistent hyperplastic primary vitreous in *NDP*, and micrognathia in *BMP2* were detected in this study, up to 60% of cases with underlying genetic causes remained undetermined.

The consequences of genetic variant might relate to pregnancy management, including both prognoses of the affected pregnancy and recurrence risk of continuing pregnancy. According to our study, a diagnostic genetic variant might be more likely to be identified in selected cases, especially skeletal anomalies as that was reported in the previous study and neuromuscular or head anomalies as shown in ours. It would be helpful for variant interpretation and might improve the diagnostic yield if WES was performed for both parents and fetuses. The multidisciplinary meeting was held among molecular geneticists, maternal–fetal medicine specialists, and an expert in obstetric ultrasound before clinical counseling. The results related to clinical phenotypes detected by ultrasound would be reported back, not secondary findings. In general, the pediatric geneticist was also included in the multidisciplinary conference, while it was not in this study, which was one of our limitations. Although 88 of 94 pregnancies were terminated, and the potential prognosis of these affected pregnancies was unknown, the genetic variants would effectively predict the recurrence risk of continuing pregnancy according to the types of inherited variants.

There are other two limitations in our study. First, our hospital is a tertiary level referral center, and abnormal cases diagnosed in downtowns are transferred to our hospital, which leads to a relatively high detection rate in rare abnormal cases, such as osteogenesis imperfecta and Apert syndromes. Second, although a large number of fetuses have structural anomalies, only a small number of parents chose to participate in our research, and cases eligible for our study included proband-only and trio groups depending on parents' willingness to provide blood samples, which might lead to inevitable selection bias. The proportion of skeletal abnormalities in the trio group might have led to a higher diagnostic yield than that in the proband-only group.

In conclusion, prenatal WES can improve the genetic diagnostic yield of anomalous fetuses with accurate phenotypes with normal karyotype and CMA results and expand the information for the prenatal genotype–phenotype spectrum. With the utility of WES, parents will gain information on the guidance and assessment of recurrence risk for the subsequent pregnancy. A preimplantation genetic diagnosis can be performed for cases with specific inherited variants.

## DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: NCBI BioProject, accession no: PRJNA713979.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by The First Affiliated Hospital of Sun Yat-sen University. The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

## AUTHOR CONTRIBUTIONS

MH and LD mainly conducted the design of this study and analysis of the results. MH mainly wrote and submitted the manuscript. HX provided the prenatal cases. LZ and YG helped with the pathogenic examination and sample collection. TL and JZ provided a few cases. DC was responsible for the sample delivery. All authors contributed to the article and approved the submitted version.

## FUNDING

This work was supported by the National Natural Science Foundation of China (Grant No. 81571687), the Youth Program of National Natural Science Foundation of China (Grant No. 81801705), and the Medical Scientific Research Foundation of Guangdong Province, China (Grant No. 2017A020214013).

## ACKNOWLEDGMENTS

We would like to thank the laboratory technicians in Guangzhou KingMed Center for Clinical Laboratory for their assistance in the sampling process.

## REFERENCES


Alamillo, C. L., Powis, Z., Farwell, K., Shahmirzadi, L., Weltmer, E. C., Turocy, J., et al. (2015). Exome sequencing positively identified relevant alterations in more than half of cases with an indication of prenatal ultrasound anomalies. *Prenat. Diagn.* 35, 1073–1078. doi: 10.1002/pd.4648

Bonafe, L., Cormier-Daire, V., Hall, C., Lachman, R., Mortier, G., Mundlos, S., et al. (2015). Nosology and classification of genetic skeletal disorders: 2015 revision. *Am. J. Med. Genet. A* 167A, 2869–2892. doi: 10.1002/ajmg.a.37365

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

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

Ely, D. M., and Driscoll, A. K. (2019). Infant mortality in the United States, 2017: data from the period linked birth/infant death file. *Natl. Vital Stat. Rep.* 68, 1–20.

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

Haliloglu, G., and Topaloglu, H. (2013). Arthrogryposis and fetal hypomobility syndrome. *Handb. Clin. Neurol.* 113, 1311–1319. doi: 10.1016/b978-0-444-59565-2.00003-4

Hall, J. G., Kimber, E., and Dieterich, K. (2019). Classification of arthrogryposis. *Am. J. Med. Genet. C Semin. Med. Genet.* 181, 300–303. doi: 10.1002/ajmg.c.31716

Han, J., Yang, Y. D., He, Y., Liu, W. J., Zhen, L., Pan, M., et al. (2020). Rapid prenatal diagnosis of skeletal dysplasia using medical trio exome sequencing: benefit for prenatal counseling and pregnancy management. *Prenat. Diagn.* 40, 577–584. doi: 10.1002/pd.5653

Lei, T. Y., Fu, F., Li, R., Wang, D., Wang, R. Y., Jing, X. Y., et al. (2017). Whole-exome sequencing for prenatal diagnosis of fetuses with congenital anomalies of the kidney and urinary tract. *Nephrol. Dial. Transplant.* 32, 1665–1675. doi: 10.1093/ndt/gfx031

Li, H., and Durbin, R. (2010). Fast and accurate long-read alignment with burrows-wheeler transform. *Bioinformatics* 26, 589–595. doi: 10.1093/bioinformatics/btp698

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

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

Ragge, N. K., Subak-Sharpe, I. D., and Collin, J. R. (2007). A practical guide to the management of anophthalmia and microphthalmia. *Eye (Lond)* 21, 1290–1300. doi: 10.1038/sj.eye.6702858

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

Sa, J., Melo, F., Tarelho, A., Branco, M., Pereira, J., Galhano, E., et al. (2018). Broad multi-gene panel or whole exome sequencing in malformed fetuses reveals eight definitive and one likely diagnoses in fifteen studied fetus, in prenatal setting. *Eur. J. Hum. Genet.* 26:129.

Skaria, P., Dahl, A., and Ahmed, A. (2019). Arthrogryposis multiplex congenita in utero: radiologic and pathologic findings. *J. Matern. Fetal Neonatal Med.* 32, 502–511. doi: 10.1080/14767058.2017.1381683

Sparks, T. N., Lianoglou, B. R., Adami, R. R., Pluym, I. D., Holliman, K., Duffy, J., et al. (2020). Exome sequencing for prenatal diagnosis in nonimmune hydrops fetalis. *N. Engl. J. Med.* 383, 1746–1756. doi: 10.1056/NEJMoa2023643

Vora, N. L., Gilmore, K., Brandt, A., Gustafson, C., Strande, N., Ramkissoon, L., et al. (2020). An approach to integrating exome sequencing for fetal structural anomalies into clinical practice. *Genet. Med.* 22, 954–961. doi: 10.1038/s41436-020-0750-4

Yadava, S. M., and Ashkinadze, E. (2019). Whole exome sequencing for prenatal diagnosis in cases with fetal anomalies: criteria to improve diagnostic yield. *J. Genet. Couns.* 28, 251–255. doi: 10.1002/jgc4.1045

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

Yates, C. L., Monaghan, K. G., Copenheaver, D., Retterer, K., Scuffins, J., Kucera, C. R., et al. (2017). Whole-exome sequencing on deceased fetuses with ultrasound anomalies: expanding our knowledge of genetic disease during fetal development. *Genet. Med.* 19, 1171–1178. doi: 10.1038/gim.2017.31

Zhang, L., Zhang, J., Yang, J., Ying, D., Lau, Y. L., and Yang, W. (2013). PriVar: a toolkit for prioritizing SNVs and indels from next-generation sequencing data. *Bioinformatics* 29, 124–125. doi: 10.1093/bioinformatics/bts627


**Conflict of Interest:** DC was employed by the company Guangzhou Kingmed Diagnostics Group.

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.

*Copyright © 2021 He, Du, Xie, Zhang, Gu, Lei, Zheng and Chen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*

frontiers in Genetics

ORIGINAL RESEARCH
published: 01 September 2021
doi: 10.3389/fgene.2021.677699

# Clinical Application of Whole Exome Sequencing for Monogenic Disorders in PICU of China

*Yingchao Liu[1,2†], Chanjuan Hao[3†], Kechun Li[1,2], Xuyun Hu[3], Hengmiao Gao[1,2], Jiansheng Zeng[1,2], Ruolan Guo[3], Jun Liu[2], Jun Guo[3], Zheng Li[1,2], Zhan Qi[3], Xinlei Jia[1,2], Wei Li[3*‡] and Suyun Qian[1,2*‡]*

[1] Pediatric Intensive Care Unit, Beijing Children's Hospital, Capital Medical University, National Center for Children's Health, Beijing, China, [2] Research Unit of Critical Infection in Children, Chinese Academy of Medical Sciences (2019RU016), Beijing, China, [3] Beijing Key Laboratory for Genetics of Birth Defects, Beijing Pediatric Research Institute, MOE Key Laboratory of Major Diseases in Children, Beijing Children's Hospital, Capital Medical University, National Center for Children's Health, Beijing, China

OPEN ACCESS

**Edited by:**
Natália Duarte Linhares, Genuity Science, Ireland

**Reviewed by:**
Ni-Chung Lee, National Taiwan University, Taiwan
Baiba Lace, Biomedical Research and Study Centre, Latvia

***Correspondence:**
Wei Li
liwei@bch.com.cn
Suyun Qian
syqian2020@163.com

†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 Genetics of Common and Rare Diseases, a section of the journal Frontiers in Genetics

**Received:** 08 March 2021
**Accepted:** 03 August 2021
**Published:** 01 September 2021

**Citation:**
Liu Y, Hao C, Li K, Hu X, Gao H, Zeng J, Guo R, Liu J, Guo J, Li Z, Qi Z, Jia X, Li W and Qian S (2021) Clinical Application of Whole Exome Sequencing for Monogenic Disorders in PICU of China. Front. Genet. 12:677699. doi: 10.3389/fgene.2021.677699

**Objectives:** Whole exome sequencing (WES) has been widely used to detect genetic disorders in critically ill children. Relevant data are lacking in pediatric intensive care units (PICUs) of China. This study aimed to investigate the spectrum of monogenic disorders, the diagnostic yield and clinical utility of WES from a PICU in a large children's hospital of China.

**Methods:** From July 2017 to February 2020, WES was performed in 169 critically ill children with suspected monogenic diseases in the PICU of Beijing Children's Hospital. The clinical features, human phenotype ontology (HPO) terms, and assessment of clinical impact were analyzed.

**Results:** The media age of the enrolled children was 10.5 months (range, 1 month to 14.8 years). After WES, a total of 43 patients (25%) were diagnosed with monogenic disorders. The most common categories of diseases were metabolic disease (33%), neuromuscular disease (19%), and multiple deformities (14%). The diagnosis yield of children with "metabolism/homeostasis disorder" and "growth delay" or "ocular anomalies" was higher than that of children without these features. In addition, the diagnosis rate increased when more features were observed in children. The results of WES had an impact on the treatment for 30 cases (70%): (1) change of treatment ($n$ = 11), (2) disease monitoring initiation ($n$ = 18), (3) other systemic evaluation ($n$ = 3), (4) family intervention ($n$ = 2), and (5) rehabilitation and redirection of care toward palliative care ($n$ = 12).

**Conclusion:** WES can be used as an effective diagnostic tool in the PICU of China and has an important impact on the treatment of patients with suspected monogenic conditions.

**Keywords: whole exome sequence, pediatric intensive care unit, monogenic disorders, clinical application, effective**

# INTRODUCTION

Congenital genetic disease is frequent in pediatric intensive care units (PICUs) (French et al., 2019). At present, almost 7,000 monogenic diseases have been listed on the OMIM website[1], and this number is growing at a rate of approximately 30–50 per month. Children with monogenic diseases who are admitted to PICU are often in critical condition and accompanied by multiple organ dysfunction (Sanford et al., 2019). Considering the genetic and clinical heterogeneity of monogenic disease, critical physicians cannot provide a rapid and clear diagnosis and make a targeted treatment plan in time. The condition of the patient may continue to progress, increasing the risk of death. Cunniff et al. (1995) reported that the mortality rate of children admitted to the hospital due to birth defects and genetic diseases was 4.5 times higher than that of children admitted for other reasons. Therefore, clinicians must diagnose critically ill children with suspected monogenic diseases in time to provide them with timely interventions, improve prognosis, reduce mortality, and minimize the burden on families and the national health system.

Next-generation sequencing provides a new strategy for diagnosing genetic disorders. In recent years, an increasing number of studies have shown that whole exome sequencing (WES) has a unique value in the diagnosis of monogenic diseases. WES could be used as a first-tier test for children with suspected monogenic disorders (Stark et al., 2016; Hu et al., 2018). More than 20% of patients could be diagnosed using WES (Yang et al., 2013; Meng et al., 2017; French et al., 2019). Farnaes et al. (2018) reported that WES could reduce mortality and hospitalization costs. This may be due to the fact that after obtaining a definite diagnosis, the best clinical care can be provided for the affected children, promoting the transition from empirical treatment to definite treatment of identified diseases where feasible.

Previous studies of WES mainly included infants in neonatal intensive care units (NICUs) (Meng et al., 2017; Farnaes et al., 2018; Petrikin et al., 2018; Swaggart et al., 2019; Gubbels et al., 2020). However, studies about PICUs either had a relatively small number of cases or were reported jointly with NICUs (Mestek-Boukhibar et al., 2018; French et al., 2019; Sanford et al., 2019; Australian Genomics Health Alliance Acute Care et al., 2020; Wang et al., 2020a). The present study focused on the clinical application of WES in critically ill patients in the PICU. The diagnostic yield, the monogenic disease spectrum of patients, and effect of WES on clinical management were systematically analyzed for the first time in a PICU of China.

# MATERIALS AND METHODS

## Study Design and Clinical Samples

We retrospectively analyzed the demographics and clinical characteristics of critically ill patients in the PICU who received WES, the diagnostic yield of WES, and the clinical management of diagnosed patients.

From July 2017 to February 2020, 3,484 patients were admitted to the PICU of Beijing Children's Hospital. A total of 169 patients who met the inclusion criteria were enrolled in this study (**Figure 1**).

The inclusion criteria were: (1) age $\geq$29 days and <18 years, (2) suspected with monogenic diseases after consulting at least one geneticist, and (3) consented and received WES (trio-WES or proband WES). The exclusion criteria were as follows: (1) diagnosed with non-monogenic etiology and (2) no response during follow-up. This study was approved by the institutional review board of Beijing Children's Hospital (No. 2020-Z-098).

## WES and Data Analysis

DNA was extracted from EDTA blood samples obtained from the probands and their parents (when available) using the Gentra Puregene Blood Kit (QIAGEN, Hilden, Germany). The SureSelect Human All Exon Kit V6 (Agilent Technologies, Santa Clara, United States) was used for whole exome capture. The target regions were sequenced on NovaSeq 6000 (Illumina, San Diego, United States). Raw reads were filtered using FastQC to remove low-quality reads. Clean reads were mapped to the reference genome GRCh37. Quality control information included the following: average read depth of >100×, accurate mapping rate of >98%, base capture rate of >55%, 20× mean depth coverage rate of >96%, and duplication rate of <25%. Single nucleotide variants were annotated and filtered by TGex[2]. Variants with a frequency over 1% in the gnomAD, ESP, or 1000G databases were excluded. Variants that lacked segregation in family members were also filtered. Variants were classified following the American College of Medical Genetics and Genomics and the Association for Molecular Pathology interpretation standards and guidelines (Richards et al., 2015). The genetic results were reviewed by a multidisciplinary team which included geneticists, pediatric intensive care physicians, neurologists as required, and compared with the phenotypes, related biochemical tests, or imaging examinations of the patients to make the final diagnosis (**Supplementary Table 1**).

## Phenotype Analysis

The clinical features and the impact on patient prognosis were evaluated by reviewing electronic medical records. The clinical features were translated into human phenotype ontology (HPO) terms through the HPO website[3]. The top-level branching of HPO phenotypes was analyzed to ensure adequate counts of participants. The diagnostic rates of patients with and without HPO terms, the relationship between the phenotype complexity and expected diagnostic yield were analyzed.

## Assessment of Clinical Impact

The changes in clinical management were determined from the medical records and follow-up by telephone until June 2020. These changes were grouped into five categories: (1) changes of treatment, (2) disease monitoring initiation, (3) other systemic evaluation, (4) family intervention, and (5) rehabilitation and redirection of care toward palliative care.

[1]https://omim.org/

[2]https://geneyx.com/geneyxanalysis/

[3]https://hpo.jax.org/app/

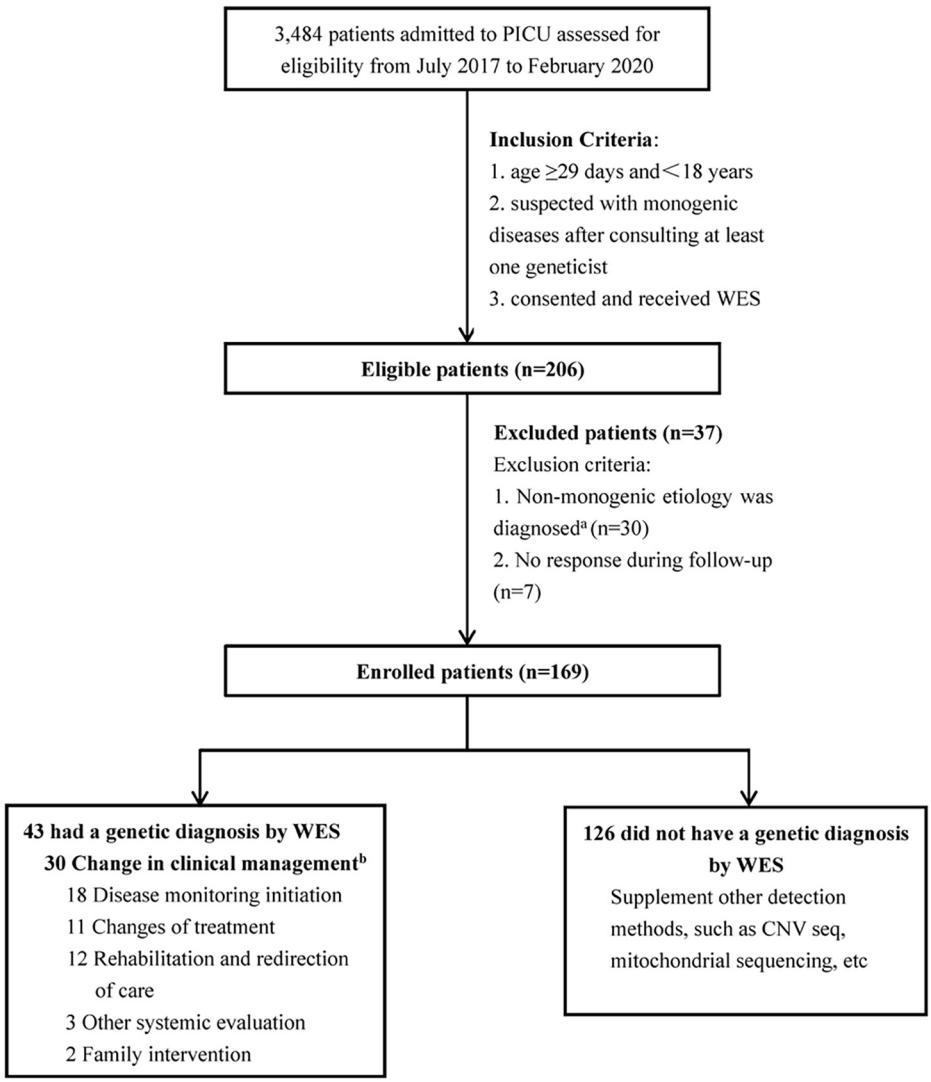



**FIGURE 1 |** Flow diagram of the PICU patients who were enrolled, received genetic diagnosis by WES, and clinical management. [a]The etiology of thirty patients was non-monogenic diseases, including: 19 acute necrotizing encephalopathy; 7 febrile infection-related epilepsy syndrome; 1 acute lymphoblastic leukemia; 1 poisoning; 1 heat stroke; 1 hemophagocytic syndrome. [b]Some patients had changed the clinical management more than 1 category.

## Statistical Analysis

All analyses were conducted in SPSS 22.0. Categorical variables were expressed as frequencies and percentages, continuous variables as median (interquartile range). Differences between groups were determined by Student *t*-test, Mann-Whitney *U* test, chi-square, if the expected counts were less than 5, Fisher's exact test. *p*-values of less than 0.05 were considered to indicate statistical significance.

# RESULTS

## General Information

We enrolled for WES analysis 169 patients, 96 (57%) males, and 73 (43%) females. The ages of patients ranged from 1 month to 14.8 years, and the median age was 10.5 months (4 months, 47.5 months). The most common primary manifestation was neurological (21%), followed by cardiac (17%) and multiple congenital anomalies (15%) (**Figure 2**). The conditions of affected patients were severe with 82% of patients requiring mechanical ventilation and 62% of patients needing endotracheal intubation. In addition, more than 20% of patients received inotropic support or continuous renal replacement therapy. The mortality rate at discharge was 12%, and the mortality at 120 days after discharge was 32%.

## Diagnostic Yield of WES in PICU

Among the 169 enrolled patients, trio WES was performed on 160 (95%) patients, while proband-only WES was performed on the remaining nine (5.3%) patients on the basis of the availability

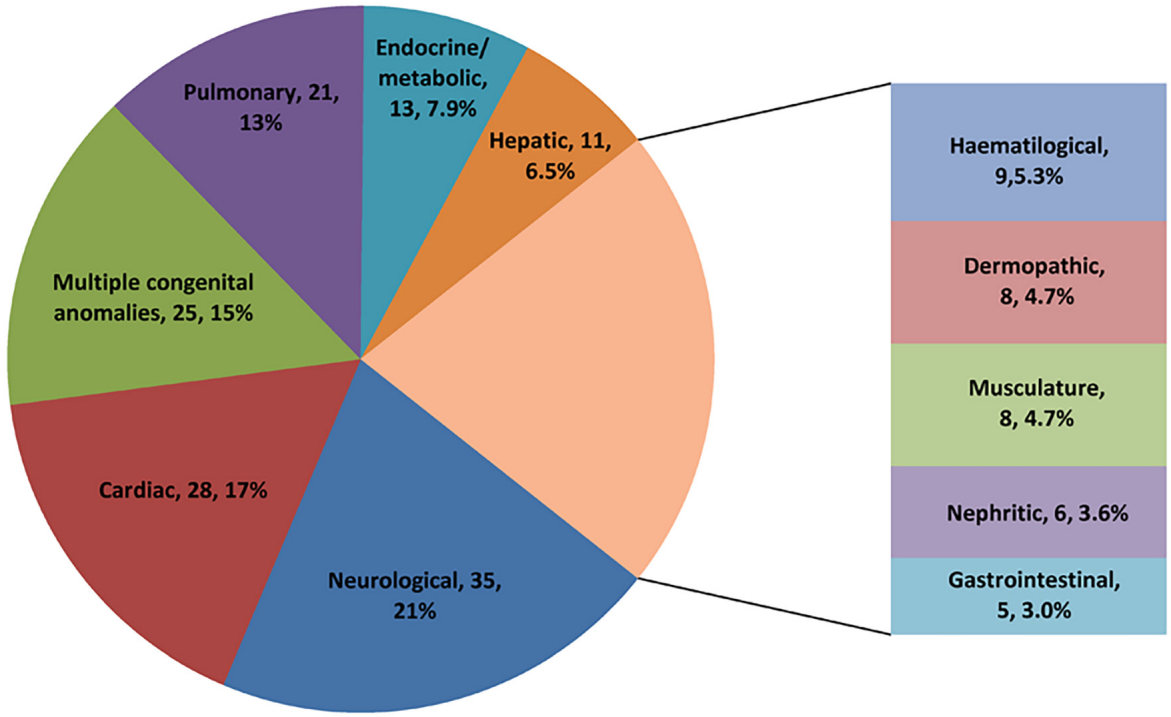



**FIGURE 2 |** The composition of primary clinical manifestations of the 169 enrolled patients.

of parental samples. Forty-three cases had monogenic disorders, and the diagnosis rate was 25%.

In the 43 diagnosed cases, 40 monogenic diseases were detected. The most common type was metabolic diseases (14/43, 33%), followed by neurological and muscular diseases (8/43, 19%), multiple malformations (6/43, 14%), cardiac diseases (4/43, 9.3%), immunodeficiency diseases (2/43, 4.7%), nephritic diseases (2/43, 4.7%), and other diseases (7/43, 16%) (**Table 1** and **Supplementary Table 1**). *ATP7B*-related Wilson disease and *KMT2D*-related Kabuki syndrome were observed in 3 (7.0%) and 2 (4.7%) patients, respectively. The three children with Wilson disease had an initial symptom of acute liver failure, and then they were admitted to the PICU. Of all diagnosed patients, 20 (46%) were autosomal dominant (with 14 *de novo*, 2 inherited, and 4 inherited unknown), 21 (49%) were autosomal recessive (with 18 compound heterozygous and 3 homozygous), one was X-linked dominant (with *de novo*), and one was X-linked recessive (with maternal carrier).

In this study, the diagnostic rates were compared among patients with different phenotypes represented by HPO terms. We found that specific clinical features were associated with a molecular diagnosis. For example, patients with neurological and cardiovascular anomalies received diagnosis rates higher than 30%. Patients with HPO terms of "Abnormality of metabolism/homeostasis" (HP: 0001939), "Growth abnormality" (HP: 0001507), and "Abnormality of the eye" (HP: 0000478) had a higher diagnostic yield than those without the terms (**Table 2**) ($P < 0.05$). Moreover, we observed the relationship between the complexity of patient's phenotype (reflected by the number of HPO terms) and the expected diagnostic rate (**Figure 3A**). The diagnostic yield of cases with only one HPO term was 17%. When more than three HPO terms were given, the diagnostic rate will increase and exceed 30%. For HPO terms greater than 6, the diagnosis rate will reach as high as 50%. However, no statistical difference was observed in the diagnosis yield between children with different HPO terms ($P = 0.24$).

The diagnostic yield of patients aged more than 12 years was the highest (45%), and that of patients aged less than 1 year was relatively low (22%) (**Figure 3B**). No statistical difference was observed in the diagnosis rate between children of different age groups ($P = 0.36$).

The media turnaround time of WES in diagnosed patients was 11 days, of which 44% of patients received definite diagnosis before discharge.

## Clinical Impact of WES Diagnosis

In this study, 30 patients (70%) were affected by the clinical management. We evaluated the impact of WES on clinical treatment through the following aspects (**Table 1** and **Supplementary Table 1**).

### Changes of Treatment

Treatment changes affected 11 patients. Due to infection-induced "hemolytic crisis" with seizure and impairment of consciousness, case 006 was admitted to the PICU. In addition, the child was hospitalized many times because of repeated jaundice. This patient was diagnosed as *GPI*-related non-spherocytic hemolytic anemia (OMIM: 613470) through WES and discharged

**TABLE 1 |** Brief clinical and genetic information of 43 patients with positive molecular diagnosis.

| Case | Age | Sex | Phenotype | Gene | Variant | Inheritance Pattern | Zygosity | Disease | Clinical impact |
|---|---|---|---|---|---|---|---|---|---|
| 004 | 2m | M | Congenital hypertrophy of left ventricle; Patent ductus arteriosus; Hypertrichosis; Coarse facial features | *ABCC9* | c.3203T > C (p.Leu1068Pro) | AD | *De novo* het | Cantu syndrome (OMIM[#]239850) | Cardiac surveillance |
| 006 | 8y1m | F | Jaundice; Pigment gallstones; Splenomegaly; Non-spherocytic hemolytic anemia; Seizure; Loss of consciousness | *GPI* | c.637T > A (p.Phe213Ile); c.1614C > G (p.His538Gln) | AR | Compound het | Hemolytic anemia, non-spherocytic, due to glucose phosphate isomerase deficiency (OMIM[#]613470) | Splenectomy; Follow-up of hematology major regularly |
| 012 | 3y | F | Renal insufficiency; Anemia | *ANKS6* | c.2420dupT (p.Thr808Aspfs*2); c.1973-1G > A | AR | Compound het | Nephronophthisis 16 (OMIM[#]615382 | Redirection of care |
| 017 | 4y6m | F | Lethargy; Coma; Seizures; Muscular hypotonia; Acute hyperammonemia; Global developmental delay | *MCCC1* | c.639 + 2T > A; c.1679dupA (p.Asn560Lysfs*10) | AR | Compound het | 3-methylcrotonyl-CoA carboxylase 1 deficiency (OMIM[#]210200) | Rehabilitation; Protein-restricted diet; Biotin supplementation; Ammonia surveillance |
| 026 | 5m | F | Pneumonia; Diarrhea; Severe combined immunodeficiency; Panhypogammaglobulinemia; Severe T lymphocytopenia; B lymphocytopenia | *RAG1* | c.1528G > T (p.Glu510*); c.2393A > G (p.His798Arg) | AR | Compound het | Severe combined immunodeficiency, autosomal recessive, T cell-negative, B cell-negative, NK cell-positive (OMIM[#]601457) | NA |
| 027 | 3y6m | M | Failure to thrive; Global developmental delay; Seizure; Hypertrophic cardiomyopathy | *MAP2K1* | c.389A > G (p.Tyr130Cys) | AD | Het | Cardiofaciocutaneous syndrome 3 (OMIM[#]615279) | NA |
| 032 | 3m | M | Weak cry; Premature birth; Intrauterine growth retardation; Small for gestational age; Limb muscle weakness; Failure to thrive; Tachypnea; Respiratory failure; Hyporeflexia; Global developmental delay; Camptodactyly of finger | *IGHMBP2* | c.1061-2A > G | AR | Homo | Spinal muscular atrophy, distal, autosomal recessive, 1 (OMIM[#]604320) | Redirection of care |
| 034 | 13y10m | M | Hepatic steatosis; Visual impairment; Abnormality of lipid metabolism; Nystagmus; Global developmental delay; Hepatomegaly; Splenomegaly; Hypertriglyceridemia; Renal insufficiency; Type II diabetes mellitus; Hypertension; Pulmonary hypertension; Diabetes insipidus; Dilated cardiomyopathy | *ALMS1* | c.8532_8533insA (p.Leu2845Thrfs*3); c.9152_9153delCT (p.Cys3053Serfs*9) | AR | Compound het | Alstrom syndrome (OMIM[#]203800) | Redirection of care |
| 041 | 3y10m | M | Hypertrophic cardiomyopathy; Muscular hypotonia; Dyspnea; Respiratory insufficiency due to muscle weakness; Myopathy; Arrhythmia | *GAA* | c.1082C > T (p.Pro361Leu); c.2051C > T (p.Pro684Leu) | AR | Compound het | Glycogen storage disease II(OMIM[#]232300) | Redirection of care; Sister diagnosed and started on enzyme replacement therapy |
| 043 | 1y8m | M | Agammaglobulinemia; Pneumonia | *BTK* | c.1781G > A (p.Gly594Glu) | XLD | *De novo* hemi | Agammaglobulinemia, X-linked (OMIM[#]300755) | Immunoglobulin replacement therapy |
| 050 | 1m | F | Coarctation of aorta; Hirsutism; Micrognathia | *KMT2D* | c.14113_14123del11(p.Ile4705Phefs*4) | AD | *De novo* het | Kabuki syndrome 1 (OMIM[#]147920) | Routine monitoring of aortic valve stenosis; Endocrinology and audiology evaluation |
| 057 | 1y3m | F | Diarrhea; Enterocolitis; Perianal abscess; Pyoderma | *IL10RA* | c.301C > T (p.Arg101Trp); c.537G > A (p.Thr179Thr) | AR | Compound het | Inflammatory bowel disease 28, autosomal recessive (OMIM[#]613148) | NA |
| 059 | 12y2m | M | Lethargy; Hyperammonemia; Vomiting; Hepatomegaly; Global developmental delay | *ASL* | c.91G > A (p.Asp31Asn); c.965_966insCAAAGACTT (p.Asn322_Lys323insLysAspPhe) | AR | Compound het | Argininosuccinic aciduria (OMIM[#]207900) | Protein-restricted diet; Arginine; Ammonia surveillance |
| 065 | 2m | M | Feeding difficulties; Dystonia; EEG abnormality; Seizures; Horizontal nystagmus | *GABRB2* | c.909G > T (p.Lys303Asn) | AD | *De novo* het | Epileptic encephalopathy, infantile or early children,2 (OMIM[#]617829) | NA |
| 075 | 8y4d | F | Jaundice; Hemolytic anemia; Hepatic failure; Kayser-Fleischer ring; Decreased serum ceruloplasmin; Increased urinary copper concentration | *ATP7B* | c.2333G > T (p.Arg778Leu) | AR | Homo | Wilson disease (OMIM[#]277900) | Penicillamine and Zinc Sulfate; Liver function and hemolysis surveillance |
| 081 | 11m7d | M | Dilated cardiomyopathy; Left ventricular non-compaction cardiomyopathy; Cardiac arrest | *CASZ1* | c.2443_2459del GTGGGCACCCCCAGCCT (p.Val815Profs*14) | AD | *De novo* het | [a]CASTOR ZINC FINGER PROTEIN 1; CASZ1 (OMIM[*]609895) | NA |
| 084 | 1y1m | F | Seizures; Global developmental delay; Microcephaly; Muscularhypotonia; Hepatomegaly | *IFIH1* | c.1852C > T (p.Arg618*) | AD | *De novo* het | Aicardi-Goutieres syndrome 7 (OMIM[#]615846) | Redirection of care |
| 091 | 10m | F | Vomiting; Elevated hepatic transaminases; Metabolic acidosis | *HMGCS2* | c.1347_1351del AGCCT (p.Ala450Profs*7) c.1201G > T (p.Glu401*) | AR | Compound het | 3-hydroxy-3-methylglutaryl-CoA synthase-2 deficiency (OMIM[#]605911) | Avoid prolonged hunger; Coenzyme Q10 and levocarnitine; Liver function and blood gas analysis surveillance |
| 096 | 2m | F | Neutropenia; Thrombocytopenia; Lethargy; Hyperhomocystinemia; Methylmalonic acidemia | *MMACHC* | c.566_567insT (p.Ile190Tyrfs*13); c.609G > A (p.Trp203*) | AR | Compound het | Methylmalonic aciduria and homocystinuria, cblC type (OMIM[#]277400) | Redirection of care; Vitamin B6, Vitamin B12, levocarnitine, folic acid and betaine |

*(Continued)*

**TABLE 1 |** Continued

| Case | Age | Sex | Phenotype | Gene | Variant | Inheritance Pattern | Zygosity | Disease | Clinical impact |
|---|---|---|---|---|---|---|---|---|---|
| 103 | 5m | F | Microcephaly; Premature birth; Small for gestational age; Intrauterine growth retardation; Recurrent infections; Bilateral single transverse palmar creases; Global developmental delay | *RNU4ATAC* | n.16G > A (non-coding transcript variant-regulatory mutation); n.51G > A (non-coding transcript variant-regulatory mutation) | AR | Compound het | Roifman syndrome (OMIM[#]616651) | NA |
| 105 | 5y11m | M | Pulmonary artery stenosis; Peripheral arterial stenosis; Supravalvular aortic stenosis | *ELN* | c.1621C > T (p.Arg541*) | AD | Het | Supravalvular aortic stenosis (OMIM[#]185500) | Redirection of care |
| 106 | 2m | M | Webbed neck; Low-set ears; Hypertelorism Pulmonic stenosis | *KRAS* | c.173C > T (p.Thr58Ile) | AD | Het | Noonan syndrome 3 (OMIM[#]609942) | Routine monitoring of pulmonic; Regular follow-up of cardiac and neurology |
| 109 | 2m | M | Type I diabetes mellitus; Diabetic ketoacidosis | *EIF2AK3* | c.641A > G (p.Tyr214Cys); c.12delC (p.Ile5Serfs*66) | AR | Compound het | Epiphyseal dysplasia, multiple, with early-onset diabetes mellitus (OMIM[#]226980) | Height, weight and growth evaluation |
| 112 | 2y2m | F | Muscular hypotonia; Hyporeflexia; Motor delay; Distal muscle weakness; Infantile onset | *MPZ* | c.380G > A (p.Cys127Tyr) | AD | *De novo* het | Hypertrophic neuropathy of Dejerine-Sottas (OMIM[#]145900) | NA |
| 113 | 2y5m | M | Motor delay; Neonatal hypotonia; Muscular hypotonia; Generalized muscle weakness; Infantile onset | *RYR1* | c.6721C > T (p.Arg2241*); c.5915A > T (p.Asn1972Ile) | AR | Compound het | Central core disease of muscle (OMIM[#]117000) | Redirection of care |
| 120 | 6m | F | Infantile onset; Generalized myoclonic seizures; Status epilepticus; Epileptic encephalopathy | *SCN1A* | c.2758C > A (p.Arg920Ser) | AD | *De novo* het | Epileptic encephalopathy, early infantile, 6 (OMIM[#]607208) | NA |
| 140 | 1m | F | Cyanosis; Bilateral tonic-clonic seizure with focal onset | *SCN2A* | c.4976C > T (p.Ala1659Val) | AD | *De novo* het | Seizures, benign familial infantile, 3(OMIM[#]607745) | Redirection of care |
| 149 | 5y3m | M | Cryptorchidism; Hypospadias; Hypertelorism; Anal atresia; Pectus excavatum | *MID1* | c.1863_1879dupGAACTCCA TCCACCTCT (p.Tyr627*) | XLR | Inherited hemi (maternal carrier) | Opitz GBBB syndrome, type I (OMIM[#]300000) | NA |
| 151 | 1m | M | Elevated hemoglobin A1c; Hyperglycemia; Diabetic ketoacidosis | *KCNJ11* | c.601C > T (p.Arg201Cys) | AD | *De novo* het | Diabetes mellitus, transient neonatal, 3(OMIM[#]610582) | NA |
| 153 | 1m | M | Defect in the atrial septum; Ventricular septal defect; Muscular hypotonia; Macrotia; Hypoglycemia | *KMT2D* | c.5269C > T (p.Arg1757*) | AD | *De novo* het | Kabuki syndrome 1 (OMIM[#]147920) | Cardiac surveillance; Endocrinology and audiology evaluation |
| 154 | 8m | M | Seizures; Lethargy; Global developmental delay; Hypoglycemia; Metabolic acidosis; Hepatomegaly; Medium chain dicarboxylic aciduria; Muscular hypotonia | *ACADM* | c.449_452delTGAC (p.Thr150Argfs*4); c.1010A > C (p.Tyr337Ser) | AR | Compound het | Acyl-CoA dehydrogenase, medium-chain, deficiency of (OMIM[#]201450) | Avoid prolonged hunger; Levocarnitine; Regular follow-up of neurology; Initiated a reproductive program |
| 159 | 13y4m | F | Glioblastoma | *TP53* | c.375G > A (p.Thr125Thr) | AD | Inherited het (Paternal carrier) | Glioma susceptibility 1 (OMIM[#]137800) | NA |
| 176 | 9y7d | M | Sinus bradycardia; Atrioventricular block; Sick sinus syndrome; Ventricular escape rhythms | *SCN5A* | c.4895G > A (p.Arg1632His) | AD | Inherited het (maternal carrier) | Sick sinus syndrome 1 (OMIM [#]608567) | Electrocardiogram and echocardiography surveillance regularly |
| 177 | 5m | M | Psoriasis; Pustule Erythema | *IL36RN* | c.115 + 6T > C | AR | Homo | Psoriasis 14, pustular (OMIM[#]614204) | Follow-up of dermatological major |
| 178 | 2y3m | M | Short stature | *P4HB* | c.236dupT (p.Leu79Phefs*11) | AD | *De novo* het | Cole-carpenter syndrome 1 (OMIM[#]112240) | NA |
| 181 | 12y5m | M | Pectus excavatum; Tall stature; Dolichocephaly; Narrow face; Long face; Myopia; Decreased subcutaneous fat; Aortic root dilation | *FBN1* | c.1995C > A (p.Tyr665*) | AD | *De novo* het | Marfan syndrome (OMIM[#]154700) | Routine monitoring of aorta |
| 186 | 2y8m | M | Lethargy; Seizures; Elevated hepatic transaminases; Hepatomegaly; Hyperammonemia; Myopathy | *CPT2* | c.1148T > A (p.Phe383Tyr); c.1749C > A (p.Asn583Lys) | AR | Compound het | Carnitine palmitoyltransferase II deficiency, infantile (OMIM[#]600649) | High-carbohydrate and low-fat diet; Levocarnitine; Regular monitoring of liver function, muscle enzyme, blood glucose, blood lipid and echocardiography |

*(Continued)*

**TABLE 1** | Continued

| Case | Age | Sex | Phenotype | Gene | Variant | Inheritance Pattern | Zygosity | Disease | Clinical impact |
|---|---|---|---|---|---|---|---|---|---|
| 187 | 1y2m | F | Dilated cardiomyopathy | *TNNI3* | c.536A > G (p.Glu179Gly) | AD | *De novo* het | Cardiomyopathy, dilated, 1FF (OMIM[#]613286) | Cardiac surveillance |
| 196 | 8y8m | F | Pheochromocytoma; Episodic hypertension; Renal artery stenosis; Elevated urinary norepinephrine; Proteinuria | *SDHB* | c.137G > A (p.Arg46Gln) | AD | Het | Pheochromocytoma (OMIM[#]171300) | NA |
| 197 | 14y7m | F | Jaundice; Hemolytic anemia; Proteinuria; Hepatic failure; Kayser-Fleischer ring; Decreased serum ceruloplasmin; Increased urinary copper concentration | *ATP7B* | c.3836A > G (p.Asp1279Gly); c.3168delT (p.Leu1057Trpfs*64) | AR | Compound het | Wilson disease (OMIM[#]277900) | Penicillamine and zinc sulfate; Liver function and hemolysis surveillance |
| 201 | 8y8m | F | Hepatic failure; Jaundice; Cholestatic liver disease; Decreased serum ceruloplasmin | *ATP7B* | c.2333G > T (p.Arg778Leu); c.2975C > T (p.Pro992Leu) | AR | Compound het | Wilson disease (OMIM[#]277900) | Redirection of care |
| 204 | 2m26d | M | Elevated hepatic transaminases; Lactic acidosis; Hypoglycemia; Hyperuricemia; Hepatomegaly; Protuberant abdomen; Hypertriglyceridemia | G6PC | c.248G > A (p.Arg83His); c.310C > T (p.Gln104*) | AR | Compound het | Glycogen storage disease Ia (OMIM[#]232200) | Avoid prolonged hunger; Monitor blood glucose, height, weight; Endocrine follow-up |
| 205 | 4m15d | M | Decreased serum complement factor H; Hematuria | *CFH* | c.3548G > T (p.Trp1183Leu) | AD | *De novo* het | Complement factor H deficiency (OMIM[#]609814) | Redirection of care |

[a]*Gene description see Guo et al. (2019). NA, not available. AD, autosomal dominant; AR, autosomal recessive; Het, heterozygous; Hemi, hemizygous; Homo, homozygous; XLD, X-linked dominant; XLR, X-linked recessive.*

from the hospital after laparoscopic cholecystectomy and splenectomy. Her condition was normal, and her jaundice was significantly alleviated.

Moreover, some metabolic diseases required diet therapy in addition to drug treatment. For example, a patient (case 059) with argininosuccinic aciduria (OMIM: 207900) received a protein-restricted diet, another child (case 186) with carnitine palmitoyltransferase II deficiency (OMIM: 600649) needed a high-carbohydrate and low-fat diet, while other children should avoid prolonged hunger.

## Disease Monitoring Initiation

Eighteen patients underwent further disease monitoring after diagnosis. For example, regular electrocardiogram monitoring and echocardiography surveillance were required for case 176 with sick sinus syndrome to monitor for arrhythmias. Liver function and blood gas monitoring was critical for case 91 with 3-hydroxy-3-methylglutaryl-CoA synthase-2 deficiency to understand the metabolic state of the body and provide timely intervention. Routine neurological examinations were required for patients with brain damage.

## Other Systemic Evaluation

After the diagnosis of WES, three patients were subjected to other systemic evaluation: cases 050 and 153 had *KMT2D*-related Kabuki syndrome 1 (OMIM: 147920) and underwent endocrinology and audiology evaluations. Case 109 had *EIF2AK3*-related multiple epiphyseal dysplasia with early-onset diabetes mellitus (OMIM: 226980). In addition to ongoing diabetes treatment, height, weight, growth, and development were evaluated and monitored.

## Family Intervention

In addition to the effect on medical services, WES also had a potential influence on family members. Case 041 was diagnosed with glycogen storage disease II (OMIM: 232300). The child's 9-year-old sister had normal activity endurance but could not tolerate activities such as dancing and running. She underwent relevant genetic testing and was subsequently diagnosed. Fortunately, she was started on enzyme replacement therapy, and lives a normal life at present. Another family initiated a reproductive program through genetic counseling.

## Rehabilitation and Redirection of Care Toward Palliative Care

Considering the severity and poor prognosis of the diseases, parents may make different decisions after the diagnosis of WES. Case 017 had 3-methylcrotonyl-CoA carboxylase 1 deficiency (OMIM: 210200). After treatment in the PICU, the child's vital signs gradually stabilized, but consciousness did not return. The parents chose to leave the PICU for rehabilitation training and accompany the child after diagnosis. The child is now conscious and able to walk. However, 11 patients' family members chose to stop invasive surgery or life support equipment to alleviate the pain of their children. Most of the children suffered from severe disorders or life-limiting conditions.

**TABLE 2 |** Comparison of diagnostic rate by whole exome sequencing in groups with and without the phenotype.

| HPO terms | Total (%) | Diagnostic rate in patients with the term (%) | Diagnostic rate in patients without the term (%) | *P*-value[a] |
|---|---|---|---|---|
| Abnormality of the eye (HP:0000478) | 11(6.5) | 7/11 (64) | 36/158 (23) | **0.006** |
| Abnormality of metabolism/homeostasis (HP:0001939) | 35 (21) | 15/35 (43) | 28/134 (21) | **0.007** |
| Growth abnormality (HP:0001507) | 11 (6.5) | 6/11 (55) | 37/158 (23) | **0.03** |
| Abnormality of head and neck (HP:0000152) | 12 (7.1) | 6/12 (50) | 37/157 (24) | 0.06 |
| Abnormality of connective tissue (HP:0003549) | 4 (2.4) | 0/4 (0) | 43/165 (26) | 0.12 |
| Abnormality of the ear (HP:0000598) | 3 (1.8) | 2/3 (67) | 41/166 (25) | 0.16[a] |
| Abnormality of the musculature (HP:0003011) | 28 (17) | 10/28 (36) | 33/141 (23) | 0.17 |
| Abnormality of the endocrine system (HP:0000818) | 12 (7.1) | 5/12 (42) | 38/157 (24) | 0.20 |
| Abnormality of the integument (HP:0001574) | 29 (17) | 10/29 (34) | 33/140 (24) | 0.22 |
| Abnormality of voice (HP:0001608) | 1 (0.6) | 1/1 (100) | 42/168 (25) | 0.25[a] |
| Abnormality of the respiratory system (HP:0002086) | 38 (22) | 7/38 (18) | 36/131 (27) | 0.26 |
| Abnormality of digestive system (HP:0025031) | 49 (29) | 15/49 (31) | 28/120 (23) | 0.32 |
| Abnormality of the cardiovascular system (HP:0001626) | 58 (34) | 13/58 (22) | 30/111 (27) | 0.51 |
| Abnormality of the skeletal system (HP:0000924) | 16 (9.5) | 5/16 (31) | 38/153 (25) | 0.58 |
| Abnormality of blood and blood-forming tissues (HP:0001871) | 31 (18) | 9/31 (29) | 34/138 (25) | 0.61 |
| Abnormality of the nervous system (HP:0000707) | 72 (43) | 17/72 (24) | 26/97 (27) | 0.64 |
| Abnormality of limbs (HP:0040064) | 6 (3.6) | 2/6 (33) | 41/163 (25) | 0.66 |
| Neoplasm (HP:0002664) | 7 (4.1) | 2/7 (29) | 41/162 (25) | 0.85 |
| Abnormality of the genitourinary system (HP:0000119) | 23 (14) | 6/23 (26) | 37/146 (25) | 0.94 |
| Abnormality of the immune system (HP:0002715) | 35 (21) | 9/35 (26) | 34/134 (25) | 0.97 |
| Abnormality of prenatal development or birth (HP:0001197) | 8 (4.7) | 2/8 (25) | 41/161 (25) | 0.98 |

[a] *Fisher's exact test. The bold values are less than 0.05, which indicate statistical significance.*

## DISCUSSION

WES has been increasingly used to diagnose genetic diseases. Previous studies on WES mainly focused on NICUs (Smith et al., 2015; Willig et al., 2015; Wang et al., 2020b). In the present study, we first reported the use of WES in a relatively large number of critically ill children in a PICU from China. The age of patients in this study ranged from 1 month to 14.8 years, with an average age of 2.8 years, which was older than those in previous similar studies (Mestek-Boukhibar et al., 2018; French et al., 2019; Australian Genomics Health Alliance Acute Care et al., 2020; Wang et al., 2020a). The most common presentation was neurological (21%), cardiac (17%), and multiple congenital anomalies (15%), which were similar to those in previous studies (Farnaes et al., 2018; Wang et al., 2020a).

The diagnosis yield varied with different phenotypes, and increased with the phenotypic complexity. Stark et al. (2016) reported that infants with characteristics of strongly suggestive monogenic diseases, such as neurometabolic status and skeletal dysplasia, received WES, and the diagnosis yield was 57%. In the present study, the indication for clinical WES that was assessed by HPO phenotype analysis as having high diagnosis value included (1) "abnormality of metabolism/homeostasis," such as metabolic acidosis; (2) "growth abnormality," such as failure to thrive; and (3) "abnormality of the eye," such as hypertelorism. Therefore, the strict screening of phenotypes can be used to improve the diagnosis yield. However, a previous study had shown that restricting the recruitment to these specific phenotypes would miss many important diagnoses (French et al., 2019). In the future work, more stringent screening should be carried out on the conditions for joining the group to improve the diagnosis rate. Interestingly, Trujillano et al. (2017) reported that complex phenotypes including several symptoms tend to have a higher diagnostic yield, which is consistent with our findings. This might suggest that patients' phenotypes with several specific symptoms have a higher probability to be explained by monogenic reasons. Therefore, we recommend physicians who request WES to include as many symptoms as possible in their clinical description, but to mark those symptoms that seem to be specific to a particular patient. WES can find the complex manifestations of monogenic diseases. With its wide application in clinical practice, it can be used to easily diagnose patients with overlapped symptoms and mixed phenotypes.

Combined with other detection methods or information supplementation, the diagnosis yield could be improved. Hu et al. (2020) reported that through parallel tests of WES and copy number variant sequencing, the diagnosis yield could be increased from 37 to 53%. In the present study, for patients who were strongly suspected of monogenic diseases but could not be diagnosed after WES, pediatric intensive care physicians and geneticists supplemented the patient's phenotype or made suggestions to add other detection methods (e.g., copy number variant sequencing, mitochondrial sequencing) through clinical/laboratory rounds and online communication. In this way, two patients were diagnosed. Case 148 with muscular hypotonia and progressive dyspnea received a diagnosis by mitochondrial genome sequencing. He was diagnosed with Leigh syndrome (OMIM: 256000). Case 206 with development delay

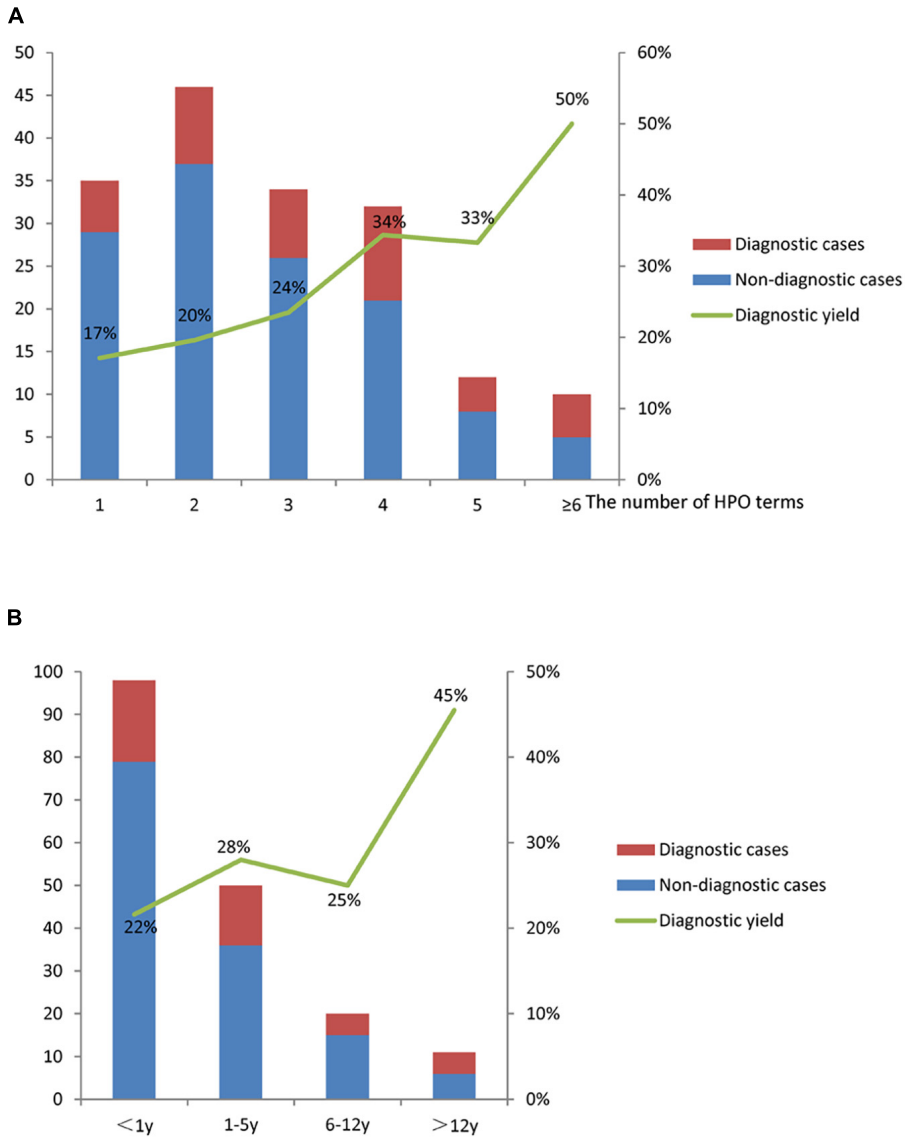



**FIGURE 3 |** The diagnostic yield of WES in the PICU. **(A)** The diagnostic yield increased with the number of HPO terms; **(B)** the diagnostic yields for different age groups were different.

was diagnosed through copy number variant sequencing (8q24.3 dup 3.6 Mb; 18q22.2-q23 del 9.9 Mb). In the future work, we will continue to strengthen clinical–laboratory interaction so as to improve the efficiency of diagnosis.

For the uncertain significance variants in this study, the following methods usually were taken to help improve the judgment result: (1) Supplement related tests, such as case 006, the final diagnosis was confirmed through supplementation of glucose phosphate isomerase activity test; (2) Test relatives of the proband with the same symptoms, such as case 041, the WES test was performed on the sister of the patient with the same symptoms, and there was a same variant, then the final diagnosis was given; (3) Perform functional verification (not covered in this article).

In addition, the diagnosis yield was also affected by socioeconomic and age factors. In this study, the diagnosis yield of children aged >12 years was the highest (45%), whereas that of children under 1 year was 22%, which was lower than that previously reported (Meng et al., 2017; Wang et al., 2020a). This may be due to the fact that more than half of the enrolled patients in this study were less than 1 year old and were treated in the PICU at the first onset, and their phenotypes were often atypical, mostly congenital heart disease and convulsions. Their parents were willing to find the etiology of the disease through relevant tests. Therefore, in this age group, the percentage of patients who received WES was high, but the diagnosis yield was relatively low. On the other hand, children who were >12 years old often had obvious phenotypes when they were hospitalized

and experienced the diagnostic odyssey outside. Thus, they had a higher chance to be diagnosed by WES.

Among the critically ill children, 48–81% of the diagnosed children had changes in clinical management (Willig et al., 2015; Meng et al., 2017; Wu et al., 2019; Wang et al., 2020a). In this study, the diagnosis of WES affected 30 cases (70%) in various aspects. First, WES had an impact on medical treatment: (1) **Implement any method, including drugs, surgery, and diet, while avoiding other ineffective or potentially harmful treatments**. For example, in case 006 who was diagnosed with *GPI*-related non-spherocytic hemolytic anemia, she was hospitalized many times in the past because of repeated jaundice. Splenectomy was performed after definite diagnosis. The available data showed that splenectomy was a common and effective treatment (Beutler et al., 1974; Fermo et al., 2019). Thereafter, the need for blood transfusions could be significantly reduced, and the quality of life of patients could be greatly improved. Although the prognosis is serious for some metabolic diseases, earlier diagnosis could lead to better results through diet control (Erez et al., 2011). (2) **Disease surveillance and evaluation of the system involved**. Single gene diseases often cause multi-system damage. After a clear diagnosis, through targeted disease surveillance, the affected organs could be identified. Once problems occur, interventions should be made in advance to achieve the purpose of long-term survival. For example, case 176 was diagnosed with sick sinus syndrome 1, and electrocardiogram and echocardiography should be monitored regularly. The genetic diagnosis also has impacts on the family: (1) **Change of attitude to treatment**. For patients with definite diagnosis, if the disease could be treated (including surgery, medicine, diet), parents are more confident to cooperate with doctors for treatment and regular re-examination. Even for patients with unclear diagnosis, some genetic diseases may be excluded. Thus, parents are urged to actively seek further diagnosis and treatment. In our study, roughly 20% of parents chose positive treatment. (2) **Withdrawing treatment**. For serious diseases that cannot be cured, parents have also found the causes of their children's illness. Some parents could consider earlier overall hospice care to alleviate their children's pain in the final stages of life and accompany them through the last days of their lives (Oberender and Tibballs, 2011). (3) **Familial genetic counseling**. Understanding the molecular cause of the patients greatly facilitated further genetic counseling, including follow-up reproductive planning and screening of siblings. Case 041 was diagnosed with glycogen storage disease II. His sister was later diagnosed, started on enzyme replacement therapy, and lives a normal life at present.

After receiving a definite diagnosis, some patients controlled their symptoms through drugs or diet and avoided ineffective treatment; through systemic assessment and regular monitoring of the disease state, some patients could be intervened in time to avoid life-threatening diseases, such as metabolic crisis; while some parents decided to withdraw the life support equipment. All of the above scenarios reduced the medical care costs from different aspects.

The limitation of this study is that the turnaround time of WES is relatively long. At present, some studies have reported the application of rapid WES in critically ill children (Carey et al., 2020; Freed et al., 2020; Wang et al., 2020b). In future works, we will continue to improve testing methods, optimize testing procedures, and reduce testing costs.

## CONCLUSION

A retrospective analysis was performed on the clinical application of WES in a single PICU of China in the past 3 years. Approximately a quarter of critically ill children with suspected monogenic diseases were diagnosed by WES. In addition, 70% of the patients had a positive impact on medical treatment and families. WES played a key role in family decision-making and clinical treatment, and may improve the prognosis of some children and reduce the economic burden on families and the society. Therefore, we recommend that critically ill children with suspected monogenic diseases in PICUs be tested for WES. With the optimization of the testing process and the reduction of the cost, an increasing number of patients in PICUs are expected to benefit from WES.

## DATA AVAILABILITY STATEMENT

According to national legislation/guidelines, specifically the Administrative Regulations of the People's Republic of China on Human Genetic Resources (http://www.gov.cn/zhengce/content/2019-06/10/content_5398829.htm, http://english.www.gov.cn/policies/latest_releases/2019/06/10/content_281476708945462.htm), no additional raw data is available at this time. Data of this project can be accessed after an approval application to the China National Genebank (CNGB, https://db.cngb.org/cnsa/). Please refer to https://db.cngb.org/, or email: CNGBdb@cngb.org for detailed application guidance. The accession code HRA001063 should be included in the application.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Institutional Review Board of the Beijing Children's Hospital. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

## AUTHOR CONTRIBUTIONS

YL designed the analysis plan, collected and analyzed the data, drafted the manuscript, and critically reviewed and revised the manuscript for important intellectual content. CH made substantial contributions to the design of the study, coordinated and supervised the data collection, and critically reviewed and revised the manuscript for important intellectual content. XH, RG, JG, and ZQ led the analysis of whole exome sequencing, interpreted results, and critically reviewed and revised the manuscript for important intellectual content. KL, HG, JZ, JL, ZL, and XJ led the clinical treatment of the patients, interpreted

the results, and critically reviewed and revised the manuscript for important intellectual content. WL and SQ conceptualized and designed the study, coordinated and supervised the data collection, and critically reviewed the manuscript for important intellectual content. All authors approved the final manuscript as submitted and agreed to be accountable for all aspects of the work.

## FUNDING

This work was funded by the Ministry of Science and Technology of China (2016YFC1000306), the CAMS Innovation Fund for Medical Sciences (2019-I2M-5-026), the Beijing Municipal Commission of Health and Family Planning Foundation (2018-2-1141 and 2020-4-1144), the Special Fund of the Pediatric Medical Coordinated Development Center of Beijing Hospitals Authority (XTCX201807), and the Beihang University and Capital Medical University Advanced Innovation Center for Big Data-Based Precision Medicine Plan (BHME-201905).

## ACKNOWLEDGMENTS

We thank all the participating families in this study.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.677699/full#supplementary-material

## REFERENCES

Australian Genomics Health Alliance Acute Care Flagship, Lunke, S., Eggers, S., Wilson, M., Patel, C., Barnett, C. P., et al. (2020). Feasibility of ultra-rapid exome sequencing in critically Ill infants and children with suspected monogenic conditions in the australian public health care system. *JAMA* 323, 2503–2511. doi: 10.1001/jama.2020.7671

Beutler, E., Sigalove, W. H., Muir, W. A., Matsumoto, F., and West, C. (1974). Glucosephosphate-isomerase (GPI) deficiency: GPI elyria. *Ann. Intern. Med.* 80, 730–732. doi: 10.7326/0003-4819-80-6-730

Carey, A. S., Schacht, J. P., Umandap, C., Fasel, D., Weng, C., Cappell, J., et al. (2020). Rapid exome sequencing in PICU patients with new-onset metabolic or neurological disorders. *Pediatr. Res.* 88, 761–768. doi: 10.1038/s41390-020-0858-x

Cunniff, C., Carmack, J. L., Kirby, R. S., and Fiser, D. H. (1995). Contribution of heritable disorders to mortality in the pediatric intensive care unit. *Pediatrics* 95, 678–681.

Erez, A., Nagamani, S. C., and Lee, B. (2011). Argininosuccinate lyase deficiency-argininosuccinic aciduria and beyond. *Am. J. Med. Genet. C Semin. Med. Genet.* 157C, 45–53. doi: 10.1002/ajmg.c.30289

Farnaes, L., Hildreth, A., Sweeney, N. M., Clark, M. M., Chowdhury, S., Nahas, S., et al. (2018). Rapid whole-genome sequencing decreases infant morbidity and cost of hospitalization. *NPJ Genom. Med.* 3:10. doi: 10.1038/s41525-018-0049-4

Fermo, E., Vercellati, C., Marcello, A. P., Zaninoni, A., Aytac, S., Cetin, M., et al. (2019). Clinical and molecular spectrum of glucose-6-phosphate isomerase deficiency. Report of 12 new cases. *Front. Physiol.* 10:467. doi: 10.3389/fphys.2019.00467

Freed, A. S., Clowes Candadai, S. V., Sikes, M. C., Thies, J., Byers, H. M., Dines, J. N., et al. (2020). The impact of rapid exome sequencing on medical management of critically Ill children. *J. Pediatr.* 226, 202–212. doi: 10.1016/j.jpeds.2020.06.020

French, C. E., Delon, I., Dolling, H., Sanchis-Juan, A., Shamardina, O., Megy, K., et al. (2019). Whole genome sequencing reveals that genetic conditions are frequent in intensively ill children. *Intensive Care Med.* 45, 627–636. doi: 10.1007/s00134-019-05552-x

Gubbels, C. S., Vannoy, G. E., Madden, J. A., Copenheaver, D., Yang, S., Wojcik, M. H., et al. (2020). Prospective, phenotype-driven selection of critically ill neonates for rapid exome sequencing is associated with high diagnostic yield. *Genet. Med.* 22, 736–744. doi: 10.1038/s41436-019-0708-6

Guo, J., Li, Z., Hao, C., Guo, R., Hu, X., Qian, S. et al. (2019). A novel de novo CASZ1 heterozygous frameshift variant causes dilated cardiomyopathy and left ventricular noncompaction cardiomyopathy. *Mol. Genet Genomic Med.* 7:e828. doi: 10.1002/mgg3.828

Hu, X., Guo, R., Guo, J., Qi, Z., Li, W., and Hao, C. (2020). Parallel tests of whole exome sequencing and copy number variant sequencing increase the diagnosis yields of rare pediatric disorders. *Front. Genet.* 11:473. doi: 10.3389/fgene.2020.00473

Hu, X., Li, N., Xu, Y., Li, G., Yu, T., Yao, R. E., et al. (2018). Proband-only medical exome sequencing as a cost-effective first-tier genetic diagnostic test for patients without prior molecular tests and clinical diagnosis in a developing country: the China experience. *Genet. Med.* 20, 1045–1053. doi: 10.1038/gim.2017.195

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

Mestek-Boukhibar, L., Clement, E., Jones, W. D., Drury, S., Ocaka, L., Gagunashvili, A., et al. (2018). Rapid paediatric sequencing (RaPS): comprehensive real-life workflow for rapid diagnosis of critically ill children. *J. Med. Genet.* 55, 721–728. doi: 10.1136/jmedgenet-2018-105396

Oberender, F., and Tibballs, J. (2011). Withdrawal of life-support in paediatric intensive care–a study of time intervals between discussion, decision and death. *BMC Pediatr.* 11:39. doi: 10.1186/1471-2431-11-39

Petrikin, J. E., Cakici, J. A., Clark, M. M., Willig, L. K., Sweeney, N. M., Farrow, E. G., et al. (2018). The NSIGHT1-randomized controlled trial: rapid whole-genome sequencing for accelerated etiologic diagnosis in critically ill infants. *NPJ Genom. Med.* 3:6. doi: 10.1038/s41525-018-0045-8

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

Sanford, E. F., Clark, M. M., Farnaes, L., Williams, M. R., Perry, J. C., Ingulli, E. G., et al. (2019). Rapid whole genome sequencing has clinical utility in children in the PICU. *Pediatr. Crit. Care Med.* 20, 1007–1020. doi: 10.1097/PCC.0000000000002056

Smith, L. D., Willig, L. K., and Kingsmore, S. F. (2015). Whole-Exome sequencing and whole-genome sequencing in critically Ill neonates suspected to have single-gene disorders. *Cold Spring Harb. Perspect. Med.* 6:a023168. doi: 10.1101/cshperspect.a023168

Stark, Z., Tan, T. Y., Chong, B., Brett, G. R., Yap, P., Walsh, M., et al. (2016). A prospective evaluation of whole-exome sequencing as a first-tier molecular test in infants with suspected monogenic disorders. *Genet. Med.* 18, 1090–1096. doi: 10.1038/gim.2016

Swaggart, K. A., Swarr, D. T., Tolusso, L. K., He, H., Dawson, D. B., and Suhrie, K. R. (2019). Making a genetic diagnosis in a level IV Neonatal intensive care unit population: who, when, how, and at what cost? *J. Pediatr.* 213:e214. doi: 10.1016/j.jpeds.2019.05.054

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

Wang, H., Lu, Y., Dong, X., Lu, G., Cheng, G., and Qian, Y. (2020a). Optimized trio genome sequencing (OTGS) as a first-tier genetic test in critically ill infants: practice in China. *Hum. Genet.* 139, 473–482. doi: 10.1007/s00439-019-02103-8

Wang, H., Qian, Y., Lu, Y., Qin, Q., Lu, G., Cheng, G., et al. (2020b). Clinical utility of 24-h rapid trio-exome sequencing for critically ill infants. *NPJ Genom. Med.* 5:20. doi: 10.1038/s41525-020-0129-0

Willig, L. K., Petrikin, J. E., Smith, L. D., Saunders, C. J., Thiffault, I., Miller, N. A., et al. (2015). Whole-genome sequencing for identification of Mendelian disorders in critically ill infants: a retrospective analysis of diagnostic and clinical findings. *Lancet Respir. Med.* 3, 377–387. doi: 10.1016/S2213-2600(15)00139-3

Wu, E. T., Hwu, W. L., Chien, Y. H., Hsu, C., Chen, T. F., Chen, N. Q., et al. (2019). Critical trio exome benefits in-time decision-making for pediatric patients with severe illnesses. *Pediatr. Crit. Care Med.* 20, 1021–1026. doi: 10.1097/PCC.0000000000002068

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:** 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.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

*Copyright © 2021 Liu, Hao, Li, Hu, Gao, Zeng, Guo, Liu, Guo, Li, Qi, Jia, Li and Qian. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*

frontiers in Genetics

BRIEF RESEARCH REPORT
published: 20 September 2021
doi: 10.3389/fgene.2021.707411

# Identification of a Rare Case With Nagashima-Type Palmoplantar Keratoderma and 18q Deletion Syndrome *via* Exome Sequencing and Low-Coverage Whole-Genome Sequencing

*Qianqian Li†, Xiaofan Zhu†, Conghui Wang, Jingjing Meng, Duo Chen and Xiangdong Kong**

*Genetics and Prenatal Diagnosis Center, Department of Obstetrics and Gynecology, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, China*

OPEN ACCESS

**Edited by:**
*Kathleen M. Gorman, Temple Street Children's University Hospital, Ireland*

**Reviewed by:**
*Caterina Marconi, Hôpitaux Universitaires de Genève, Switzerland*
*Ni-Chung Lee, National Taiwan University, Taiwan*

***Correspondence:**
*Xiangdong Kong*
*kongxd@263.net*

*† These authors have contributed equally to this work*

**Specialty section:**
*This article was submitted to Genetics of Common and Rare Diseases, a section of the journal Frontiers in Genetics*

**Received:** *10 May 2021*
**Accepted:** *16 August 2021*
**Published:** *20 September 2021*

**Citation:**
*Li Q, Zhu X, Wang C, Meng J, Chen D and Kong X (2021) Identification of a Rare Case With Nagashima-Type Palmoplantar Keratoderma and 18q Deletion Syndrome via Exome Sequencing and Low-Coverage Whole-Genome Sequencing. Front. Genet. 12:707411. doi: 10.3389/fgene.2021.707411*

Nagashima-type palmoplantar keratoderma (NPPK) is characterized by non-progressive, diffuse, and cross-gradient hyperkeratosis caused by mutations in the *SERPINB7* gene on chromosome 18q21.33. Chromosome 18q deletion syndrome (18q- syndrome) is a terminal deletion or microdeletion syndrome characterized by intellectual disability and congenital malformations. This paper describes an 18-year-old man with palmoplantar keratoderma and diffuse white matter abnormalities in the brain. Trio-based exome sequencing (ES) revealed a suspected mosaic compound heterozygous mutation for c.796C>T (p.Arg266*) in exon 8 inherited from the mother and a *de novo* exons 4–6 deletion of *SERPINB7*. Additional copy number variant (CNV) analysis of the ES data indicated a heterozygous gross deletion of 18q22.3-q23. The two *SERPINB7* gene variants were verified by Sanger sequencing and quantitative real-time polymerase chain reaction (qRT-PCR). Finally, low-coverage whole-genome sequencing (WGS) confirmed the 18q22.3-q23 deletion and additionally detected a mosaic 18q21.33-q22.3 deletion, together explaining NPPK and the neurological phenotypes of the proband. The gross deletion of all exons of *SERPINB7* was revealed for the first time. More rarely, c.796C>T (p.Arg266*) was likely to be mosaic, while the exon deletion was mosaic. In conclusion, the combination of multiple molecular genetic testing methods provides comprehensive informative molecular findings and promotes the diagnosis of complex diseases, as in this case.

**Keywords: Nagashima-type palmoplantar keratoderma, 18q deletion syndrome, exome sequencing, low-coverage whole-genome sequencing, *SERPINB7*, mosaicism**

## INTRODUCTION

Nagashima-type palmoplantar keratoderma (NPPK, MIM# 615598) is the most common type of palmoplantar keratoderma in Asian populations (Kubo, 2014), with a prevalence in Japan and China of 1.2/10000 and 3.1/10000, respectively (Kubo et al., 2013). The clinical manifestations of NPPK include mild diffuse palmoplantar hyperkeratosis, diffuse erythema with clear boundaries

on the dorsum of the hands, feet, forearm, and elbow, and Achilles tendon and knee. NPPK is caused by homozygous or compound heterozygous mutations in the serpin family B member 7 gene (*SERPINB7*) (Hannula-Jouppi et al., 2020) mapping on chromosome 18q21.33. To date, 14 variants of *SERPINB7* have been incorporated into HGMD® Professional 2021.2 (**Supplementary Table 1**). More than 90% of patients with NPPK carry the founder nonsense mutation c.796C>T (p.Arg266*). Gross deletions, insertions, complex re-arrangements, and repeats have not yet been reported.

Chromosome 18q deletion syndrome (18q- syndrome, MIM# 601808), a rare autosomal-dominant disorder, is a terminal deficiency or microdeletion syndrome characterized by intellectual disability and congenital malformations (Finley et al., 1969). The clinical symptoms are highly variable, including cognitive impairment from normal intelligence to severe intellectual disability, diffuse white matter abnormalities of the brain (Linnankivi et al., 2003), short stature, delayed myelination (Linnankivi et al., 2006), ear canal abnormalities (Feenstra et al., 2011), genital abnormalities, and foot deformities (Versacci et al., 2005).

This paper presents a rare case of NPPK and diffuse white matter abnormalities in the brain, with dual diagnosis of a suspected mosaic *SERPINB7* gene mutation and an exon deletion along with 18q deletion syndrome *via* trio-based exome sequencing (ES) and low-coverage whole-genome sequencing (WGS).

## MATERIALS AND METHODS

### Ethics Statement

An affected man with palmoplantar keratoderma and brain white matter abnormality and his family provided written informed consent for genetic studies. The study was approved by the appropriate local institutional review boards on human subject research at the First Affiliated Hospital of Zhengzhou University and conformed to the guidelines set forth by the Declaration of Helsinki.

### DNA Extraction

Genomic DNA from peripheral blood was obtained from 500 μL of whole blood using the Lab-Aid Nucleic Acid (DNA) Isolation Kit (Zeesan, Xiamen, China) in accordance with the manufacturer's instructions. Genomic DNA from other types of samples, including hair follicle cells, oral swabs, and urine, was extracted using the QIAamp DNA Blood & Tissue Kit (Qiagen, Germany).

### Quantitative Fluorescent Polymerase Chain Reaction

The genetic relationship of the family members was confirmed by quantitative fluorescent polymerase chain reaction (QF-PCR) using the GoldeneyeTM DNA ID System 20A Kit (Peoplespot, Beijing, China).

### Trio-Based Exome Sequencing

Trio-based exome sequencing (trio-ES) was performed using Illumina library construction and capture kits (Illumina, San Diego, CA, United States) in accordance with standard instructions (Document# 1000000048601v03), and 150 bp pair-end sequencing was conducted on NovaSeq 6000 (Illumina).

### Mapping, Variant Calling, and Variant Annotation

The Efficient Genosome Interpretation System (EGIS; Sierra Vast Bio-Medical, Shanghai, China) was used for mapping, variant calling, and variant annotation. The reads were aligned to the hg19/GRCh37 human reference genome sequence.

For single-nucleotide variants (SNVs) and small indels selecting, palmoplantar keratoderma (HP: 0000982) was entered into the EGIS, and 146 OMIM genes (**Supplementary Table 2**) were obtained based on the human phenotype ontology (Kohler et al., 2019). For CNV selecting, the bpCNV scan tool in the EGIS was used. The background library was constructed by calculating the correlation coefficient ($R > 0.94$) according to the average sequencing depth and exon fragment length of the target sample and reference samples (20 healthy subjects in the same batch). XHMM was used for exon and chromosome CNV calling in the ES data (Fromer et al., 2012). The copy number ratio of exon CNVs was obtained by dividing the exon reads per kilobase per million mapped reads (RPKM) value of the target sample by the average RPKM value of background library samples.

The potential impact of SNVs, small indels, and CNVs was investigated using Ensemble[1], RefSeq[2], and other databases, including OMIM[3], ClinVar[4], ClinGen[5], and DECIPHER[6]. The pathogenicity of all variations was evaluated in accordance with the latest guidelines of the American College of Medical Genetics and Genomics (ACMG; Riggs et al., 2020) and the ClinGen Sequence Variant Interpretation (SVI) Workgroup. Finally, polymerase chain reaction (PCR) amplification and Sanger sequencing were performed to confirm the SNV in *SERPINB7* by using the primer pairs listed in **Supplementary Table 3**.

### Quantitative Real-Time PCR (qRT-PCR)

The DNA samples, including those obtained from the proband, parents, and a healthy control subject (unrelated to this family without any skin diseases), were diluted to 50 ng $\mu L^{-1}$. qRT-PCR was conducted with KAPA SYBR FAST qPCR Kits (KK4601, Roche, Salt River Cape Town, South Africa) on the ABI QuantStudio5 Real-Time PCR System using the primer pairs exhibited in **Supplementary Table 3**. Human *GAPDH* was used as a reference gene, and the f fold-change was calculated using the $2^{-\Delta\Delta Ct}$ method.

[1]http://grch37.ensembl.org/index.html
[2]https://www.ncbi.nlm.nih.gov/refseq/
[3]https://omim.org/about
[4]https://www.ncbi.nlm.nih.gov/clinvar/
[5]https://www.clinicalgenome.org/
[6]https://www.Deciphergenomics.org/

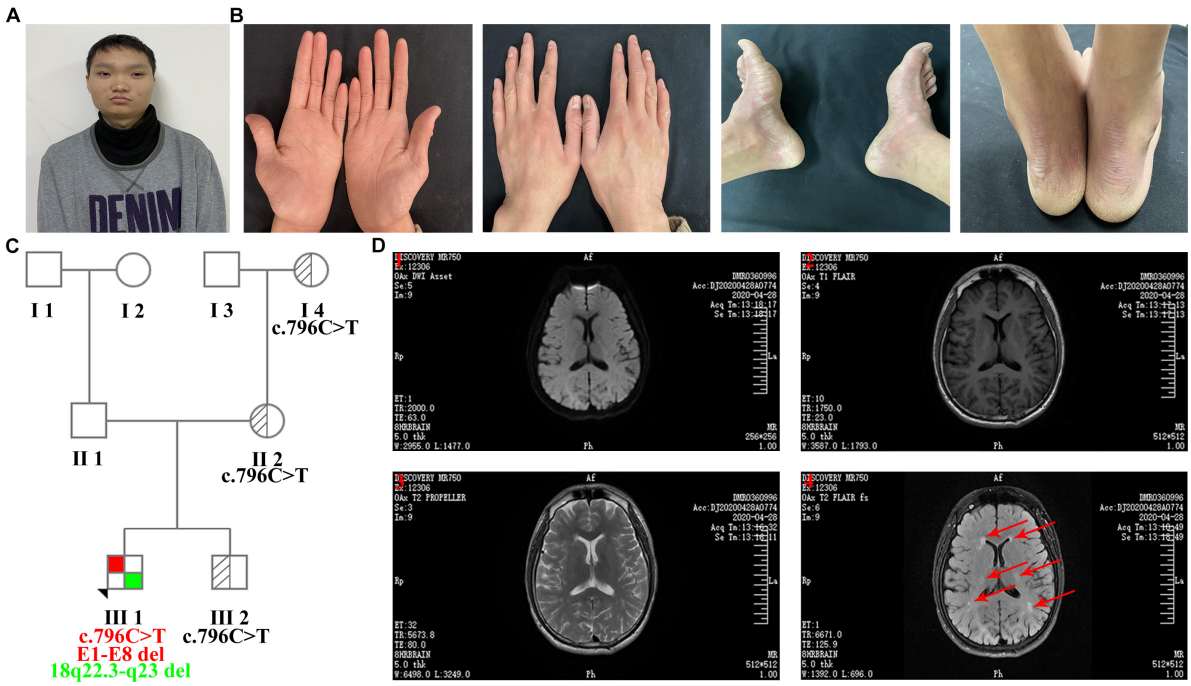



**FIGURE 1 |** Clinical features of the proband. **(A)** Facial features: wide-set eyes and short philtrum. **(B)** Bilateral erythema and hyperkeratosis of the hands and feet. **(C)** Family tree (E1-E8 del: exon 1-exon 8 deletion). **(D)** MRI images of transverse views. (1) Diffusion-weighted imaging (DWI); (2) T1 weighted image (T1WI); (3) T2 weighted image (T2WI); (4) Water image. MRI images indicated that the proband had multiple abnormal signals in the bilateral frontal and parietal lobes, and red arrows represent the regions of brain white matter abnormality.

## Low-Coverage Whole-Genome Sequencing

The experimental methods and data analysis of low-coverage WGS have been previously described (Wang et al., 2018). CNVs (GRCh37.p13) were analyzed and queried against public databases, including DGV[7], gnomAD[8], DECIPHER, OMIM, UCSC[9], and ClinGen. Pathogenicity was assessed according to the latest guidelines outlined by the ACMG (Riggs et al., 2020).

$CopyN_{bin}$ is the product of $obsRC_{bin}$ and $refRC_{bin}$ multiplied by $copyN_{chrom}$ ($CopyN_{bin}$, copy number of a bin; $obsRC_{bin}$, read number of the observed sample in this bin; $refRC_{bin}$, read number of the reference sample in this bin; $copyN_{chrom}$, theoretical copy number of the chromosome in this bin). Finally, the mosaic level was assessed based on the copy number value of the chromosomes.

# RESULTS

## Clinical Report

The proband was an 18-year-old man with slight facial abnormalities (wide-set eyes and short philtrum) (**Figure 1A**), mild intellectual disability, and developmental delay. The typical clinical symptoms of mild diffuse palmoplantar hyperkeratosis and diffuse erythema with clear boundaries on the dorsum of the hands and feet and Achilles tendon appeared in the first year after birth (**Figure 1B**). The palms and soles of the feet had a tendency to peel, and the latter were accompanied by a foot odor. The skin on the sole of the foot was thick and hard, and the skin on the palm of the hand was slightly thinner. These symptoms have been observed since palmoplantar hyperkeratosis was observed. At the age of 10 years, the proband was diagnosed with white matter abnormalities in the brain *via* magnetic resonance imaging (MRI) (data not shown). At the age of 17 years, the proband underwent a brain MRI scan (ID: DMR0360996) in our hospital. MRI images indicated that the proband had multiple abnormal signals in the bilateral frontal and parietal lobes (**Figure 1D**), which were similar to the previous MRI images, indicating that the abnormal areas of the white matter did not increase. Electrocardiogram (ECG) preformed in the same year revealed that the proband had sinus bradycardia (**Supplementary Figure 1**). The patient's parents, sibling, and other family members had no clinical manifestations. The family tree is presented in **Figure 1C**.

## Trio-ES Results

The results of QF-PCR confirmed the genetic relationship among the four family members (**Supplementary Tables 4,5** and **Supplementary Figures 2,3**). The quality control of trio-ES data is revealed in **Supplementary Table 6**. After filtering, a heterozygous variant of *SERPINB7* (NM_001040147)

[7]http://dgv.tcag.ca/dgv/app/home

[8]https://www.nature.com/collections/afbgiddede/gnomad-website

[9]http://genome.ucsc.edu/

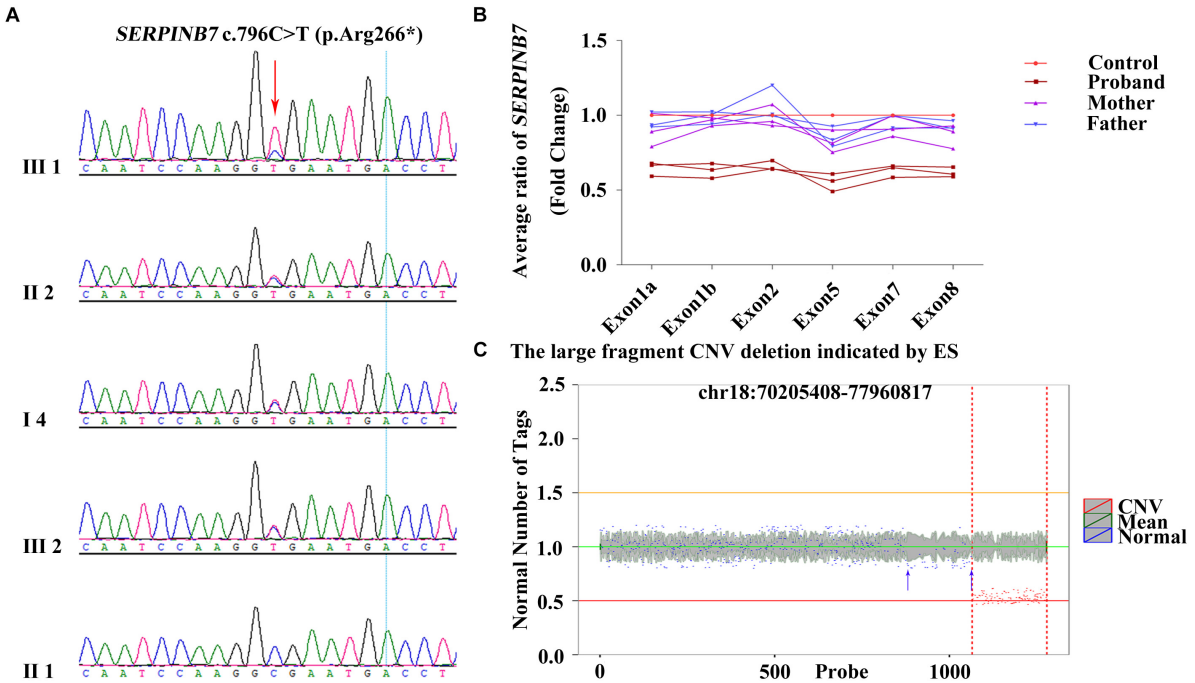



**FIGURE 2 |** c.796C>T (p.Arg266*) heterozygous variant and exon deletion of *SERPINB7* along with and 18q deletion identified by trio-ES. **(A)** Results of Sanger sequencing confirmed that the heterozygous variant c.796C>T (p.Arg266*) in the proband was inherited from the mother. **(B)** qRT-PCR results demonstrated that exons 1–8 deletion in the proband was a *de novo* deletion. **(C)** Large fragment chromosome CNV deletion indicated by trio-ES in the proband. *X*-axis: the position on the chromosome corresponding to the region of variation currently displayed. The red and blue marks indicate the areas of variant and normal, respectively. *Y*-axis: ratio of target sample RPKM value to the mean value of background library RPKM.

c.796C>T (p.Arg266*) in exon 8 (reference allele/alternative allele, ref/alt: 5/13) inherited from the mother (ref/alt: 12/28) was identified (Sheet 1 in **Supplementary Excel 1** and **Supplementary Figures 4–6**), and c.796C>T (p.Arg266*) was verified by Sanger sequencing (**Figure 2A**). Moreover, c.796C>T (p.Arg266*) was further identified in other types of samples from the proband, including hair follicle cells, oral swabs, and urine using Sanger sequencing, and the results were similar to those of the peripheral blood (**Supplementary Figure 7**). All *de novo* variants (irrespective of the candidate genes) in the proband identified by trio-ES are listed in **Supplementary Excel 2**.

By analyzing the gene exon-CNV data of trio-ES, the deletion of exons 4–6 of *SERPINB7* was indicated (Sheet 2 in **Supplementary Excel 1** and **Supplementary Table 7**). qRT-PCR was carried out to further confirm whether the exons of *SERPINB7* were deleted. The results of qRT-PCR confirmed that exons 1–8 were all deleted, which was a *de novo* deletion with an average ratio of 0.62 (**Figure 2B** and **Supplementary Table 7**). Consequently, c.796C>T (p.Arg266*) and exons 1–8 deletion in *SERPINB7* constituted a compound heterozygous state, leading to the occurrence of NPPK. Moreover, CNV analysis of the ES data indicated an approximately 7.75-Mb heterozygous deletion of 18q22.3-q23 (chr18:70205408-77960817) (**Figure 2C** and **Supplementary Table 8**), which was located in the key region of the 18q deletion syndrome, covering 28 coding genes, including 4 OMIM genes: *CYB5A*, *CTDP1*, *TSHZ1*, and *TXNL4A*. More interestingly, the ES-CNV scatter plot (**Figure 2C**, blue arrow) suggested that the deletion may be larger than 7.75 Mb, which covers the upstream region including *SERPINB7*. Furthermore, the copy number ratios were suggestive of a mosaic state in the upstream region.

## 18q21.33-q22.3 and 18q22.3-q23 Heterozygous Deletions Verified by Low-Coverage Whole-Genome Sequencing

An approximately 9.18-Mb mosaic heterozygous deletion of 18q21.33-q22.3 (chr18: 60480000-69660000) with a mosaic level of ∼40% (**Figure 3B** and **Supplementary Table 7**) and an 8.36-Mb heterozygous deletion of 18q22.3-q23 (chr18:69660000-78020000) were identified by low-coverage WGS (**Figure 3B** and **Supplementary Table 8**). However, no chromosomal abnormalities were found in the parents (**Figure 3**), indicating that the gross heterozygous deletions of 18q21.33-q22.3 and 18q22.3-q23 in the proband were both *de novo*.

# DISCUSSION

Nagashima-type palmoplantar keratoderma is an autosomal recessive PPK, and c.796C>T (p.Arg266*) in *SERPINB7* is the most common disease-causing variant in the Asian population.

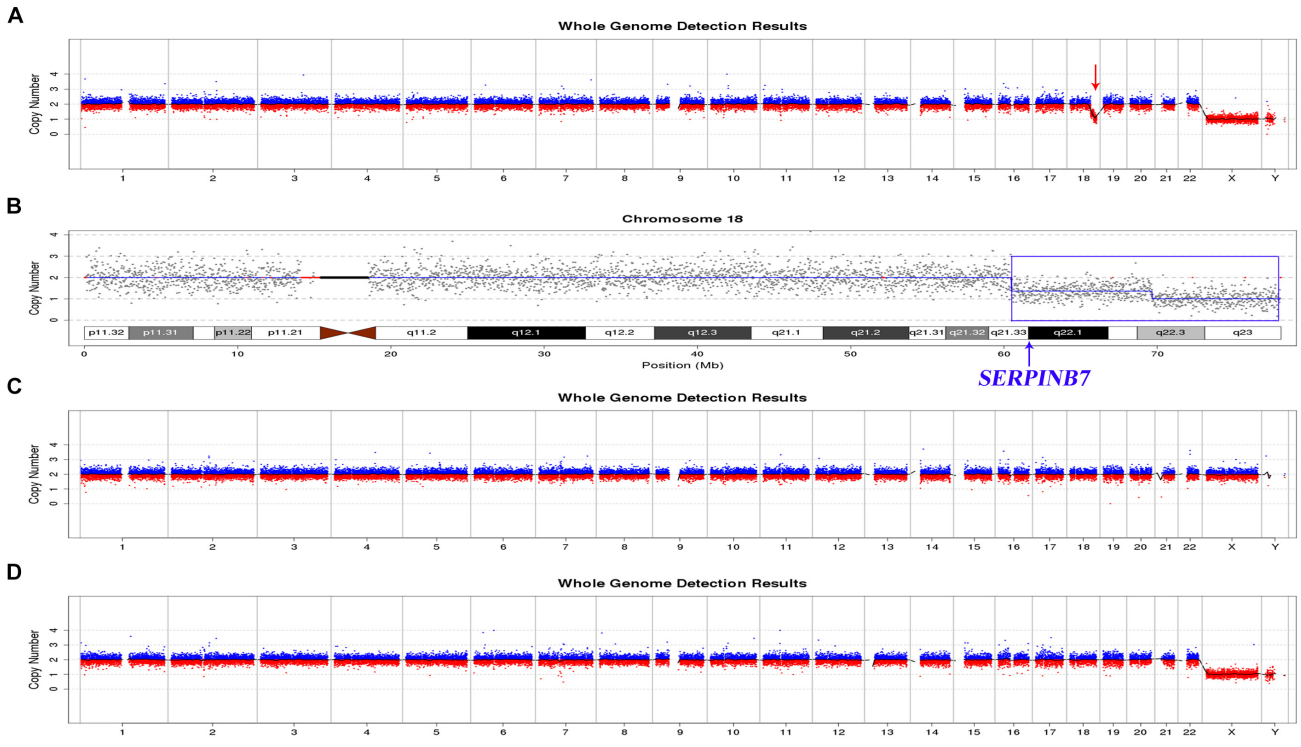



**FIGURE 3 |** Results of low-coverage WGS of the proband and parents. **(A)**, **(C)**, **(D)** Overview of the low-coverage WGS results of the proband, mother, and father, respectively. The red arrow indicates the region with abnormal copy numbers in the proband. The mother and the father both had normal copy numbers for all the chromosomes. **(B)** Low-coverage WGS results for chromosome 18 of the proband. The blue box shows the 18q21.33-q22.3 mosaic deletion and 18q22.3qter full deletion.

Other types of variants have been rarely reported. In the present family, the trio-ES results suggested that the proband carries the heterozygous variant c.796C>T (p.Arg266*) in exon 8 and deletion of exons 4–6 (confirmed as deletion of all exons by qRT-PCR), which may constitute a compound heterozygous state causing NPPK.

An approximately 9.18-Mb mosaic heterozygous deletion of 18q21.33-q22.3 (chr18:60480000-69660000) with a mosaic level of ∼40% was confirmed by low-coverage WGS, and this region contained 23 OMIM genes, including *SERPINB7* (**Supplementary Figure 8**). The low-coverage WGS result indicated that the *SERPINB7* whole gene deletion was part of the larger 9.18-Mb mosaic deletion. Moreover, the large CNV (18q22.3-q23 deletion at chr18:70205408-77960817) identified by trio-ES was confirmed by low-coverage WGS (deletion at 18:69660000-78020000), although the boundaries of the CNV regions were slightly different.

The ES ref/alt ratio and the Sanger sequencing results are not conclusive about the mosaic state of c.796C>T (p.Arg266*), since the decreased representation of mutant alleles could be due to technical reasons. However, the c.796C>T (p.Arg266*) homozygous variant should be detected if the whole exons of *SERPINB7* are deleted. The fact that the *SERPINB7* deletion is mosaic (for which on the contrary, the data are convincing) could explain the fact that some wild-type (wt) alleles were detected. This indicates that c.796C>T (p.Arg266*) is probably present in a mosaic form. An important limitation of this study is that no direct evidence could prove the mosaic state of c.796C>T (p.Arg266*). Considering that c.796C>T (p.Arg266*) was inherited from the mother, we speculated that three cell populations of wt + wt, wt + deletion, and wt + c.796C>T (p.Arg266*) were present in the proband. However, this would not represent a compound heterozygosity and, therefore, would not explain the phenotype of NPPK. In addition, for the proband, the heterozygosity states of c.796C>T (p.Arg266*) in hair follicle cells, oral swabs, and urine were similar to those in the peripheral blood (**Supplementary Figure 7**). Therefore, c.796C>T (p.Arg266*) may be mosaic, and together with the mosaic exon deletion, result in NPPK in the proband. However, if the mechanism by which a germline-inherited variant of c.796C>T (p.Arg266*) can become mosaic does exist, it would be much less likely than a situation in which the inherited c.796C>T (p.Arg266*) is present in all cells (heterozygous) and only the deletion at chr18:60480000-69660000 is in a mosaic state. Although we cannot demonstrate this, it is a possible scenario.

In this study, the gross deletion of eight exons in *SERPINB7* was revealed for the first time. The observation that this exon deletion in *SERPINB7* was mosaic is highly rare. Although c.796C>T (p.Arg266*) may also be mosaic, there was no direct evidence to prove its mosaic form. Moreover, the exact mosaic level of the 9.18-Mb heterozygous deletion will need to be

investigated further in future studies, such as fluorescence *in situ* hybridization. Although the autosomal recessive inheritance pattern of one of the SNV in one allele and the other of CNV in other allele has been reported in other types of diseases (Miksch et al., 2005; Miousse et al., 2011), this has yet to be reported for NPPK.

Furthermore, the proband was found to have *de novo* deletions of the 18q21.33-q22.3 region (chr18:60480000-69660000) in a mosaic state and the 18q22.3qter region (chr18:69660000-78020000) in full form by low-coverage WGS. The DECIPHER database includes similar cases with overlapping deletions. Case ID (272505) (chr18:60679560-68497696) presented global developmental delay, protruding ears, and short stature, whereas no phenotype was recorded for Case ID (254160, *de novo*) (chr18:63567516-70799806). Case ID (248930, *de novo*) (chr18:69191420-78014582) showed small nails, abnormal toes, and frontal bossing. Case ID (267134, *de novo*) (chr18:69199778-77982186) showed delayed puberty, proportionate short stature, ptosis, and stenosis of the external auditory canal. However, the DGV database does not include cases with similar CNV fragment deletions (chr18:60480000-69660000 and chr18:69660000-78020000). Therefore, it is possible that the two deletions might together contribute to the neurological phenotype of the proband. Interestingly, the 18q22.3qter deletion was indicated by the ES data, but the 18q21.33-q22.3 mosaic deletion was missed. Although the scatter plot of this region (**Figure 2C**) was slightly skewed from the mean copy number, this finding suggests that low-coverage WGS is more sensitive in the detection of CNV mosaicism (> 30%) than ES. In combination with the *SERPINB* gene variants, the results of low-coverage WGS provide accurate information for the genetic counseling and management of the patient, including the possibility of prenatal and pre-implantation diagnosis in the future.

Exome sequencing is a state-of-the-art method that enables the direct assessment of variants in the protein-coding region, which has been successfully used as a diagnostic approach to investigate the underlying genetic etiology of complex phenotypes (Tammimies et al., 2015). CNVs range in size from changes of a few hundred base pairs to enlargements or deletions of millions of base pairs of DNA (Conrad et al., 2010). Large CNVs are detected in severe pediatric cases, including neurological and congenital birth defects and neuropsychiatric disorders. ES may indicate the existence of large fragment deletions, and low-coverage WGS can verify the results (Bradshaw et al., 2018) and even uncover additional findings, as shown in the case presented in this study. The combined use of ES and low-coverage WGS may provide comprehensive genetic information for the molecular diagnosis of complex single- or multi-system diseases.

In conclusion, we report a rare case of NPPK and brain white matter abnormalities caused by *SERPINB7* mutations and 18q deletion syndrome. This study demonstrated that the combination of multiple genetic testing methods allowed for the accurate diagnosis of complex diseases.

## DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: the data supporting the findings of this research are available within the article/**Supplementary Material**.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the appropriate local institutional review boards on human subject research at the First Affiliated Hospital of Zhengzhou University and also conformed to the guidelines set forth by the Declaration of Helsinki. The proband and his family provided written informed consent to participate in this research. The project was conducted in accordance with the International Ethical Guidelines for Biomedical Research Involving Human Subjects (CIOMS). 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) for the publication of any potentially identifiable images or data included in this manuscript.

## AUTHOR CONTRIBUTIONS

QL, XZ, and XK designed the research and wrote the manuscript. QL, XZ, and JM carried out the experiments and data analysis. CW and DC revised the manuscript. All authors contributed to the article and approved the submitted version.

## FUNDING

National Natural Science Foundation of China (No. 82000321).

## ACKNOWLEDGMENTS

We were thankful for the family who participated and assisted us in the study.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.707411/full#supplementary-material

**Supplementary Excel 1 |** Sheet 1: Detailed information of c.796C>T (p.Arg266*) in *SERPINB7* identified by trio-ES. Sheet 2: Exon-CNV results of *SERPINB7* for the proband, mother, and father identified by trio-ES. Left: For the proband, exons 4–6 may be deleted. Middle: For mother, exon 5 may be duplicated. Right: For father, the copy number ratio may be normal.

**Supplementary Excel 2 |** All *de novo* variants (irrespective of candidate genes) in the proband identified by trio-ES.

## REFERENCES


Bradshaw, G., Lualhati, R. R., Albury, C. L., Maksemous, N., Roos-Araujo, D., Smith, R. A., et al. (2018). Exome sequencing diagnoses X-linked moesin-associated immunodeficiency in a primary immunodeficiency case. *Front. Immunol.* 9:420. doi: 10.3389/fimmu.2018.00420

Conrad, D. F., Pinto, D., Redon, R., Feuk, L., Gokcumen, O., Zhang, Y., et al. (2010). Origins and functional impact of copy number variation in the human genome. *Nature* 464, 704–712. doi: 10.1038/nature08516

Feenstra, I., Vissers, L. E., Pennings, R. J., Nillessen, W., Pfundt, R., Kunst, H. P., et al. (2011). Disruption of teashirt zinc finger homeobox 1 is associated with congenital aural atresia in humans. *Am. J. Hum. Genet.* 89, 813–819. doi: 10.1016/j.ajhg.2011.11.008

Finley, S. C., Cooper, M. D., Finley, W. H., Uchida, I. A., Noto, T. A., and Roddam, R. F. (1969). Immunological profile in a chromosome 18 deletion syndrome with IgA deficiency. *J. Med. Genet.* 6, 388–393. doi: 10.1136/jmg.6.4.388

Fromer, M., Moran, J. L., Chambert, K., Banks, E., Bergen, S. E., Ruderfer, D. M., et al. (2012). Discovery and statistical genotyping of copy-number variation from whole-exome sequencing depth. *J. Hum. Genet.* 91, 597–607. doi: 10.1016/j.ajhg.2012.08.005

Hannula-Jouppi, K., Harjama, L., Einarsdottir, E., Elomaa, O., Kettunen, K., Saarela, J., et al. (2020). Nagashima-type palmoplantar keratosis in Finland caused by a SERPINB7 founder mutation. *J. Am. Acad. Dermatol.* 83, 643–645. doi: 10.1016/j.jaad.2019.11.004

Kohler, S., Carmody, L., Vasilevsky, N., Jacobsen, J., Danis, D., Gourdine, J. P., et al. (2019). Expansion of the Human Phenotype Ontology (HPO) knowledge base and resources. *Nucleic Acids Res.* 47, D1018–D1027.

Kubo, A. (2014). Nagashima-type palmoplantar keratosis: a common Asian type caused by SERPINB7 protease inhibitor deficiency. *J. Invest. Dermatol.* 134, 2076–2079. doi: 10.1038/jid.2014.156

Kubo, A., Shiohama, A., Sasaki, T., Nakabayashi, K., Kawasaki, H., Atsugi, T., et al. (2013). Mutations in SERPINB7, encoding a member of the serine protease inhibitor superfamily, cause Nagashima-type palmoplantar keratosis. *Am. J. Hum. Genet.* 93, 945–956. doi: 10.1016/j.ajhg.2013.09.015

Linnankivi, T., Tienari, P., Somer, M., Kahkonen, M., Lonnqvist, T., Valanne, L., et al. (2006). 18q deletions: clinical, molecular, and brain MRI findings of 14 individuals. *Am. J. Med. Genet. A* 140, 331–339. doi: 10.1002/ajmg.a.31072

Linnankivi, T. T., Autti, T. H., Pihko, S. H., Somer, M. S., Tienari, P. J., Wirtavuori, K. O., et al. (2003). 18q-syndrome: brain MRI shows poor differentiation of gray and white matter on T2-weighted images. *J. Magn. Reson. Imaging* 18, 414–419. doi: 10.1002/jmri.10383

Miksch, S., Lumsden, A., Guenther, U. P., Foernzler, D., Christen-Zach, S., Daugherty, C., et al. (2005). Molecular genetics of pseudoxanthoma elasticum: type and frequency of mutations in ABCC6. *Hum Mutat.* 26, 235–248. doi: 10.1002/humu.20206

Miousse, I. R., Watkins, D., and Rosenblatt, D. S. (2011). Novel splice site mutations and a large deletion in three patients with the cblF inborn error of vitamin B12 metabolism. *Mol. Genet. Metab.* 102, 505–507. doi: 10.1016/j.ymgme.2011.01.002

Riggs, E. R., Andersen, E. F., Cherry, A. M., Kantarci, S., Kearney, H., Patel, A., et al. (2020). Technical standards for the interpretation and reporting of constitutional copy-number variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics (ACMG) and the Clinical Genome Resource (ClinGen). *Genet Med.* 22, 245–257. doi: 10.1038/s41436-019-0686-8

Tammimies, K., Marshall, C. R., Walker, S., Kaur, G., Thiruvahindrapuram, B., Lionel, A. C., et al. (2015). Molecular diagnostic yield of chromosomal microarray analysis and whole-exome sequencing in children with autism spectrum disorder. *JAMA* 314, 895–903. doi: 10.1001/jama.2015.10078

Versacci, P., Digilio, M. C., Sauer, U., Dallapiccola, B., and Marino, B. (2005). Absent pulmonary valve with intact ventricular septum and patent ductus arteriosus: a specific cardiac phenotype associated with deletion 18q syndrome. *Am. J. Med. Genet. A* 138, 185–186. doi: 10.1002/ajmg.a.30916

Wang, J., Chen, L., Zhou, C., Wang, L., Xie, H., Xiao, Y., et al. (2018). Prospective chromosome analysis of 3429 amniocentesis samples in China using copy number variation sequencing. *Am. J. Obstet. Gynecol.* 219, 281–287.


**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.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

*Copyright © 2021 Li, Zhu, Wang, Meng, Chen and Kong. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*

frontiers in Genetics

ORIGINAL RESEARCH
published: 27 September 2021
doi: 10.3389/fgene.2021.743230

# Characterization of *FMR1* Repeat Expansion and Intragenic Variants by Indirect Sequence Capture

*Valentina Grosso[1†], Luca Marcolungo[1†], Simone Maestri[1], Massimiliano Alfano[1], Denise Lavezzari[1], Barbara Iadarola[1], Alessandro Salviati[1,2], Barbara Mariotti[3], Annalisa Botta[4], Maria Rosaria D'Apice[5], Giuseppe Novelli[4,6,7], Massimo Delledonne[1,2‡] and Marzia Rossato[1,2*‡]*

[1]Department of Biotechnology, University of Verona, Verona, Italy, [2]GENARTIS srl, Verona, Italy, [3]Department of Medicine, Section of General Pathology, University of Verona, Verona, Italy, [4]Department of Biomedicine and Prevention, Medical Genetics Section, University of Rome "Tor Vergata", Rome, Italy, [5]Laboratory of Medical Genetics, Tor Vergata Hospital, Rome, Italy, [6]IRCCS Neuromed Mediterranean Neurological Institute, Pozzilli, Italy, [7]Department of Pharmacology, School of Medicine, University of Nevada, Reno, NV, United States

OPEN ACCESS

**Edited by:**
Alfredo Brusco,
University of Turin, Italy

**Reviewed by:**
Kishore Raj Kumar,
Garvan Institute of Medical Research, Australia
Nicole Ziliotto,
Humanitas University, Italy
Zhining Wen,
Sichuan University, China

***Correspondence:**
Marzia Rossato
marzia.rossato@univr.it

[†]These authors share first authorship
[‡]These authors share last authorship

**Specialty section:**
This article was submitted to Genetics of Common and Rare Diseases, a section of the journal Frontiers in Genetics

**Received:** 17 July 2021
**Accepted:** 26 August 2021
**Published:** 27 September 2021

**Citation:**
Grosso V, Marcolungo L, Maestri S, Alfano M, Lavezzari D, Iadarola B, Salviati A, Mariotti B, Botta A, D'Apice MR, Novelli G, Delledonne M and Rossato M (2021) Characterization of FMR1 Repeat Expansion and Intragenic Variants by Indirect Sequence Capture. Front. Genet. 12:743230. doi: 10.3389/fgene.2021.743230

Traditional methods for the analysis of repeat expansions, which underlie genetic disorders, such as fragile X syndrome (FXS), lack single-nucleotide resolution in repeat analysis and the ability to characterize causative variants outside the repeat array. These drawbacks can be overcome by long-read and short-read sequencing, respectively. However, the routine application of next-generation sequencing in the clinic requires target enrichment, and none of the available methods allows parallel analysis of long-DNA fragments using both sequencing technologies. In this study, we investigated the use of indirect sequence capture (Xdrop technology) coupled to Nanopore and Illumina sequencing to characterize *FMR1*, the gene responsible of FXS. We achieved the efficient enrichment (> 200×) of large target DNA fragments (~60–80 kbp) encompassing the entire *FMR1* gene. The analysis of Xdrop-enriched samples by Nanopore long-read sequencing allowed the complete characterization of repeat lengths in samples with normal, pre-mutation, and full mutation status (> 1 kbp), and correctly identified repeat interruptions relevant for disease prognosis and transmission. Single-nucleotide variants (SNVs) and small insertions/deletions (indels) could be detected in the same samples by Illumina short-read sequencing, completing the mutational testing through the identification of pathogenic variants within the *FMR1* gene, when no typical CGG repeat expansion is detected. The study successfully demonstrated the parallel analysis of repeat expansions and SNVs/indels in the *FMR1* gene at single-nucleotide resolution by combining Xdrop enrichment with two next-generation sequencing approaches. With the appropriate optimization necessary for the clinical settings, the system could facilitate both the study of genotype–phenotype correlation in FXS and enable a more efficient diagnosis and genetic counseling for patients and their relatives.

**Keywords: long fragment enrichment, indirect sequence capture, repeat expansion, single nucleotide variants, *FMR1***

## INTRODUCTION

The expansion of unstable short tandem repeats is the causal DNA mutation in almost 40 genetic human diseases (Paulson, 2018). This group includes neurological and neuromuscular disorders, such as fragile X syndrome (FXS; MIM# 300624), which is caused by the expansion of CGG trinucleotide repeats in the 5′ untranslated region of the *fragile X mental retardation 1* gene (*FMR1*; MIM# 309550; Eichler et al., 1994; Yrigollen et al., 2012; Nolin et al., 2014). Normal alleles carry 5–44 CGG repeats, whereas expanded alleles are classified as intermediate (45–54 repeats), pre-mutation (55–200 repeats), or full mutation (> 200 repeats). Females with pre-mutations have approximately a 20% risk for fragile X-associated primary ovarian insufficiency (FXPOI; MIM#311360). Older males and females with pre-mutations are at risk for fragile X-associated tremor/ataxia syndrome (FXTAS; MIM#300623). The pre-mutation allele often expands to a full mutation during female germline transmission, thus giving rise to FXS in the progeny. The risk of pre-mutation expansion depends mainly on the number of CGG repeats (with shorter alleles being less likely to expand to a full mutation than larger ones) and the presence of AGG interruptions in the tandem array. Such AGG interruptions increase repeat stability, reduce the risk of expansions (Eichler et al., 1994; Nolin et al., 2003; Yrigollen et al., 2012), and can modulate the disease phenotype (Matsuyama et al., 1999; Sakamoto et al., 2001; Charles et al., 2007; Braida et al., 2010). Moreover, recent evidence has suggested pronounced repeat variability between individuals and within them (mosaicism) that also modulates the disease phenotype (van Blitterswijk et al., 2013; Tabolacci et al., 2020). Similar mechanisms have been observed in the transmission/phenotype of related diseases, such as Myotonic Dystrophy type 1 and Huntington's disease (Rodriguez and Todd, 2019). Although much less frequent than microsatellite expansions, intragenic single-nucleotide variants (SNVs) and short insertions or deletions (indels) are significant mutational mechanisms leading to FXS and other repeat-associated diseases (Quartier et al., 2017). Accordingly, accurate risk prediction in genetic counseling not only requires the precise characterization of repeats, but also the mapping and counting of interruptions within the repeat array and the ability to map additional intragenic variants (Loomis et al., 2013).

Conventional diagnostic testing to assess repeat length involves triplet repeat primed PCR (TP-PCR) or Southern blotting (Spector et al., 2021). These methods are imprecise when dealing with long expansions, are severely limited in their ability to detect minor alleles, and lack single-nucleotide resolution (Warner et al., 1996; Nolin et al., 2003; Saluto et al., 2005; Filipovic-Sadic et al., 2010; Adler et al., 2011; Bastepe and Xin, 2015; Hayward et al., 2016; Ardui et al., 2018). More recently, third-generation sequencing technologies, such as Oxford Nanopore Technologies (ONT) and PacBio SMRT sequencing, have shown consistent benefits for the characterization of short tandem repeats in FXS and related disorders (McFarland et al., 2014, 2015; Tsai et al., 2017; Giesselmann et al., 2019; Mantere et al., 2019). These approaches can sequence DNA fragments several kbp in length, facilitating the accurate genotyping of repeat expansion alleles and the identification of interruptions and mosaicism (Tsai et al., 2017; Giesselmann et al., 2019). The combination of third-generation sequencing with enrichment strategies can reduce costs while ensuring sufficient coverage for accurate repeat characterization by focusing on the target site. In the first such report, *FMR1* repeat arrays were amplified by PCR for PacBio sequencing (Loomis et al., 2013). However, PCR is unsuitable in patients heterozygous for normal and large expansion alleles because only the normal allele may be amplified (Chakraborty et al., 2016), and polymorphisms surrounding the repeat region can lead to allele bias, dropout, or the misinterpretation of results (Bastepe and Xin, 2015).

More recently, both third-generation sequencing technologies have been coupled to an enrichment method based on CRISPR/Cas9, where Cas9 cuts at sites flanking the repeats allowing the ligation of sequencing adapters for the accurate characterization of repeat length, interruptions, and mosaicism in *FMR1* (Tsai et al., 2017). Although this removes the reliance on PCR, remaining limitations include the large amount of starting material required, typically 1–10 µg DNA (Gilpatrick et al., 2020; Stangl et al., 2020), which makes it difficult to work with low-abundant samples, as, for examples, those from prenatal/pre-implant testing or clinical biopsies. Moreover, sequencing is confined to a few kbp surrounding the repeat, thus preventing the analysis of mutations along the full length of the causative gene. Finally, the system lacks flexibility, because commonly utilized protocols to sequence the Cas9-enriched DNA rely only on long-read sequencing and not on short-read sequencing platforms, such as Illumina, which show higher accuracy. Despite recent improvements strongly increased long-read accuracy, ONT still fails at accurately detecting indels (Maestri et al., 2020), while PacBio High-Fidelity mode still requires the use of high-capacity SMRT cells, that makes the analysis very expensive when only few samples are multiplexed. These are critical drawbacks, especially when the repeat characterization is inconclusive and the analysis of the entire gene is necessary to identify other mutations, namely, SNVs or indels (Sitzmann et al., 2018). Although the analysis of tandem repeats using Illumina technology is challenging due to the large size and typically high GC content of the fragments, it has nevertheless proven valuable for the identification of causative intragenic variants in patients with a negative standard workup based on the analysis of repeat expansions (Quartier et al., 2017). To address these limitations and exploit the advantages of both short-read and long-read sequencing, we investigated the use of Xdrop technology (Samplix, Birkerød, Denmark) for the characterization of the *FMR1* locus. The approach uses so-called "indirect sequence capture" to enrich for long fragments (several kbp) starting with limited DNA input (10–15 ng). High-molecular-weight (HMW) DNA molecules (50–100 kbp) are initially encapsulated in individual droplets, and droplet PCR (dPCR) is used to amplify a detection sequence (DS) of 100–150 bp located near the target of interest. Positive droplets are revealed by staining with a DNA-intercalating dye and are recovered by flow sorting. A few hundred target DNA

molecules are recovered for multiple displacement amplification after their encapsulation in individual droplets (dMDA) to minimize amplification biases (Madsen et al., 2020; Blondal et al., 2021). We took advantage of Xdrop technology to enrich the *FMR1* locus and used ONT long-read sequencing to characterize the *FMR1* repeat length/features with parallel Illumina sequencing to determine the presence of intragenic variants within the *FMR1* gene body.

## MATERIALS AND METHODS

### DNA Samples

Genomic DNA (NA12878, NA06891, NA07537, and NA20241, representing cells with diverse *FMR1* alleles) was purchased from the Coriell Institute for Medical Research. All the other samples were isolated from the whole blood of unrelated healthy donors (Blood Center, Verona Hospital) following informed written consent. Venous blood samples were collected in EDTA tubes, de-identified immediately after collection, and stored at −80°C until use. The study was approved by the Ethics Committee for Clinical Research of Verona and Rovigo Provinces and all the investigations were conducted according to the Declaration of Helsinki. Genomic DNA was extracted using the Genomic Tip 100/G kit (Qiagen, Hilden, Germany), Nanobind CBB Big DNA Kit (Circulomics, Baltimore, MD, United States), NucleoSpin Blood Mini kit (Macherey-Nagel, Düren, Germany), or the Miller's protocol (Miller et al., 1988). All protocols were carried out according to the manufacturer's instructions, and for the Circulomics kit, we used either the HMW or ultra-HMW protocol. The different DNA extraction methods were tested on samples from distinct donors, an aspect that may represent a weakness of the study.

### Droplet Generation and dPCR

Before enrichment, DNA samples were purified using 1× HighPrep MagBio beads (MagBio Genomics, Gaithersburg, MD, United States) and diluted with DNase-free water to 5 ng/µl. Detection sequence-specific primers for *FMR1* enrichment were designed using the Samplix primer design tool[1]: forward primer 5'-GAG CCC TAG TCC TCA CCC AAT-3' and reverse primer 5'-CCC TAC CTA TCA GGC AAA GCT-3' (**Supplementary Figure S1**). The dPCR reaction consisted of 20 µl 2× dPCR mix (Samplix), 0.8 µl of each primer (10 µM), 2 µl 5 ng/µl DNA, and water to 40 µl. Droplets were generated using a dPCR cartridge and Xdrop droplet generator (both from Samplix). Droplets were then transferred to four tubes and dPCR was carried out by heating to 94°C for 2 min followed by 40 cycles of 94°C for 3 s and 60°C for 30 s at a ramping rate of 1.5°C/s.

### Positive Droplet Sorting

Following dPCR, droplets were collected in a single tube, diluted with 1 ml dPCR buffer (Samplix), and stained with 10 µl droplet dye (Samplix). Droplets were sorted on a FACS Aria Fusion II (Becton Dickinson, Franklin Lakes, NJ, United States), with instrument settings adjusted to FSC = 210, SSC = 250, and FL1 = 370. The positive droplets were gated on FL1 fluorescence and the sorting mode was set to "Yield." Sorted droplets were collected in 15 µl water.

### dMDA

Sorted droplets were mixed with 20 µl Break solution and 2 µl Break color (Samplix), and 10 µl of the resulting aqueous phase was used as a template for dMDA. The reaction mix consisted of 4 µl dMDA buffer, 1 µl dMDA enzyme, 10 µl template, and water to 20 µl. Droplets were generated as above, while running the dMDA program. Afterward, the droplets were incubated for 16 h at 30°C (lid at 75°C) followed by 10 min at 65°C to terminate the reaction. The dMDA droplets were broken using 20 µl Break solution and 1 µl Break color as above.

### qPCR Analysis

Total DNA released from dMDA droplets was quantified using a Qubit fluorimeter and the Qubit HS DNA quantification kit (Thermo Fisher Scientific, Waltham, MA, United States). The size range of the amplified DNA was analyzed on a TapeStation 4,150 using the Genomic DNA ScreenTape assay (both from Agilent Technologies, Santa Clara, CA, United States). Fold enrichment of target DNA was assessed by qPCR using the KAPA library Quant qPCR mix (Roche, Basel, Switzerland), 10 ng DNA, and 2 mM each of forward (5′-TCA TTG GTG GTC GGG TGT AC-3′) and reverse (5′-AGC GAC ACC TCA CAT TCC TT-3′) validation primers (**Supplementary Figure S1**). Fold enrichment was determined using an online calculator.[2] Usually, samples with ≥100-fold enrichment at qPCR showed also robust enrichment and breath of coverage after sequencing and thus were selected for downstream analysis.

### ONT Sequencing

We sequenced 1–1.5 µg of the enriched DNA samples from the Xdrop workflow using the ONT platform, pooling two replicates when necessary. Amplified DNA was initially debranched using 15 units of T7 endonuclease I in 30 µl for 15 min. Debranched DNA fragments were isolated by size selection using AmPure XP beads (Beckman Coulter, High Wycombe, United Kingdom) in the presence of 15% polyethylene glycol (Sigma–Aldrich, St Louis, MO, United States). The ONT sequencing library was generated using the Oxford Nanopore Ligation Sequencing Kit SQK-LSK109 (ONT, Oxford, United Kingdom) according to the manufacturer's instructions with minor modifications. Briefly, DNA was end-repaired using the NEBNext FFPE DNA Repair Mix (New England Biolabs, Ipswich, MA, United States) at 20°C for 10 min and subsequently end-prepped with the NEBNext End repair/dA-tailing Module (New England Biolabs) at 20°C for 20 min. Sequencing adapters were ligated at room temperature for 10 min. Finally, the 30–50

[1]https://samplix.com/primer

[2]https://samplix.com/calculations

fmol library was loaded into a MinION R9.4.1 flowcell (ONT) and standard settings were applied for a run time of ~16 h.

## ONT Data and Repeat Analysis

Base calling was applied to the raw ONT fast5 files using Guppy v4.2.2 in high-accuracy mode, with parameters "-r -i $FAST5_DIR -s $BASECALLING_DIR --flowcell FLO-MIN106 --kit SQK-LSK109." Reads were quality filtered using NanoFilt v2.7.1 (De Coster et al., 2018), with a minimum quality score of 7. Reads were then mapped to the hg38 human reference genome using Minimap2 v2.17-r941 (Li, 2018). The ONT datasets showed a large fraction of bases (59.3%) mapping as supplementary alignments within the same genomic region, but not recurrent at the same position, suggesting the presence of chimeric reads, possibly derived from dMDA as previously reported (Gawad et al., 2016; Zhou et al., 2020). To exploit the full sequencing dataset, ONT read mapping was therefore adjusted by also considering supplementary read alignments. Bedtools intersect v2.29.2 (Quinlan and Hall, 2010) was used to extract primary or supplementary alignments completely spanning the *FMR1* repetitive region defined in a bed file, containing repeat coordinates plus 400 bp flanking the repeat on each side (chrX:147911849–147,912,310). Sequences corresponding to alignments of interest were extracted in forward orientation from the bam alignment file using a combination of Samtools v1.10 (Li et al., 2009) and awk scripting language and were realigned to the hg38 reference using Minimap2. A combination of PcrClipReads and SamExtractClip from jvarkit v1f97a3401[3] and seqtk subseq v1.3-r106[4] was then used to trim the portions of sequences outside the bed file, allowing us to retrieve all sequences fully spanning the repeat, including supplementary alignments.

Repeat length was determined from consensus sequences obtained by the *de novo* assembly of the extracted sequences using the *CharONT* pipeline (**Supplementary Figure S2**). First, the sequences were clustered using VSEARCH v2.15.1_linux_x86_64 (Rognes et al., 2016) with an 85% minimum identity threshold. Reads in the most abundant cluster were then aligned to each other using MAFFT v7.475 (Katoh et al., 2002) with parameters "--auto –adjustdirectionaccurately." A draft consensus sequence was called using EMBOSS cons v6.6.6.0,[5] setting the "--plurality" parameter to the value obtained by multiplying the number of aligned reads by 0.15 (Maestri et al., 2019). This process generated a preliminary consensus sequence for one allele. All sequences were then mapped to the consensus sequence, and a bidimensional score was calculated for each sequence, extracting the size of the biggest DEL, and the biggest INS from the CIGAR string in the bam file. If soft clipping occurred, the length of the soft-clipped sequence contributed to the score calculation by exploiting the presence of flanking sequences. Candidate outliers were then identified (with either component of the score exceeding a predefined threshold based on the interquartile range of scores assigned to all sequences) and were excluded from the clustering process. Scores were used to cluster the sequences in two groups, corresponding to the two alleles, using the *k*-means function of the "stats" R package (R Core Team, 2013). Outliers with either component of the score exceeding a predefined threshold were then identified based on the interquartile range of scores assigned to sequences within the cluster and were saved to a new file. Sequences assigned to each allele were processed separately. Up to 200 sequences were randomly subsampled using seqtk sample, and a draft consensus sequence was called by combining MAFFT and EMBOSS cons, as previously described (Footnote 5). Another set of up to 200 reads was subsampled using seqtk sample to polish the draft consensus sequence, and read overlaps were found with Minimap2 (Li, 2018). Racon v1.4.13 (Vaser et al., 2017) was then used to perform a first round of polishing with parameters "-m 8 -x−6 -g−8 -w 500 --no-trimming." A second round of polishing was performed using the medaka_consensus program of Medaka v1.2.1[6] specifying the "r941_min_high_g360" model. The polished consensus sequences for each allele were finally searched for repeat motifs using Tandem Repeat Finder v4.09 (Benson, 1999). The scripts used to generate consensus sequences and repeat annotations are available online.[7]

The presence of somatic mosaicism was investigated by aligning reads to sequences flanking the repeat, searching for repeat motifs, and visualizing alignments in a genome browser using the *MosaicViewer_FMR1* pipeline (**Supplementary Figure S3**). The msa.sh and cutprimers.sh programs from BBMap suite v38.87 were used to trim one of the two sequences flanking the repeat expansion, and trimmed reads were aligned to the other flanking sequence using Minimap2. Alignments were visualized in the IGV genome browser v2.8.3 (Robinson et al., 2011). Mapped sequences were extracted from the bam file in the forward orientation using Samtools and a custom script, and the ID of reads in reverse orientation was extracted from the SAM flag. Extracted sequences were searched for repeats with the motif "CGG" using the NCRF script in the Noise-cancelling repeat finder package v1.01.02 (Harris et al., 2019) with parameters "--scoring = nanopore --minlength = 12 CGG_repeat:CGG --minmratio = 0.90 --stats = events –positionalevents." Repeats were sorted in a single repeat summary file using the scripts ncrf_cat.py, ncrf_sort.py, and ncrf_summary.py. Reads were then aligned to the flanking sequence using Minimap2 and visualized in the IGV genome browser. The scripts used to investigate somatic mosaicism are available online.[8]

## Illumina Sequencing

Amplified DNA was fragmented using a Covaris sonicator to achieve an average size of 400 bp, and Illumina PCR-free libraries were prepared from ~200–400 ng DNA using the KAPA Hyper prep kit and unique dual-indexed adapters (5 µl of a 15 µm stock) according to the supplier's protocol (Roche). The library concentration and size distribution were assessed on a Bioanalyzer (Agilent Technologies). Barcoded libraries were pooled at

[3]https://github.com/lindenb/jvarkit
[4]https://github.com/lh3/seqtk
[5]http://emboss.open-bio.org/rel/dev/apps/cons.html
[6]https://github.com/nanoporetech/medaka
[7]https://github.com/MaestSi/CharONT
[8]https://github.com/MaestSi/MosaicViewer_FMR1

equimolar concentrations and sequenced on a NovaSeq6000 instrument (Illumina, San Diego, CA, United States) to generate 150-bp paired-end reads.

## Illumina Data Analysis and Variant Calling

Illumina fastq files were quality checked using FastQC,[9] and low-quality nucleotides and adaptors were trimmed using fastp (Chen et al., 2018). Reads were then aligned to the reference human genome version GRCh38/hg38 using BWA-MEM v0.7.17.[10] All bam files were cleaned by local realignment around indel sites, followed by duplicate marking and recalibration using Genome Analysis Toolkit v3.8.1.6. BamUtil v1.4.14 was used to clip overlapping regions of the bam file in order to avoid counting multiple reads representing the same fragment. The genotypability of the *FMR1* gene was calculated using CallableLoci in GATK v3.8, with a minimum read depth of 10. CollectHsMetrics by Picard v2.17.10 was used to calculate fold enrichment to determine enrichment quality. Variants were called using HaplotypeCaller (GATK v4.1.8.0). Variant filtering was then carried out according to the GATK Best Practices for exomes. Variants were also filtered by quality (filter PASS) and by location within the *FMR1* gene. The accuracy of variant calling for each replicate was calculated using SNPSift, comparing their genotypes with the GIAB NA12878_HG001 annotated VCF file,[11] based on variants called by at least two different pipelines. Variants were annotated using VarSeq (GoldenHelix, Bozeman, MT, United States) to screen clinical databases of germline mutations: ClinVar and HGMD Professional v2020.1.

# RESULTS

## *FMR1* Enrichment Using Xdrop Technology

A specific primer pair was designed to amplify a DS by dPCR ~5 kbp from the microsatellite repeat in exon 1 of the *FMR1* gene (**Supplementary Figure S1**). Another primer pair was designed to anneal ~500 bp from the latter in order to monitor enrichment by qPCR (**Supplementary Figure S1**).

The Xdrop *FMR1* assay was tested on samples comprising DNA fragments >60 kbp extracted using five different methods (**Supplementary Figure S4A**). Following Xdrop-mediated encapsulation and dPCR, a clear cloud of positive droplets was visible by FACS for all but one of the extraction methods (**Supplementary Figure S4B**). We sorted an average of ~500 positive droplets for each sample, allowing the recovery of ~1.3 µg of enriched DNA after dMDA (**Figures 1A,B**), each of which was 12–15 kbp in length (**Supplementary Figure S4C**). The *FMR1* target showed a median enrichment of 170× across all samples based on qPCR analysis (**Figure 1C**). Although the Circulomics ultra-HMW protocol resulted in highly variable enrichments, no significant differences were observed among the extraction methods on average, with the exception of Qiagen columns (which did not achieve successful enrichment).

A subset of Xdrop-enriched DNA samples was sequenced using the Illumina and ONT platforms, generating on average 11,493,290 and 170,532 reads, with average lengths of 150 and 4,098 bp, respectively (**Supplementary Table S1**). Both sequencing methods achieved low genome-wide coverage (~0.2×) but significant enrichment was reproducibly observed for all samples on the *FMR1* gene: 462× for Illumina and 357× for ONT (**Figure 1D**; **Supplementary Table S1**). Maximum enrichment for both sequencing technologies was observed on the DS, and progressively decreased moving away from the target site, with a coverage >10× maintained for up to ±40 kbp flanking the DS (**Figures 1E,F**).

## Analysis of *FMR1* Repeat Characteristics by Xdrop Enrichment and ONT Sequencing

Next, we analyzed ONT sequencing data representing samples with known repeat features and showing expansions of 100–1,000 bp (**Table 1**). The consistent enrichment achieved on the target (range 33–330×) facilitated the extraction of sufficient reads spanning the entire tandem array (22 to 257) and allowed us to determine allele counts and features for every sample (**Figure 2** and **Table 1**). Sample NA12878 showed the anticipated normal pattern of 28 CGG repeats in both alleles, interrupted by the AGG trinucleotide at two sites. Sample NA06891 was derived from a male patient in the pre-mutation stage, with 118–121 CGG repeats according to previous sequencing data (Amos Wilson et al., 2008; Lim et al., 2017). Consistently, our analysis counted an average of 119 CGG repeats and highlighted the presence of a single AGG trinucleotide interrupting the array. Sample NA20241 was obtained from a female patient heterozygous for normal and pre-mutated alleles. The expanded allele was reported to contain 93–110 repeats based on traditional methods (Amos Wilson et al., 2008), whereas more recent PacBio sequencing analysis revealed two groups of molecules with 90 and 120 repeats, respectively (Tsai et al., 2017). In agreement with the latter study, our analysis demonstrated the presence of mosaicism in this sample, evident as a bimodal distribution of sequencing read lengths, with modal values of 92 and 113 repeats. The CGG repeat count of the normal allele was also confirmed as 29, interrupted by two AGG trinucleotides. Sample NA07537 was previously reported to be heterozygous with 29 CGG repeats in the normal allele and > 200 in the expanded allele, corresponding to a full mutation (Adler et al., 2011). The expanded allele was also characterized by PacBio sequencing, revealing a broad size distribution of 272–400 CGG repeats, which was confirmed by our data. Specifically, ONT sequencing reads ranged from a minimum of 196 to a maximum of 402 repeats, with a modal value of 342. Overall, the analysis of Xdrop-enriched samples by ONT sequencing allowed the accurate assessment of *FMR1* repeat length for each allele, and their correct classification as normal, pre-mutation, or full mutation. Moreover, the per-base analysis revealed repeat interruptions and mosaicism in agreement with previous reports.

[9]http://www.bioinformatics.babraham.ac.uk/projects/fastqc/
[10]https://arxiv.org/abs/1303.3997
[11]https://ftp-trace.ncbi.nlm.nih.gov/ReferenceSamples/giab/release/NA12878_HG001/latest/GRCh38/supplementaryFiles/HG001_GRCh38_GIAB_highconf_CG-IllFB-IllGATKHC-Ion-10X-SOLID_CHROM1-X_v.3.3.2_annotated.vcf.gz

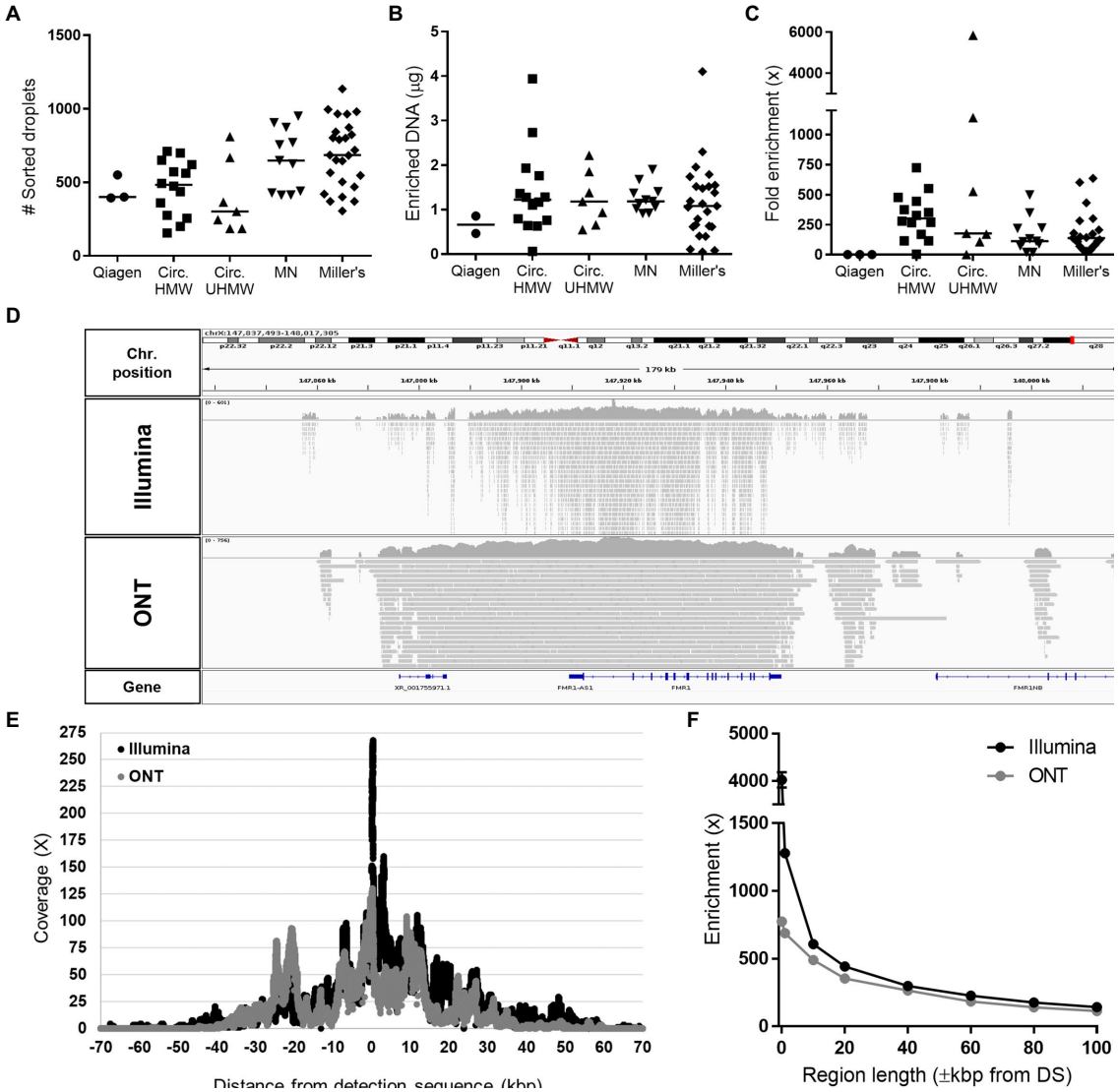



**FIGURE 1 |** Statistics of *FMR1* enrichment using the Xdrop technology. **(A)** Number of positive sorted droplets, **(B)** quantity of amplified DNA recovered after dMDA, and **(C)** fold enrichment of *FMR1* determined by qPCR after applying the Xdrop workflow to DNA samples extracted with different methods: Genomic Tip kit (Qiagen), Circulomics Nanobind CBB Big DNA Kit using either the HMW protocol (Circ. HMW) or the ultra-HMW protocol (Circ. UHMW), the NucleoSpin Blood Mini kit (Macherey-Nagel, MN), or Miller's protocol (Coriell samples). **(D)** Integrative Genomics Viewer (IGV) visualization of Illumina and ONT mapped reads obtained from a representative Xdrop-enriched sample. **(E)** Average coverage and **(F)** fold enrichment of the *FMR1* gene after sequencing the Xdrop-enriched samples listed in **Supplementary Table S1** on the ONT and Illumina platforms.

## Analysis of *FMR1* Intragenic Variants by Xdrop Enrichment and Illumina Sequencing

Xdrop allowed the enrichment of a genomic region containing the entire *FMR1* gene, so we next analyzed intragenic SNVs and indels in the same four samples discussed above (**Supplementary Table S1**). At this aim, we exploited Illumina sequencing (**Supplementary Table S1**), because with ONT most of gene body (74%) had coverage <60X (**Figure 1F**), i.e., lower than the minimum threshold required to accurately call SNV using this technology (Maestri et al., 2020). Analysis of the five distinct dMDA replicates of the NA12878 sample, for which genotypes are available, demonstrated most of the GIAB variants were properly called by each replicate, with 93% sensitivity on average (**Supplementary Table S2**). A minor fraction of false-positive (FP, 9%) and false-negative (FN, 7%) variants was also identified, but not reproducibly detected among replicates. FN variants were caused by non-callable/non-covered positions or allele dropout, whereas FP variants were usually supported by low read depth (<15 reads) and characterized by low Variant Allele Frequency (VAF < 25%, in 67% of the FP cases).

Based on these results, to avoid FNs, variants were called on the other three *FMR1* cases considering both available

**TABLE 1 |** Characterization of *FMR1* repeats from Xdrop-enriched samples by ONT sequencing.

| Sample ID | Condition | Sex | Total ONT reads | Mean coverage on repeat | Fold enrichment on repeat | Reads spanning repeat | Allele | Average number of expected repeats | Average number of observed repeats |
|---|---|---|---|---|---|---|---|---|---|
| NA12878 | Normal | F | 96,549 | 33.0X | 713.9 | 22 | 1 | 28 CGG + 2 AGG | 28 CGG + 2 AGG |
| | | | | | | | 2 | 28 CGG + 2 AGG | 28 CGG + 2 AGG |
| NA06891 | Pre-mutation | M | 265,556 | 55.5X | 791.2 | 36 | 1 | 118 (Amos Wilson et al., 2008) | 119 CGG + 1 AGG |
| | | | | | | | 2 | 121 (Lim et al., 2017) | – |
| NA20241 | Pre-mutation | F | 94,262 | 330.8X | 930.3 | 257 | 1 | – | – |
| | | | | | | | | 29(Chen et al. 2010; Juusola et al. 2012; Lim et al., 2017) | 27 CGG + 2 AGG |
| | | | | | | | | 27CGG + 2AGG (Tsai et al., 2017) | |
| | | | | | | | | 93–110(Amos Wilson et al., 2008) | |
| | | | | | | | 2 | 125(Lim et al., 2017) | 114 CGG + 1 AGG (mosaicism) |
| | | | | | | | | 119, mosaicism (Tsai et al., 2017) | |
| | | | | | | | | 28–29 (Adler et al., 2011) | |
| NA07537 | Mutation | F | 737,359 | 146.8X | 912.9 | 87 | 1 | 27 CGG + 2 AGG (Tsai et al., 2017) | 27 CGG +2 AGG |
| | | | | | | | 2 | >200, mosaicism (Adler et al., 2011; Lim et al., 2017; Tsai et al., 2017) | 342 CGG + 1 AGG (mosaicism) |

*For each sample, ONT sequencing statistics covering the FMR1 repeat are shown, along with the anticipated repeat features based on earlier reports and data generated from Xdrop-enriched samples.*

dMDA replicates (**Table 2**). Each sample showed an average coverage breadth >5× and genotypability ranging from 91 to 100% on *FMR1*. The consideration of both replicates allowed the entire *FMR1* gene length to be genotyped (99.99%), including the 34 positions of pathogenic/likely pathogenic variants listed in clinical databases (**Table 2** and **Supplementary Table S3**). These positions could be genotyped in all samples by both replicates, except for two variants in sample NA06891 that could be called based on only a single replicate (**Table 2** and **Supplementary Table S3**). No variant was identified in these positions, in agreement with the absence of pathogenic SNVs/indels reported within the *FMR1* gene for these samples (**Table 2**). These results confirmed that Xdrop enrichment coupled to Illumina sequencing allows the analysis of clinically relevant variants in the *FMR1* gene, but the use of technical dMDA replicates is necessary for complete, high-confidence variant calling.

## DISCUSSION

The accurate characterization of short tandem repeats, which underlie numerous inherited diseases, is challenging to achieve using traditional methods and even with the most recent sequencing technologies. Yet the correct diagnosis of these diseases and informed prognosis requires the precise determination of the number of repeats as well as the complete and accurate characterization at single-nucleotide resolution of both the repetitive site and the surrounding regions.

In this study, we demonstrated that the size and composition of triplet repeats in the *FMR1* gene can be determined accurately by Xdrop enrichment coupled to ONT long-read sequencing. The approach allowed us to classify the full range of *FMR1* alleles (normal, pre-mutation, and full mutation), with accurate size estimates comparable to previous results. Furthermore, the enrichment of sequencing data at the target site was sufficient to compensate for the consistent frequency of ONT sequencing errors, thus allowing the high-confidence identification of AGG interruptions. This aspect is essential because the presence of one or zero interruptions within a pre-mutated allele confers a high risk of expansion into a full mutation (Nolin et al., 2003; Yrigollen et al., 2012). In FXPOI patients, presence of AGG interruptions has also an effect on the fragile X-associated ovarian dysfunction (Lekovich et al., 2018). The precise determination of interruption patterns in female (pre-mutation) carriers is therefore critical because it influences their reproductive planning. Depending on the expansion risk, women might opt for preimplantation genetic diagnosis or normal conception, optionally combined with invasive prenatal diagnosis to screen the fragile X status of their fetus (Coskun and Alsmadi, 2007; Chen et al., 2020). In this context, Xdrop technology offers advantages over the Cas9 approach, because 500–1,000 times less DNA is required, allowing the application of long-read sequencing to limited samples, such as those derived from prenatal testing (Mosca-Boidron et al., 2013). In addition, the inclusion of an additional dMDA step before the conventional Xdrop workflow may

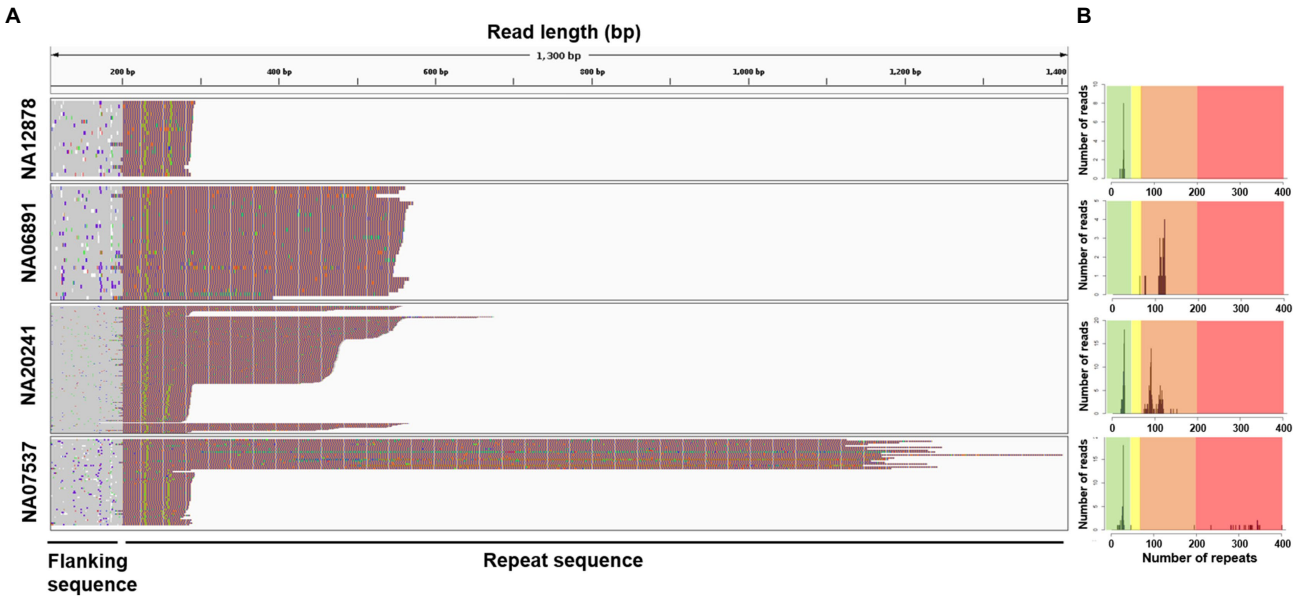



**FIGURE 2 |** Visualization of repeat structure and length after sequencing Xdrop-enriched samples on the ONT platform. **(A)** Individual ONT reads were trimmed to include only the *FMR1* repeat region plus 400 bp flanking sequence and aligned at the repeat 5'-end. Each line represents a single read, colored according to: A = green, T = red, G = orange, and C = blue. **(B)** Repeat count histograms showing the number of reads reporting a certain repeat length: shaded background in each plot represents risk ranges for disease development. Green = normal; yellow = intermediate; orange = pre-mutation; and red = full mutation.

**TABLE 2 |** Genotyping data for the *FMR1* gene from Xdrop-enriched samples based on Illumina sequencing.

| Sample ID | Replicate | Average coverage FMR1 | %5x | %10x | %PASS | %PASS DP10 | Total %PASS | Number of variants identified | Pathogenic variants | Reported pathogenic variants genotypable |
|---|---|---|---|---|---|---|---|---|---|---|
| NA12878 | R1 | 293 | 100 | 98.74 | 98.99 | 97.17 | 100 | 20 | 0 | 34/34 |
| | R2 | 244 | 99.58 | 95.54 | 96.44 | 91.41 | | 22 | | |
| | R3 | 707 | 95.58 | 93.01 | 93.61 | 91.1 | | 18 | | |
| | R4 | 52 | 95.05 | 85.62 | 96.89 | 84.04 | | 22 | | |
| | R5 | 114 | 100 | 99.37 | 100 | 97.58 | | 20 | | |
| NA06891 | R1 | 257 | 91.3 | 90.4 | 91.3 | 90.2 | 99.9 | 35 | 0 | 34/34 |
| | R2 | 164 | 100 | 100 | 99.9 | 99.9 | | 52 | | |
| NA20241 | R1 | 544 | 100 | 100 | 100 | 100 | 100 | 37 | 0 | 34/34 |
| | R2 | 873 | 100 | 100 | 100 | 100 | | 38 | | |
| NA07537 | R1 | 336 | 98.3 | 96.6 | 98 | 96.2 | 100 | 42 | 0 | 34/34 |
| | R2 | 400 | 100 | 99.3 | 100 | 99.2 | | 41 | | |

*For each replicate, columns show the average coverage of Illumina data on FMR1, the percentage of the FMR1 gene covered by at least 5 or 10 reads, the percentage of callable bases on the target for standard read depth (> 3, %PASS) or read depth > 10 (%PASS DP10), and the total number of variants identified, considering the combination of replicates, the total percentage of callable bases on the target for read depth > 3, the pathogenic variants identified, and the fraction of known pathogenic/likely pathogenic variants in PASS positions.*

enable target enrichment even from a single cell, as required for preimplantation testing (Hård et al., 2021).

In addition to repeat interruptions, we also detected a consistent level of mosaicism affecting the size of tandem repeats in pre-mutated and fully mutated alleles. Repeat instability is a hallmark of repeat expansion disorders (Nolin et al., 1999), and it may also explain why the accurate sizing of repeats in previous studies using traditional methods has been so challenging (Tsai et al., 2017). Assessing the variability in CGG repeats within and between tissues is another important aspect of FXS and FXTAS diagnosis because this can influence the clinical phenotype of affected individuals (Pretto et al., 2014). Although our results confirmed the presence of somatic mosaicism, in agreement with previous reports based on long-read sequencing (Tsai et al., 2017; Hafford-Tear et al., 2019; Sone et al., 2019), we also observed some variability within clusters, which may reflect the accumulation of indel errors along ONT reads.

From the technical perspective, the level of enrichment achieved with the Xdrop technology, typically >200x, was

comparable to other targeted enrichment strategies associable to long-read sequencing (Tsai et al., 2017; Gilpatrick et al., 2020). Also, the genome-wide noise of Xdrop samples (˜0.21X) was similar to that obtained with the Cas9 system coupled to ONT (˜0.3X) in our hands. Currently, the workflow for Xdrop-based enrichment takes ˜1.5 days and ˜150 € per sample, both anticipated to be reduced with the expected release of a Xdrop system integrated with a flow sorter. After enrichment, sequencing costs are comparable for the Xdrop- and the Cas9-system, that are both compatible with ONT and PacBio long-read technologies, and Xdrop also with Illumina. In our experiments, the high enrichment achieved using Xdrop technology facilitated downstream analysis with a limited amount of sequencing data (< 1 million ONT reads and ~ 10 million Illumina reads). The Xdrop workflow also provides an option to assess enrichment by qPCR before proceeding with sequencing. This should be considered solely as a qualitative test to ensure successful results (when >100x), because there was no full correlation between the enrichment level determined by sequencing and qPCR in our experiments. One potential drawback is the generation of chimeric reads, despite the robust limitation of this phenomenon by droplet-based Phi29 amplification (Zhou et al., 2020; Gonzalez-Pena et al., 2021; Hård et al., 2021). Chimeras can be overcome by considering supplementary alignments, even if such adjustments reduce the length of ONT read mapping, thus limiting the ability to accurately assess repeat lengths expanding over primary alignment sizes. This issue could be addressed by reducing the dMDA reaction time or using more efficient sorting systems. The breadth of enrichment achieved using Xdrop technology spanned a ~ 60–80 kbp region flanking a single "point of view" (DS) sequence of 100 bp, and probably corresponded to the average length of the initial DNA molecules encapsulated in the droplets. This is an advantageous feature because it allows the analysis of the entire *FMR1* gene, far beyond the triplet repeat stretch. This option is not readily available when using the Cas9 enrichment approach, which is typically limited to the distance specified by the guide RNA location (≤ 20 kbp). To maximize the breadth of coverage covering the whole gene, HMW starting DNA is preferred, but ultra-HMW molecules should be avoided because viscous DNA is difficult to dilute down to nanograms with any accuracy, resulting in variable enrichments (as we experienced with the Circulomics ultra-HMW protocol). In our hands, a wide set of DNA extraction methods provided similar enrichment results, even when not specifically designed for HMW DNA extraction (i.e., MN, based on standard silica columns). The possibility to use extraction kits routinely used in diagnostic procedures as well as frozen blood, as our starting samples could facilitate the broad application of the technology in the clinic. An exception was the Qiagen Gtip kit that did not properly work in combination with Xdrop. The latter may reflect the carryover of contaminants that interfere with DNA encapsulation/staining and could not be removed using bead-based cleanup methods.

Investigation of the full FMR1 gene is beneficial when the analysis of repeat expansion is inconclusive and the exclusion of other mutations within the gene body is desirable either to complete genetic testing or to prevent disease transmission. Indeed, besides repeat expansion, FXS can be also caused by point mutations or deletions, as those recently reported to occur in the 5'UTR of FMR1, that can challenge genetic diagnosis (Erbs et al., 2021). Although rare, these variants could be fine characterized at the breakpoint level by the Xdrop method, thus facilitating the segregation analysis and genetic counseling. We genotyped SNVs along the entire *FMR1* gene by coupling Xdrop enrichment to Illumina sequencing. Because the dMDA step yields sufficient DNA for both protocols, the same Xdrop-enriched DNA can therefore be sequenced on the ONT platform for the analysis of repeat expansions, and then with Illumina technology to assess the presence of intragenic variants. Based on this approach, we could accurately genotype the entire *FMR1* gene and analyze all pathogenic/likely pathogenic variants reported therein. The small fraction of FP and FN calls was not reproducible in multiple samples and could therefore be identified by analyzing two sequencing replicates from different dMDA reactions. Accordingly, the same initial sorted sample could be split in two aliquots for independent downstream amplifications. Because FN calls were mainly due to coverage drops, these could be minimized by adding a second DS in the same reaction, to allow more uniform enrichment over the 80 kbp length of the gene, especially when using not-HMW DNA. FP calls were probably caused by dMDA errors, because the constant amplification of a few hundred molecules obtained by sorting may result in this frequency of artifacts, as well documented for single-cell sequencing (Gawad et al., 2016). We excluded the possibility that FP calls were derived from contamination prior to the dMDA step, such as sorting processing or the operator, because negative controls (sequencing the sheath fluid from the flow sorter) did not reveal the presence of human DNA. Preliminary data also suggested that FP calls may be exacerbated when the efficiency of sorting is suboptimal because this reduces the number of target molecules collected and thus increases the chance of Phi29 errors. The anticipated launch of an Xdrop system integrated with a flow sorter should maximize the efficiency of this method, thus overcoming such technical limitations.

## CONCLUSION

Our study demonstrated the simultaneous characterization of challenging microsatellite expansions and SNV/indels within the *FMR1* gene, which has not been achieved before. This was possible thanks to the implementation of a novel targeted sequencing approach, in which Xdrop enrichment was combined with the analysis of large DNA fragments by short-read and long-read sequencing. Although technical improvements are required to implement this approach in the clinic, our proof-of-concept study should be easily adapted for the analysis of other genes characterized by repeat expansions, or other genomic loci where the analysis of structural variations combined with the detection of SNVs and indels is desirable for complete genetic counseling.

## DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: https://www.ncbi.nlm.nih.gov/, PRJNA745542.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethics Committee for Clinical Research of Verona and Rovigo Provinces. The patients/participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

MR, MD, VG, and SM contributed to the conception and design of the study. VG, LM, SM, MA, DL, BI, and BM performed the experiments and analyzed the data. SM developed bioinformatic pipelines for ONT data analysis. AS, AB, MD'A, and GN revised the intellectual and clinical content of the manuscript. MR, VG, and SM wrote the manuscript. MR and MD acquired funding. All authors contributed to manuscript revision, read, and approved the submitted version.

## FUNDING

The work was supported by the University of Verona Joint Project JPVR18RY8Y. The founding sponsor had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; and in the decision to publish the results.

## ACKNOWLEDGMENTS

Flow sorting was performed at the Technological Platforms (CPT) of the University of Verona. We acknowledge Samplix's team for the good technical support.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.743230/full#supplementary-material

## REFERENCES

Adler, K., Moore, J. K., Filippov, G., Wu, S., Carmichael, J., and Schermer, M. (2011). A novel assay for evaluating fragile X locus repeats. *J. Mol. Diagn.* 13, 614–620. doi: 10.1016/j.jmoldx.2011.06.002

Amos Wilson, J., Pratt, V. M., Phansalkar, A., Muralidharan, K., Highsmith, W. E. Jr., Beck, J. C., et al. (2008). Consensus characterization of 16 FMR1 reference materials: a consortium study. *J. Mol. Diagn.* 10, 2–12. doi: 10.2353/jmoldx.2008.070105

Ardui, S., Race, V., de Ravel, T., Van Esch, H., Devriendt, K., Matthijs, G., et al. (2018). Detecting AGG interruptions in females With a FMR1 Premutation by long-read single-molecule sequencing: A 1 year clinical experience. *Front. Genet.* 9:150. doi: 10.3389/fgene.2018.00150

Bastepe, M., and Xin, W. (2015). Huntington disease: molecular diagnostics approach. *Curr. Protoc. Hum. Genet.* 87, 9.26.1–9.26.23. doi: 10.1002/0471142905.hg0926s87

Benson, G. (1999). Tandem repeats finder: a program to analyze DNA sequence. *Nucleic Acids Res.* 27, 573–580. doi: 10.1093/nar/27.2.573

Blondal, T., Gamba, C., Jagd, L. M., Ling, S., Demirov, D., Guo, S., et al. (2021). Verification of CRISPR editing and finding transgenic inserts by Xdrop indirect sequence capture followed by short- and long-read sequencing. *Methods* 191, 68–77. doi: 10.1016/j.ymeth.2021.02.003

Braida, C., Stefanatos, R. K., Adam, B., Mahajan, N., Smeets, H. J., Niel, F., et al. (2010). Variant CCG and GGC repeats within the CTG expansion dramatically modify mutational dynamics and likely contribute toward unusual symptoms in some myotonic dystrophy type 1 patients. *Hum. Mol. Genet.* 19, 1399–1412. doi: 10.1093/hmg/ddq015

Chakraborty, S., Vatta, M., Bachinski, L. L., Krahe, R., Dlouhy, S., and Bai, S. (2016). Molecular diagnosis of myotonic dystrophy. *Curr. Protoc. Hum. Genet.* 91, 9.29.1–9.29.19. doi: 10.1002/cphg.22

Charles, P., Camuzat, A., Benammar, N., Sellal, F., Destée, A., Bonnet, A. M., et al. (2007). Are interrupted SCA2 CAG repeat expansions responsible for parkinsonism? *Neurology* 69, 1970–1975. doi: 10.1212/01.wnl.0000269323.21969.db

Chen, L., Hadd, A., Sah, S., Filipovic-Sadic, S., Krosting, J., Sekinger, E., et al. (2010). An information-rich CGG repeat primed PCR that detects the full range of Fragile X expanded alleles and minimizes the need for southern Blot analysis. *J. Mol. Diagn.* 12:5. doi: 10.2353/jmoldx.2010.090227

Chen, S., Yin, X., Zhang, S., Xia, J., Liu, P., Xie, P., et al. (2020). Comprehensive preimplantation genetic testing by massively parallel sequencing. *Hum. Reprod.* 36, 236–247. doi: 10.1093/humrep/deaa269

Chen, S., Zhou, Y., Chen, Y., and Jia, G. (2018). Fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34, i884–i890. doi: 10.1093/bioinformatics/bty560

R Core Team (2013). *R: A Language and Environment for Statistical Computing*. (Vienna: Austria).

Coskun, S., and Alsmadi, O. (2007). Whole genome amplification from a single cell: a new era for preimplantation genetic diagnosis. *Prenat. Diagn.* 27, 297–302. doi: 10.1002/pd.1667

De Coster, W., D'Hert, S., Schultz, D. T., Cruts, M., and Van Broeckhoven, C. (2018). NanoPack: visualizing and processing long-read sequencing data. *Bioinformatics* 34, 2666–2669. doi: 10.1093/bioinformatics/bty149

Eichler, E. E., Holden, J. J., Popovich, B. W., Reiss, A. L., Snow, K., Thibodeau, S. N., et al. (1994). Length of uninterrupted CGG repeats determines instability in the FMR1 gene. *Nat. Genet.* 8, 88–94. doi: 10.1038/ng0994-88

Erbs, E., Fenger-Grøn, J., Jacobsen, C. M., Lildballe, D. L., and Rasmussen, M. (2021). Spontaneous rescue of a FMR1 repeat expansion and review of deletions in the FMR1 non-coding region. *Eur. J. Med. Genet.* 64:104244. doi: 10.1016/j.ejmg.2021.104244

Filipovic-Sadic, S., Sah, S., Chen, L., Krosting, J., Sekinger, E., Zhang, W., et al. (2010). A novel FMR1 PCR method for the routine detection of low abundance expanded alleles and full mutations in fragile X syndrome. *Clin. Chem.* 56, 399–408. doi: 10.1373/clinchem.2009.136101

Gawad, C., Koh, W., and Quake, S. R. (2016). Single-cell genome sequencing: current state of the science. *Nat. Rev. Genet.* 17, 175–188. doi: 10.1038/nrg.2015.16

Giesselmann, P., Brändl, B., Raimondeau, E., Bowen, R., Rohrandt, C., Tandon, R., et al. (2019). Analysis of short tandem repeat expansions and their methylation state with nanopore sequencing. *Nat. Biotechnol.* 37, 1478–1481. doi: 10.1038/s41587-019-0293-x

Gilpatrick, T., Lee, I., Graham, J. E., Raimondeau, E., Bowen, R., Heron, A., et al. (2020). Targeted nanopore sequencing with Cas9-guided adapter ligation. *Nat. Biotechnol.* 38, 433–438. doi: 10.1038/s41587-020-0407-5

Gonzalez-Pena, V., Natarajan, S., Xia, Y., Klein, D., Carter, R., Pang, Y., et al. (2021). Accurate genomic variant detection in single cells with primary

template-directed amplification. *Proc. Natl. Acad. Sci.* 118:e2024176118. doi: 10.1073/pnas.2024176118

Hafford-Tear, N. J., Tsai, Y.-C., Sadan, A. N., Sanchez-Pintado, B., Zarouchlioti, C., Maher, G. J., et al. (2019). CRISPR/Cas9-targeted enrichment and long-read sequencing of the Fuchs endothelial corneal dystrophy–associated TCF4 triplet repeat. *Genet. Med.* 21, 2092–2102. doi: 10.1038/s41436-019-0453-x

Hård, J., Mold, J. E., Eisfeldt, J., Tellgren-Roth, C., Häggqvist, S, Bunikis, I, et al. (2021). Long-read whole genome analysis of human single cells, *bioRxiv*. [Preprint]. doi:10.1101/2021.04.13.439527

Harris, R. S., Cechova, M., and Makova, K. D. (2019). Noise-cancelling repeat finder: uncovering tandem repeats in error-prone long-read sequencing data. *Bioinformatics* 35, 4809–4811. doi: 10.1093/bioinformatics/btz484

Hayward, B. E., Zhou, Y., Kumari, D., and Usdin, K. (2016). A set of assays for the comprehensive analysis of FMR1 alleles in the fragile X-related disorders. *J. Mol. Diagn.* 18, 762–774. doi: 10.1016/j.jmoldx.2016.06.001

Juusola, J. S., Anderson, P., Sabato, F., Wilkinson, D. S., Pandya, A., and Ferreira-Gonzalez, A. (2012). Performance evaluation of two methods using commercially available reagents for PCR-based detection of FMR1 mutation. *J. Mol. Diagn.* 14, 476–486. doi: 10.1016/j.jmoldx.2012.03.005

Katoh, K., Misawa, K., Kuma, K., and Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* 30, 3059–3066. doi: 10.1093/nar/gkf436

Lekovich, J., Man, L., Xu, K., Canon, C., Lilienthal, D., Stewart, J. D., et al. (2018). CGG repeat length and AGG interruptions as indicators of fragile X-associated diminished ovarian reserve. *Genet. Med.* 20, 957–964. doi: 10.1038/gim.2017.220

Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 34, 3094–3400. doi: 10.1093/bioinformatics/bty191

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

Lim, G. X., Yeo, M., Koh, Y. Y., Winarni, T. I., Rajan-Babu, I. S., Chong, S. S., et al. (2017). Validation of a commercially available test that enables the quantification of the numbers of CGG trinucleotide repeat expansion in FMR1 gene. *PLoS One* 12:e0173279. doi: 10.1371/journal.pone.0173279

Loomis, E. W., Eid, J. S., Peluso, P., Yin, J., Hickey, L., Rank, D., et al. (2013). Sequencing the unsequenceable: expanded CGG-repeat alleles of the fragile X gene. *Genome Res.* 23, 121–128. doi: 10.1101/gr.141705.112

Madsen, E. B., Höijer, I., Kvist, T., Ameur, A., and Mikkelsen, M. J. (2020). Xdrop: targeted sequencing of long DNA molecules From low input samples using droplet sorting. *Hum. Mutat.* 41, 1671–1679. doi: 10.1002/humu.24063

Maestri, S., Cosentino, E., Paterno, M., Freitag, H., Garces, J. M., Marcolungo, L., et al. (2019). A rapid and accurate MinION-based workflow for tracking species biodiversity in the field. *Genes* 10:468. doi: 10.3390/genes10060468

Maestri, S., Maturo, M. G., Cosentino, E., Marcolungo, L., Iadarola, B., Fortunati, E., et al. (2020). A long-read sequencing approach for direct haplotype phasing in clinical settings. *Int. J. Mol. Sci.* 21:9177. doi: 10.3390/ijms21239177

Mantere, T., Kersten, S., and Hoischen, A. (2019). Long-read sequencing emerging in medical genetics. *Front. Genet.* 10:426. doi: 10.3389/fgene.2019.00426

Matsuyama, Z., Izumi, Y., Kameyama, M., Kawakami, H., and Nakamura, S. (1999). The effect of CAT trinucleotide interruptions on the age at onset of spinocerebellar ataxia type 1 (SCA1). *J. Med. Genet.* 36, 546–548.

McFarland, K. N., Liu, J., Landrian, I., Godiska, R., Shanker, S., Yu, F., et al. (2015). SMRT sequencing of long tandem nucleotide repeats in SCA10 reveals unique insight of repeat expansion structure. *PLoS One* 10:e0135906. doi: 10.1371/journal.pone.0135906

McFarland, K. N., Liu, J., Landrian, I., Zeng, D., Raskin, S., Moscovich, M., et al. (2014). Repeat interruptions in spinocerebellar ataxia type 10 expansions are strongly associated with epileptic seizures. *Neurogenetics* 15, 59–64. doi: 10.1007/s10048-013-0385-6

Miller, S. A., Dykes, D. D., and Polesky, H. F. (1988). A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res.* 16:1215. doi: 10.1093/nar/16.3.1215

Mosca-Boidron, A.-L., Faivre, L., Aho, S., Marle, N., Truntzer, C., Rousseau, T., et al. (2013). An improved method to extract DNA from 1 ml of uncultured amniotic fluid from patients at less than 16 weeks' gestation. *PLoS One* 8:e59956. doi: 10.1371/journal.pone.0059956

Nolin, S. L., Brown, W. T., Glicksman, A., Houck, G. E. Jr., Gargano, A. D., Sullivan, A., et al. (2003). Expansion of the fragile X CGG repeat in females with premutation or intermediate alleles. *Am. J. Hum. Genet.* 72, 454–464. doi: 10.1086/367713

Nolin, S. L., Glicksman, A., Ersalesi, N., Dobkin, C., Brown, W. T., Cao, R., et al. (2014). Fragile X full mutation expansions are inhibited by one or more AGG interruptions in premutation carriers. *Genet. Med.* 17, 358–364. doi: 10.1038/gim.2014.106

Nolin, S. L., Houck, G. E., Gargano, A. D., Blumstein, H., Dobkin, C. S., and Ted Brown, W. (1999). FMR1 CGG-repeat instability in single sperm and lymphocytes of fragile-X Premutation males. *Am. J. Hum. Genet.* 65, 680–688. doi: 10.1086/302543

Paulson, H. (2018). "Chapter 9- repeat expansion diseases," in *Handbook of Clinical Neurology*. eds. D. H. Geschwind, H. L. Paulson and C. Klein (Amsterdam: Elsevier)

Pretto, D., Yrigollen, C. M., Tang, H.-T., Williamson, J., Espinal, G., Iwahashi, C. K., et al. (2014). Clinical and molecular implications of mosaicism in FMR1 full mutations. *Front. Genet.* 5:318. doi: 10.3389/fgene.2014.00318

Quartier, A., Poquet, H., Gilbert-Dussardier, B., Rossi, M., Casteleyn, A.-S., Portes, V. d., et al. (2017). Intragenic FMR1 disease-causing variants: a significant mutational mechanism leading to fragile-X syndrome. *Eur. J. Hum. Genet.* 25, 423–431. doi: 10.1038/ejhg.2016.204

Quinlan, A. R., and Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841–842. doi: 10.1093/bioinformatics/btq033

Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., et al. (2011). Integrative genomics viewer. *Nat. Biotechnol.* 29, 24–26. doi: 10.1038/nbt.1754

Rodriguez, C. M., and Todd, P. K. (2019). New pathologic mechanisms in nucleotide repeat expansion disorders. *Neurobiol. Dis.* 130:104515. doi: 10.1016/j.nbd.2019.104515

Rognes, T., Flouri, T., Nichols, B., Quince, C., and Mahé, F. (2016). VSEARCH: a versatile open source tool for metagenomics. *PeerJ* 4:e2584. doi: 10.7717/peerj.2584

Sakamoto, N., Larson, J. E., Iyer, R. R., Montermini, L., Pandolfo, M., and Wells, R. D. (2001). GGA*TCC-interrupted triplets in long GAA*TTC repeats inhibit the formation of triplex and sticky DNA structures, alleviate transcription inhibition, and reduce genetic instabilities. *J. Biol. Chem.* 276, 27178–27187. doi: 10.1074/jbc.M101852200

Saluto, A., Brussino, A., Tassone, F., Arduino, C., Cagnoli, C., Pappi, P., et al. (2005). An enhanced polymerase chain reaction assay to detect pre- and full mutation alleles of the fragile X mental retardation 1 gene. *J. Mol. Diagn.* 7, 605–612. doi: 10.1016/S1525-1578(10)60594-6

Sitzmann, A. F., Hagelstrom, R. T., Tassone, F., Hagerman, R. J., and Butler, M. G. (2018). Rare FMR1 gene mutations causing fragile X syndrome: A review. *Am. J. Med. Genet. A* 176, 11–18. doi: 10.1002/ajmg.a.38504

Sone, J., Mitsuhashi, S., Fujita, A., Mizuguchi, T., Hamanaka, K., Mori, K., et al. (2019). Long-read sequencing identifies GGC repeat expansions in NOTCH2NLC associated with neuronal intranuclear inclusion disease. *Nat. Genet.* 51, 1215–1221. doi: 10.1038/s41588-019-0459-y

Spector, E., Behlmann, A., Kronquist, K., Rose, N. C., Lyon, E., and Reddi, H. V. (2021). Laboratory testing for fragile X, 2021 revision: a technical standard of the American College of Medical Genetics and Genomics (ACMG). *Genet. Med.* 23, 799–812. doi: 10.1038/s41436-021-01115-y

Stangl, C., de Blank, S., Renkens, I., Westera, L., Verbeek, T., Valle-Inclan, J. E., et al. (2020). Partner independent fusion gene detection by multiplexed CRISPR-Cas9 enrichment and long read nanopore sequencing. *Nat. Commun.* 11:2861. doi: 10.1038/s41467-020-16641-7

Tabolacci, E., Pietrobono, R., Maneri, G., Remondini, L., Nobile, V., Della Monica, M., et al. (2020). Reversion to Normal of FMR1 expanded alleles: A rare event in two independent fragile X syndrome families. *Genes* 11:248. doi: 10.3390/genes11030248

Tsai, Yu-Chih, Greenberg, David, Powell, James, Höijer, Ida, Ameur, Adam, Strahl, Maya, et al. (2017). Amplification-free, CRISPR-Cas9 Targeted Enrichment and SMRT Sequencing of Repeat-Expansion Disease Causative Genomic Regions, *bioRxiv*. [Preprint]. doi:10.1101/203919

van Blitterswijk, M., DeJesus-Hernandez, M., Niemantsverdriet, E., Murray, M. E., Heckman, M. G., Diehl, N. N., et al. (2013). Association between repeat sizes and clinical and pathological characteristics in carriers of C9ORF72 repeat expansions (Xpansize-72): a cross-sectional cohort study. *Lancet Neurol.* 12, 978–988. doi: 10.1016/S1474-4422(13)70210-2

Vaser, R., Sović, I., Nagarajan, N., and Šikić, M. (2017). Fast and accurate de novo genome assembly from long uncorrected reads. *Genome Res.* 27, 737–746. doi: 10.1101/gr.214270.116

Warner, J. P., Barron, L. H., Goudie, D., Kelly, K., Dow, D., Fitzpatrick, D. R., et al. (1996). A general method for the detection of large CAG repeat expansions by fluorescent PCR. *J. Med. Genet.* 33, 1022–1026. doi: 10.1136/jmg.33.12.1022

Yrigollen, C. M., Durbin-Johnson, B., Gane, L., Nelson, D. L., Hagerman, R., Hagerman, P. J., et al. (2012). AGG interruptions within the maternal FMR1 gene reduce the risk of offspring with fragile X syndrome. *Genet. Med.* 14, 729–736. doi: 10.1038/gim.2012.34

Zhou, X., Xu, Y., Zhu, L., Zhen, S., Han, X., Zhang, Z., et al. (2020). Comparison of multiple displacement amplification (MDA) and multiple annealing and looping-based amplification cycles (MALBAC) in limited DNA sequencing based on tube and droplet. *Micromachines* 11:645. doi: 10.3390/mi11070645

**Conflict of Interest:** AS, MD and MR are partners of Genartis srl.

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.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

*Copyright © 2021 Grosso, Marcolungo, Maestri, Alfano, Lavezzari, Iadarola, Salviati, Mariotti, Botta, D'Apice, Novelli, Delledonne and Rossato. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*

frontiers in Genetics

CASE REPORT
published: 05 October 2021
doi: 10.3389/fgene.2021.705808

# Molecular Diagnosis and Prenatal Phenotype Analysis of Eight Fetuses With Ciliopathies

*Yuefang Liu*[1†], *Hui Wang*[2,3†], *Xin Jin*[1], *Qixiang Shao*[2,3,4*] *and Qiong Pan*[1*]

[1]*Department of Clinical Genetics, Huai'an Maternity and Child Clinical College of Xuzhou Medical University, Huai'an, China,* [2]*Reproductive Sciences Institute, Jiangsu University, Zhenjiang, China,* [3]*Department of Immunology, Jiangsu Key Laboratory of Medical Science and Laboratory Medicine, School of Medicine, Zhenjiang, China,* [4]*Jiangsu College of Nursing, School of Medical Science and Laboratory Medicine, Huai'an, China*

Human ciliopathies are hereditary conditions caused by variants in ciliary-associated genes. Ciliopathies are often characterized by multiple system defects. However, it is not easy to make a definite diagnosis in the prenatal period only based on the imageology. In this report, eight new prenatal cases from five unrelated families diagnosed with ciliopathies were systematically examined. The clinical manifestations of these fetuses showed such prenatal diagnostic features as occipital encephalocele, and polydactyly and polycystic kidneys. *Situs inversus* caused by *CPLANE1* variant was first reported. In Family 1 and Family 3, homozygous variants of *CPLANE1* and *NPHP4* caused by consanguineous marriage and uniparental disomy were detected by whole-exome sequencing, respectively. In Family 2, Family 4 and Family 5, compound heterozygotes of *TMEM67* and *DYNC2H1* including two novel missense variants and one novel nonsense variant were identified. The distribution of pathogenic missense variants along *TMEM67* gene mainly clustered in the extracellular cysteine rich region, extracellular area with unknown structure, and the transmembrane regions. Genotype-phenotype relationship between *CPLANE1* and *TMEM67* genes was concluded. This report describes new clinical manifestations and novel variants in *CPLANE1*, *TMEM67*, *NPHP4*, and *DYNC2H1*.

**Keywords: ciliopathies, prenatal clinical phenotype, occipital encephalocele, polydactyly, polycystic kidneys, whole-exome sequencing**

OPEN ACCESS

***Edited by:***
*Kathleen M. Gorman, Temple Street Children's University Hospital, Ireland*

***Reviewed by:***
*Maria Paola Lombardi, University of Amsterdam, Netherlands*
*Qifei Li, Boston Children's Hospital and Harvard Medical School, United States*

***Correspondence:**
*Qixiang Shao*
*shao_qx@ujs.edu.cn*
*Qiong Pan*
*jonespan@163.com*

*†These authors have contributed equally to this work*

***Specialty section:***
*This article was submitted to Genetics of Common and Rare Diseases, a section of the journal Frontiers in Genetics*

***Received:*** *06 May 2021*
***Accepted:*** *30 August 2021*
***Published:*** *05 October 2021*

***Citation:***
*Liu Y, Wang H, Jin X, Shao Q and Pan Q (2021) Molecular Diagnosis and Prenatal Phenotype Analysis of Eight Fetuses With Ciliopathies. Front. Genet. 12:705808. doi: 10.3389/fgene.2021.705808*

## INTRODUCTION

Primary cilia are highly conserved organelles located on the surface of almost all polar cells, which play important roles in tissue morphogenesis, and chemical and mechanical signal transduction (Hirokawa et al., 2006; Copeland, 2020). Genetic variants affecting the structure or function of primary cilia can lead to a broad range of developmental diseases known as ciliopathies. The whole-exome sequencing (WES) has been widely used in the clinical molecular diagnosis of ciliopathies in adults and children. Presently, approximately 187 established ciliopathy-related genes have been identified in humans, variants in which can be associated with 35 ciliopathy syndromes (Reiter and Leroux, 2017; McConnachie et al., 2021). It has been clarified that variant in *CPLANE1* causes Joubert syndrome (JBS; MIM#614615), which is characterized by a unique cerebellar and brainstem malformation, also known as molar tooth sign (MTS). Moreover, *TMEM67*-related ciliopathies are mainly JBS and Meckel syndrome (MKS; MIM#607361). While MKS is a lethal disorder with typical renal cystic dysplasia, polydactyly, and occipital encephalocele. Variants in *NPHP4* cause nephronophthisis (NPH; MIM#606966) characterized

**TABLE 1 |** Clinical phenotype and related genetic variants in eight fetuses with ciliopathies from five unrelated families.

| Family | Case | Age | OE | PD | PK | Other | Gene | Variant | Inheritance | ACMG | Disease |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | $18^{+4}$ w | – | – | – | Situs inversus, CHD, DWM CVA | *CPLANE1* | c.7939delC (p.H2647IfsTer51) | Hom | P | JBS |
| | 2 | 20 w | + | – | – | IUGR, banana-shaped cerebellum, lemon head | *CPLANE1* | c.7939delC (p.H2647IfsTer51) | Hom | P | JBS |
| 2 | 3 | $23^{+4}$ w | – | – | + | Pericardial effusion, LV separation, cerebellar dysplasia, lemon head | *TMEM67* | c.1175C > G (p.P392R)<br>c.2439G > T (p.A813A, splice) | F<br>M | LP<br>P | MKS |
| | 4 | 24 w | + | + | + | Balkes pouch | *TMEM67* | c.1175C > G (P392R)<br>c.2439G > T (p.A813A, splice) | F<br>M | LP<br>P | MKS |
| 3 | 5 | $26^{+4}$ w | – | – | – | Abnormal nasal bone, increased renal cortical echogenicity IUGR | *NPHP4* | c.3730C > T (p.Q1244X) | Hom | P | NPH |
| 4 | 6 | 17 w | – | + | – | Dilation of the renal pelvis, narrow chest, short limbs, CPC | *DYNC2H1* | c.152T > G (p.L51R)<br>c.988C > T (p.R330C) | M<br>F | LP<br>P | SRTD |
| | 7 | 28 w | – | + | – | Narrow chest, short limbs | *DYNC2H1* | c.152T > G (p.L51R)<br>c.988C > T (p.R330C) | M<br>F | LP<br>P | SRTD |
| 5 | 8 | 17 w | – | + | – | Short limbs, low-set ears, CHD | *DYNC2H1* | c.4267C > T (p.R1423C)<br>c.7858C > T (p.R2620X) | F<br>M | P<br>P | SRTD |

*OE, occipital encephalocele; PD, polydactyly; PK, polycystic kidneys; DWM, Dandy-Walker malformation; CVA, cerebellar vermis agenesis; CHD, congenital heart disease; IUGR, intrauterine growth retardation; CPC, choroid plexus cyst; LV, left ventricle; F, father, M, mother; Hom, homozygous; ACMG, American College of Medical Genetics and Genomics; P, pathogenic; LP, likely pathogenic; JBS, Joubert syndrome; MKS, Meckel syndrome; SRTD, short-rib thoracic dysplasia; NPH, nephronophthisis.*

by end-stage renal disease in the first 2 decades of life. In addition, the typical feature of short-rib thoracic dysplasia (SRTD; MIM#613091) caused by *DYNC2H1* variants is skeletal dysplasia. However, the different etiology of ciliopathies and the wide range of genetic variations lead to phenotypic variability. It is challenging to choose appropriate molecular testing in the prenatal period. Therefore, it is important to have a comprehensive understanding of the prenatal phenotypes of different ciliopathy syndromes before the WES testing.

Here we report the typical and atypical features of eight fetuses with pathogenic variants in *CPLANE1, TMEM67, NPHP4*, and *DYNC2H1* found *via* WES. New clinical manifestations and the discovery of novel genetic variants are helpful for the prenatal diagnosis of ciliopathies.

# PATIENTS AND METHODS

## Patients

Eight fetuses from five unrelated families diagnosed with ciliopathies were collected. Informed consents for research investigations were obtained from the relatives of the fetuses. The research protocol was approved by the local ethics committee of Jiangsu Huai'an Maternity and Child Healthcare Hospital (2019036).

## Whole-Exome Sequencing

The whole exomes were captured by using Agilent's SureSelect Whole Exome Gene Detection Kit. High-throughput sequencing was performed by using the Novoseq sequencer from Illumina. The obtained sequence was aligned on the human genome GRCh37/hg19 reference sequence by BWA (Burrows-Wheeler Aligner) software. A BAM (binary sequence alignment map format) file was produced *via* Picard software. GATK4 (Genome Analysis Toolkit) Realigner Target Creator software and Haplotype Caller software were used to adjust the sequence, extract variants, and generate VCF (Variant Call Format) files. The Annovar software was used to filter and annotate the variant.

## Analysis of Variants

All nonsense, frameshift, and canonical splice site variants were considered to be deleterious. The pathogenic potential of missense variants was predicted by PolyPhen2, SIFT, PROVEAN, and Mutation Taster. The frequency of putative variants was obtained from the Human Gene Mutation Database (HGMD), Genome Aggregation Database (gnomAD), and the 1000 Genomes (1000G) database. Conservation of mutated amino acid residues in different species was compared by UCSC. SpliceAI was used to evaluate a destroyed splice site. The deleteriousness of variants was assessed according to American College of Medical Genetics (ACMG) standards and guidelines.

## Sanger Sequencing

Sanger sequencing was performed to confirm suspected variant segregation within probands' family, and the authenticity of variants identified by WES. Primer 5 was used for primer designs. Target DNA of the fetus and its parents was amplified by PCR. Sanger sequencing results were compared with standard sequence in GenBank by SeqMan software.

# CASE PRESENTATION

## Family 1

Family 1 was a consanguineous marriage with a healthy girl and five adverse obstetric outcomes. In addition to two early miscarriages, the third time was due to a widened posterior fossa and vermis hypoplasia at the 27th week of pregnancy without genetic detection. Case 1 was the fifth pregnancy, whose prenatal ultrasound in $18^{+4}$th week of pregnancy

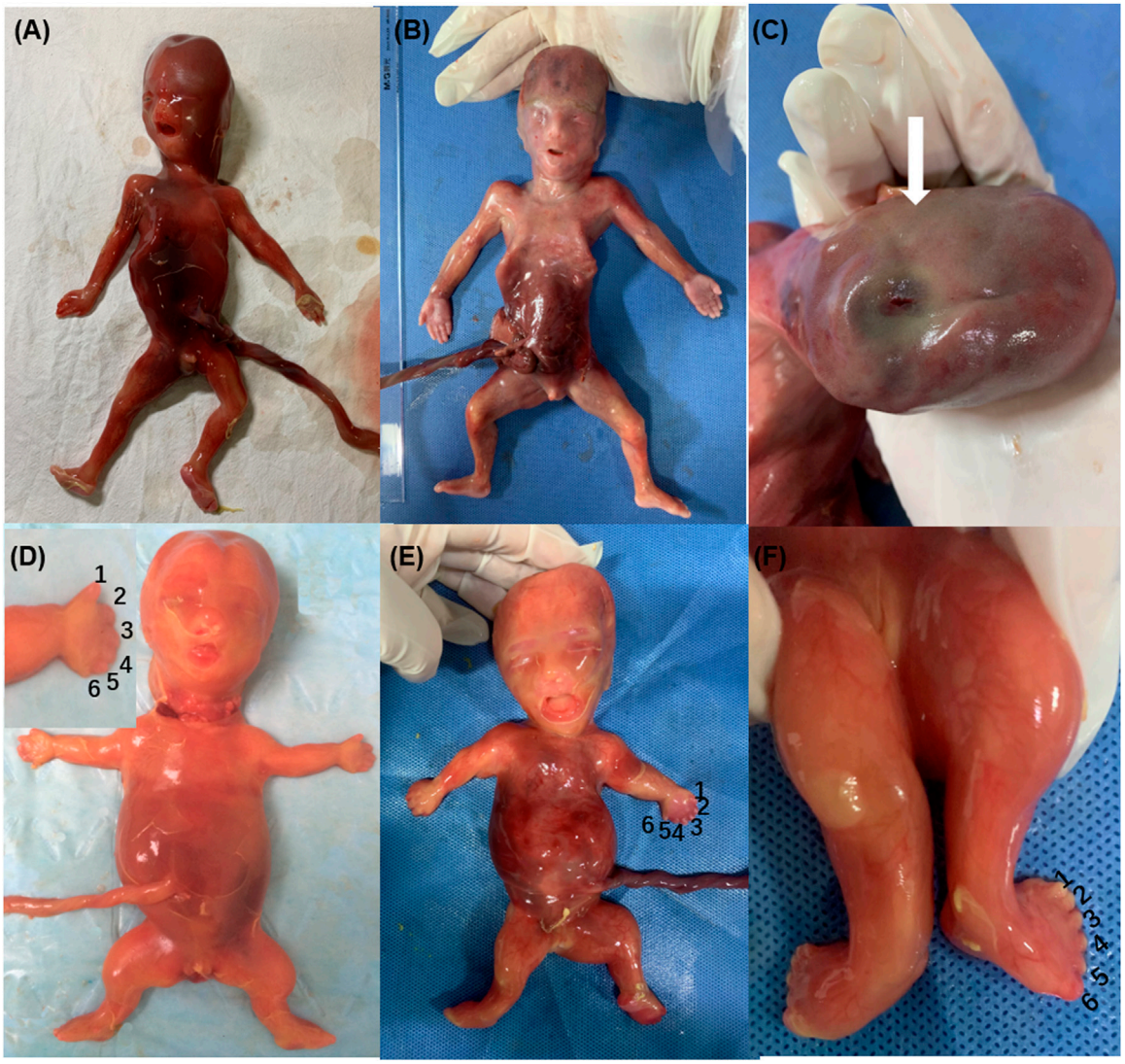



**FIGURE 1 |** Phenotypes of ciliary diseases {Family 1 [**(A)**: case 1; **(B,C)**: case 2], Family 4 [**(D)**: case 6], Family 5 [**(E,F)**: case 8], occipital encephalocele **(C)**, polydactyly of hands and feet **(D–F)**, short limb and ribs **(D,E)**}.

showed absence of cerebellar vermis, dandy-walker malformation, hydrocephalus, dextrocardia, *situs inversus*, ventricular septal defect (0.3 cm), and double outlet of fetal right ventricle with pulmonary artery stenosis (**Table 1**). Pregnancy was terminated at the 23rd week and no obvious abnormality in appearance was observed (**Figure 1A**). The sixth pregnancy (case 2) showed recurrence of malformations by prenatal ultrasound at the 20th week of fetal development including occipital encephalocele (1.3 × 0.8 cm) (**Figure 2A**) and intra-uterine growth restriction (**Table 1**) (**Figures 1B,C**).

WES revealed case 1 with a novel homozygous frame shift variant of c.7939delC (p.H2647IfsTer51) in *CPLANE1* (NM_023073.3). Sanger sequencing confirmed homozygosity in the proband and the heterozygosity in each parent (**Table 1**). Case 2 was also confirmed with same variant in case 1. The c.7939delC was not present in HGMD, gnomAD, or 1000G databases.

## Family 2

Case 3 and case 4 were affected siblings conceived of unrelated healthy parents with two early unexplained miscarriages. Prenatal ultrasound in $23^{+4}$rd week of case 3 showed separation of left lateral ventricles (0.9 cm), abnormality of the cerebellum, low amniotic fluid volume (amniotic fluid index 1.5 cm, maximum depth 1.0 cm), pericardial effusion (0.3 cm), and enlarged echogenic kidneys (left kidney 5.7 × 2.8 cm, right kidney 5.4 × 2.8 cm) with renal cysts (**Table 1**). Case 4 was found with congenital balkes pouch, cystic kidneys, polydactyly, and occipital encephalocele (**Table 1**). The pregnancy was terminated at the 24th week of gestation. Because prenatal detection and labor induction were performed in another hospital, we only obtained DNA of the two cases without ultrasound images and postnatal fetus examination.

WES was selected by Family 2 and the result showed that case 4 had compound heterozygous variants in *TMEM67* gene

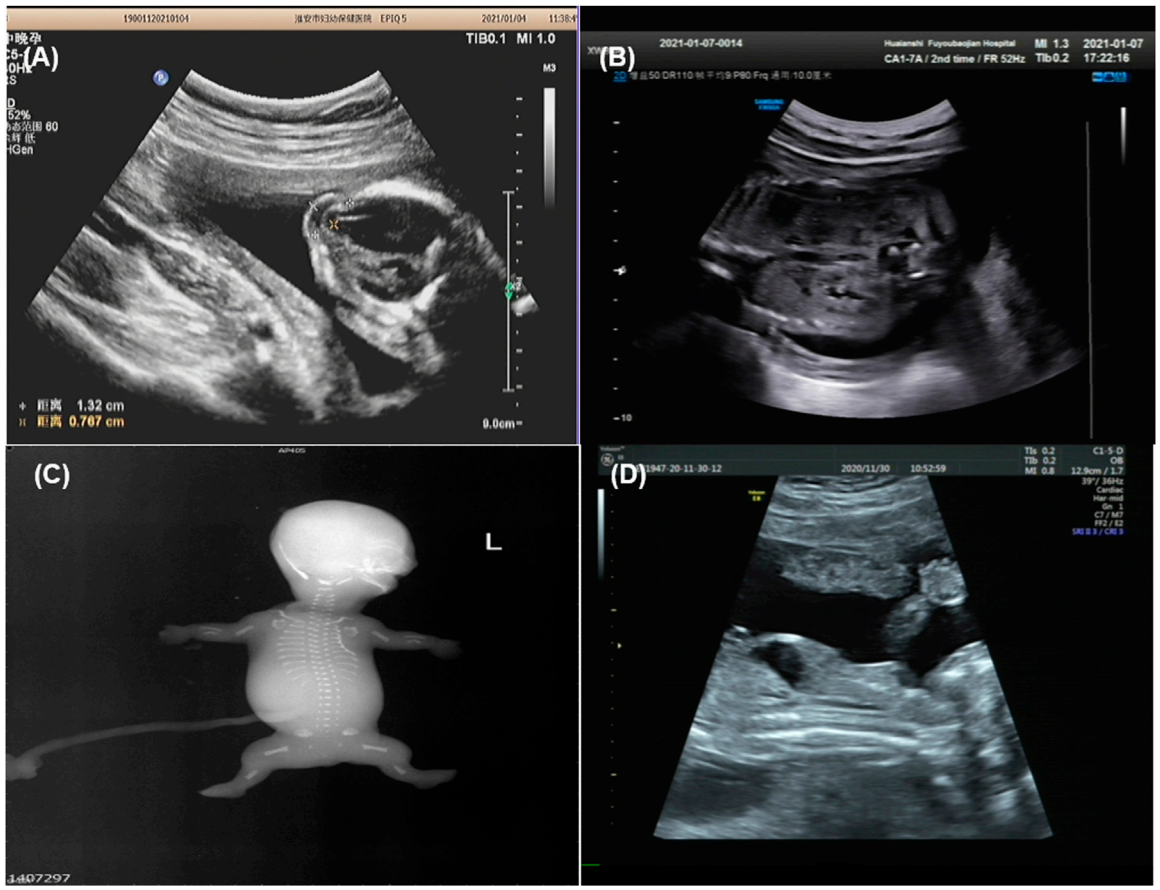



**FIGURE 2 |** Imaging examination of ciliary diseases {Family 1 [**(A)**: case 2], Family 3 [**(B)**: case 5], Family 4 [**(C)**: case 6], Family 5 [**(D)**: case 8], occipital encephalocele **(A)**, increased echogenicity in both renal cortexes **(B)**, fetal X-ray showing bilateral shortened curved femora and disproportionately shortened tibiae **(C)**, bone dysplasia and narrow chest **(D)**}.

(NM_153704.6): a missense variant at c.1175C > G (p.P392R) and a synonymous variant at c.2439G > T (p.A813A), which were inherited from the parents, respectively (**Table 1**). The c.1175C > G, not reported previously, was predicted as damaging by PolyPhen2, SIFT, PROVEAN, and Mutation Taster. Total population frequency of this variant is 3.23185e-05. The c. 2439G > T variant is located at the last base of exon 23 and was classified as a disease-causing variant in HGMD. It might abolish the donor splice site following *in silico* analysis via SpliceAI. The predicted score of a donor site loss via SpliceAI was 0.64. After this, the DNA of case 3 was subjected to molecular analysis by Sanger sequencing. The result revealed the c.1175C > G and c.2439G > T variants in the *TMEM67* gene in case 3.

## Family 3

Case 5 was the first pregnancy of unrelated healthy parents. Prenatal diagnosis was performed in the 20th week of pregnancy due to a 12 Mb deletion in chromosome 1 as indicated by noninvasive prenatal testing (NIPT). The Affymetrix CytoScan 750K SNP array was used for chromosomal microarray analysis (CMA) and the result showed a region of homozygosity in 1p36.33-p36.13 (arr[hg19] 1p36.33p36.13 (888,658_18,337,268) x2 hmz) involving no established imprinted genes. In $26^{+4}$th week, missing nasal bone, delayed growth and development, and increased echogenicity in both renal cortexes were indicated by ultrasound (**Figure 2B** and **Table 1**). Prenatal trio-WES was chosen for further detection. A homozygous variant of c.3730C > T (p.Q1244X) (chr1:5,925,248) in *NPHP4* (NM_001184.3) (**Table 1**) was detected. This variant, occurring in the region of homozygosity, was inherited from the father. These results indicated that a segmental uniparental disomy at 1p36.33-p36.13 inherited from the father caused the autosomal recessive NPH.

## Family 4

Case 6 and case 7 were affected sibling fetuses from a non-consanguineous family with a healthy 6-year-old daughter. Ultrasound imaging of case 6 revealed abdominal circumference (10.3 cm), a femur length of 1.3 cm (<5th percentile), a humerus length of 1.0 cm (<5th percentile), and polydactyly. The pregnancy was subsequently terminated at the 18th week of gestation (**Figure 1D**). Fetal X-ray (**Figure 2C**) showed bilateral shortened curved femora and disproportionately shortened tibiae. Metaphyseal flaring of femora, tibiae, and fibulae were indicated. Ultrasound imaging of case 7 at the

**TABLE 2 |** The comparison of clinical phenotypes between *CPLANE1* and *TMEM67*.

| Phenotype | *CPLANE1* | *TMEM67* |
|---|---|---|
| MTS | 13/14 (JBS) | 1/20 (MKS) |
| | 12/12 (OFD6) | 10/13 (JBS) |
| OE | 6/26 (OFD6) | 10/20 (MKS) |
| | 2/15 (JBS) | 1/13 (JBS) |
| CHD | 1/12 (OFD6) | / |
| PD | 1/15 (JBS) | 3/20 (MKS) |
| | 12/12 (OFD6) | 0/22 (JBS) |
| RD | 1/14 (JBS) | 1/20 (MKS) |
| | 0/12 (OFD6) | 0/22 (JBS) |
| KD | 0/14 (JBS) | 11/22 (JBS) |
| | 0/12 (OFD6) | 14/20 (MKS) |
| LD | 6/15 (JBS) | 21/22 (JBS) |
| | | 11/12 (MKS) |
| DD | 14/14 (JBS) | 2/20 (MKS) |
| | 4/4 (OFD6) | 11/13 (JBS) |

*MTS, molar tooth sign; OE, occipital encephalocele; CHD, congenital heart disease; PD, polydactyly; RD, retinal disease; KD, kidney disease; LD, liver disease; DD, developmental delay; JBS, Joubert syndrome; MKS, Meckel syndrome; OFD6, orofaciodigital syndrome VI.*

28th week of gestation from this family exhibited similar sonographic features including small short femur (1.2 cm), narrow chest, light band separation in double kidney collection system (0.8 and 0.9 cm), peritoneal effusion (deepest 0.5 cm), and polydactyly. We only obtained DNA of case 7 but failed to perform postnatal fetus observation.

WES of case 6 revealed the compound heterozygous variants of *DYNC2H1* (NM_001377.3): c.152T > G (p.L51R) in exon 1 and c.988C > T (p.R330C) in exon 6. Sanger sequencing confirmed that the c.152T > G was present in the mother, and the c.988C > T in the father (**Table 1**). Case 7 was also confirmed with variants as case 6 through Sanger sequencing (**Table 1**). The c.988C > T has been reported as a pathogenic variant (El Hokayem et al., 2012; Schmidts et al., 2013). The c.152T > G was not present in HGMD, gnomAD, or 1000G databases, and was predicted as deleterious following *in silico* analysis. In addition, the c.152T > G is localized to the region of nucleotide-binding site of the *DYNC2H1* gene, which is highly conserved in mouse, rat, and dog orthologs of *DYNC2H1* gene.

## Family 5

Prenatal ultrasound of case 8 (the first fetus from non-relative parents) in the 12th week of pregnancy showed abnormal NT (4.4 mm). Recheck in the 17th week showed multiple malformations with large abdominal circumference (12.1 cm), ventricular septal defect (0.08 cm), left ventricular dysplasia, valve thickening, and transposition of the great arteries (**Table 1**). Besides, increased bowel and kidney echo, and bone dysplasia (femur = 1.3 cm, humerus = 0.9 cm, narrow chest) (**Figure 2D**) were detected. Examination of the aborted fetus showed low-set ears, polydactyly, a swollen abdomen, and short limbs (**Figures 1E,F**). Compound heterozygous variants of c.4267C > T (p.R1423C) and c.7858C > T (p.R2620X) in *DYNC2H1* (NM_001377.3) were identified. We confirmed a heterozygosity in each parent by Sanger sequencing (**Table 1**). The c.4267C > T was a pathogenic variant in short-rib polydactyly syndromes (SRPS) (Zhang et al., 2018). The c.7858C > T was a novel nonsense variant causing a truncated protein.

## Genotypes–Phenotypes Analysis

Since *CPLANE1* and *TMEM67* variants show significant clinical overlap, genotype-phenotype relationship analysis between them could help discriminate related disorders.

*CPLANE1* was identified as causative of OFD6 (orofaciodigital syndrome VI) and JBS. More than 60 pathogenic variants in the *CPLANE1* gene have been identified (Srour et al., 2012; Lopez et al., 2014; Vilboux et al., 2017; Bonnard et al., 2018) and 66% of cases carried biallelic truncating variants (Bonnard et al., 2018). Penetrance of common organ systems involving brain, skeleton, and kidney diseases in *CPLANE1* patients were compared (**Table 2**) (Lopez et al., 2014; Vilboux et al., 2017; Bonnard et al., 2018). MTS was nearly 100% detected in both JBS and OFD6. Polydactyly was with higher penetrance (100%) in OFD6 than JBS. The penetrance of occipital encephalocele in both JBS and OFD6 was estimated to be about 20%. However, cystic kidneys, short limbs, and congenital heart disease (CHD) were rarely observed.

*TMEM67* was identified as causative of JBS and MKS. A genotype-phenotype correlation analysis based on a literature review shows that combination of two truncating variants in *TMEM67* gene is more common in lethal MKS than milder JBS (Iannicelli et al., 2010; Szymanska et al., 2012; Bachmann-Gagescu et al., 2015; Vilboux et al., 2017), which suggests that severity of

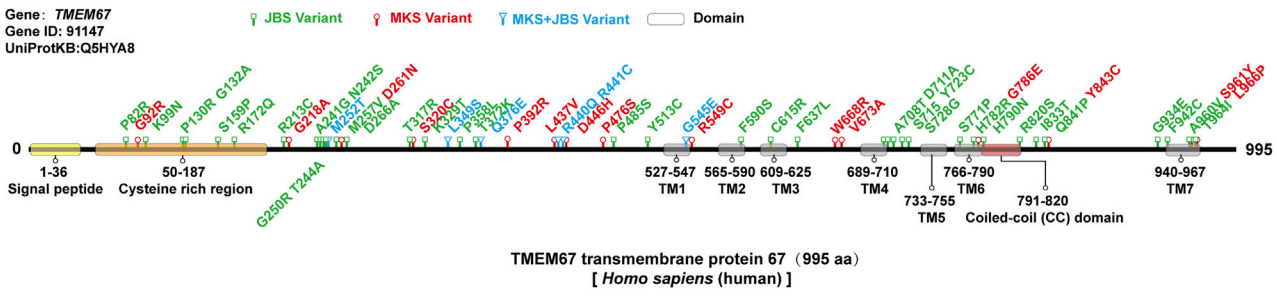



**FIGURE 3 |** Missense variants of the *TMEM67* gene in MKS and JBS. There were about 58 (likely) pathogenic missense variants in the *TMEM67* gene identified thus far including 37 in JBS (green), 15 in MKS (red), and 6 in both (blue). The topology of *TMEM67*: a signal peptide (1–36 aa), a cysteine-rich region (50–187 aa), an extracellular area with unknown structure (188–526 aa), seven transmembrane regions (527–967 aa), and a short cytoplasmic tail (968–995 aa).

variant changes disease outcomes. Thus, it is reasonable to presume that missense variant of *TMEM67* identified in MKS might bring more severe damage to protein function. There were about 58 (likely) pathogenic missense variants in *TMEM67* identified thus far including 37 in JBS, 15 in MKS, and 6 in both (Baala et al., 2007; Consugar et al., 2007; Khaddour et al., 2007; Brancati et al., 2009; Doherty et al., 2010; Iannicelli et al., 2010; Suzuki et al., 2016; Vilboux et al., 2017; Radhakrishnan et al., 2019) (**Supplemental Table S1**). Analysis of the topology of human *TMEM67* suggests the presence of a signal peptide (1–36 aa), a cysteine-rich region (50–187 aa), an extracellular area with unknown structure (188–526 aa), seven transmembrane regions (527–967 aa), and a short cytoplasmic tail (968–995 aa) (**Figure 3**) (Smith et al., 2006). Pathogenic missense variants of *TMEM67* mainly fall in the extracellular cysteine rich region, extracellular area with unknown structure, and the seven transmembrane regions (**Figure 3**). No pathogenic missense variant in the signal peptide and cytoplasmic tail has been reported (**Figure 3**). The comparison of their distribution along the gene in MKS versus JBS showed no obvious difference.

Two studies with larger *TMEM67* gene mutated cases were reviewed to analyze the characteristic of clinical phenotypes (**Table 2**) (Brancati et al., 2009; Szymanska et al., 2012; Vilboux et al., 2017). Incidence of liver and kidney diseases caused by *TMEM67* variants was high in both JBS and MKS. The prevalence of occipital encephalocele in MKS was higher than in JBS.

## DISCUSSION AND CONCLUSION

A clinical diagnosis of JBS can be made by brain imaging showing MTS. However, prenatal diagnosis of JBS has been proved to be difficult because of the relatively non-specific prenatal ultrasound findings. The clinical phenotypes of *CPLANE1* mutated cases in this report were highly heterogeneous, with intrafamilial variability. Meanwhile, case 1 and case 2 had no obvious MTS. Consistent with the absence of MTS in our fetuses, Doherty et al. described a JBS fetus with no obvious MTS phenotype instead of a deepening of the interpedullary fossa on MRI (Doherty et al., 2005). Zhu et al. reported a prenatal JBS case with lemon sign and encephalocele (Zhu et al., 2021). *Situs inversus,* a specific feature in some MKS cases resulted from *MKS1* variants (Khaddour et al., 2007), was reported for the first time in our JBS case.

MKS is a perinatally lethal autosomal recessive condition characterized by central nervous system anomalies, hepatic defects, polycystic kidneys, and polydactyly (Radhakrishnan et al., 2019). Hepatic phenotype in the prenatal period could be detected at about the 15th to 26th week (Consugar et al., 2007). MKS fetuses in Family 2 were found with abnormal nervous system, polycystic kidneys, polydactyly but not hepatic defects at the 23rd and 24th weeks of pregnancy.

NPH is a major cause of pediatric end-stage renal disease and the phenotypes were mostly observed in children (Konig et al., 2017). It may be limited to the kidneys or can be associated with extrarenal organ. A study including 250 NPH patients showed that only 6/250 patients (2.4%) were detected with homozygous or compound heterozygous variants of the *NPHP4* gene (Hoefele et al., 2005). The prenatal diagnosis is one of the effective ways to avoid the birth of NPH children. However, prenatal NPH cases are rarely described. An NPH fetus with *NPHP1* variants manifested bilateral polycystic renal dysplasia and oligohydramnios at $16^{+}$th gestational week (Wu et al., 2020). As for the fetus in this study, antenatal ultrasonography showed missing nasal bone and increased echogenicity in both renal cortex in $26^{+4}$th week.

There are 111 pathogenic variants of the *DYNC2H1* gene that have been identified in 73 families (Zhang et al., 2018). Missense variants in the *DYNC2H1* gene had the highest frequency (71/111). Characterized phenotypes were polydactyly with a prevalence 30/73 (Zhang et al., 2018) and short limbs with a prevalence of 23/29 (Schmidts et al., 2013), while abnormal retina, kidney, and liver were rare (Schmidts et al., 2013). Prenatal cases in our report were distinguished by profound abnormalities of the skeleton, including markedly short ribs, extremely short limbs, and polydactyly.

In conclusion, this report expands prenatal clinical manifestations of ciliopathies and adds novel variants in *CPLANE1*, *TMEM67*, *NPHP4*, and *DYNC2H1* to the literature. Furthermore, detailed prenatal phenotypes of different ciliopathies provide evidence for prenatal WES testing.

## DATA AVAILABILITY STATEMENT

The data presented in the study are deposited in the China National GeneBank DataBase (CNGBdb) repository, accession number CNP0002142.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Medical Ethics Committee of Huai'an Maternity and Child Health Hospital. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

## AUTHOR CONTRIBUTIONS

YL, HW, and XJ contributed to the conception. YL collected relevant studies and wrote the manuscript. QS and QP. contributed to the revision, and final approval of the manuscript.

## FUNDING

The present study was supported by grants from by the Maternal and Child Health project of Jiangsu Province (No. F201714, F201707); Open project of Jiangsu Province Key Laboratory (XZSYSKF2020024).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.705808/full#supplementary-material

## REFERENCES


Baala, L., Romano, S., Khaddour, R., Saunier, S., Smith, U. M., Audollent, S., et al. (2007). The Meckel-Gruber Syndrome Gene, MKS3, Is Mutated in Joubert Syndrome. *Am. J. Hum. Genet.* 80 (1), 186–194. doi:10.1086/510499

Bachmann-Gagescu, R., Dempsey, J. C., Phelps, I. G., O'Roak, B. J., Knutzen, D. M., Rue, T. C., et al. (2015). Joubert Syndrome: a Model for Untangling Recessive Disorders with Extreme Genetic Heterogeneity. *J. Med. Genet.* 52 (8), 514–522. doi:10.1136/jmedgenet-2015-103087

Bonnard, C., Shboul, M., Tonekaboni, S. H., Ng, A. Y. J., Tohari, S., Ghosh, K., et al. (2018). Novel Mutations in the Ciliopathy-Associated Gene CPLANE1 (C5orf42) Cause OFD Syndrome Type VI rather Than Joubert Syndrome. *Eur. J. Med. Genet.* 61 (10), 585–595. doi:10.1016/j.ejmg.2018.03.012

Brancati, F., Iannicelli, M., Travaglini, L., Mazzotta, A., Bertini, E., Boltshauser, E., et al. (2009). MKS3/TMEM67mutations Are a Major Cause of COACH Syndrome, a Joubert Syndrome Related Disorder with Liver Involvement. *Hum. Mutat.* 30 (2), E432–E442. doi:10.1002/humu.20924

Consugar, M. B., Kubly, V. J., Lager, D. J., Hommerding, C. J., Wong, W. C., Bakker, E., et al. (2007). Molecular Diagnostics of Meckel-Gruber Syndrome Highlights Phenotypic Differences between MKS1 and MKS3. *Hum. Genet.* 121 (5), 591–599. doi:10.1007/s00439-007-0341-3

Copeland, J. (2020). Actin-based Regulation of Ciliogenesis - the Long and the Short of it. *Semin. Cel Developmental Biol.* 102, 132–138. doi:10.1016/j.semcdb.2019.12.005

Doherty, D., Glass, I. A., Siebert, J. R., Strouse, P. J., Parisi, M. A., Shaw, D. W. W., et al. (2005). Prenatal Diagnosis in Pregnancies at Risk for Joubert Syndrome by Ultrasound and MRI. *Prenat. Diagn.* 25 (6), 442–447. doi:10.1002/pd.1145

Doherty, D., Parisi, M. A., Finn, L. S., Gunay-Aygun, M., Al-Mateen, M., Bates, D., et al. (2010). Mutations in 3 Genes (MKS3, CC2D2A and RPGRIP1L) Cause COACH Syndrome (Joubert Syndrome with Congenital Hepatic Fibrosis). *J. Med. Genet.* 47 (1), 8–21. doi:10.1136/jmg.2009.067249

El Hokayem, J., Huber, C., Couvé, A., Aziza, J., Baujat, G., Bouvier, R., et al. (2012). NEK1andDYNC2H1are Both Involved in Short Rib Polydactyly Majewski Type but Not in Beemer Langer Cases. *J. Med. Genet.* 49 (4), 227–233. doi:10.1136/jmedgenet-2011-100717

Hirokawa, N., Tanaka, Y., Okada, Y., and Takeda, S. (2006). Nodal Flow and the Generation of Left-Right Asymmetry. *Cell* 125 (1), 33–45. doi:10.1016/j.cell.2006.03.002

Hoefele, J., Sudbrak, R., Reinhardt, R., Lehrack, S., Hennig, S., Imm, A., et al. (2005). Mutational Analysis of theNPHP4 Gene in 250 Patients with Nephronophthisis. *Hum. Mutat.* 25 (4), 411. doi:10.1002/humu.9326

Iannicelli, M., Brancati, F., Mougou-Zerelli, S., Mazzotta, A., Thomas, S., Elkhartoufi, N., et al. (2010). NovelTMEM67mutations and Genotype-Phenotype Correlates in Meckelin-Related Ciliopathies. *Hum. Mutat.* 31 (5), a–n. doi:10.1002/humu.21239

Khaddour, R., Smith, U., Baala, L., Martinovic, J., Clavering, D., Shaffiq, R., et al. (2007). Spectrum ofMKS1andMKS3mutations in Meckel Syndrome: a Genotype-Phenotype Correlation. *Hum. Mutat.* 28 (5), 523–524. doi:10.1002/humu.9489

König, J., Kranz, B., König, S., Schlingmann, K. P., Titieni, A., Tönshoff, B., et al. (2017). Phenotypic Spectrum of Children with Nephronophthisis and Related Ciliopathies. *Cjasn* 12 (12), 1974–1983. doi:10.2215/CJN.01280217

Lopez, E., Thauvin-Robinet, C., Reversade, B., Khartoufi, N. E., Devisme, L., Holder, M., et al. (2014). C5orf42 Is the Major Gene Responsible for OFD Syndrome Type VI. *Hum. Genet.* 133 (3), 367–377. doi:10.1007/s00439-013-1385-1

McConnachie, D. J., Stow, J. L., and Mallett, A. J. (2021). Ciliopathies and the Kidney: A Review. *Am. J. Kidney Dis.* 77 (3), 410–419. doi:10.1053/j.ajkd.2020.08.012

Radhakrishnan, P., Nayak, S. S., Shukla, A., Lindstrand, A., and Girisha, K. M. (2019). Meckel Syndrome: Clinical and Mutation Profile in Six Fetuses. *Clin. Genet.* 96 (6), 560–565. doi:10.1111/cge.13623

Reiter, J. F., and Leroux, M. R. (2017). Genes and Molecular Pathways Underpinning Ciliopathies. *Nat. Rev. Mol. Cel Biol* 18 (9), 533–547. doi:10.1038/nrm.2017.60

Schmidts, M., Arts, H. H., Bongers, E. M. H. F., Yap, Z., Oud, M. M., Antony, D., et al. (2013). Exome Sequencing identifiesDYNC2H1mutations as a Common Cause of Asphyxiating Thoracic Dystrophy (Jeune Syndrome) without Major Polydactyly, Renal or Retinal Involvement. *J. Med. Genet.* 50 (5), 309–323. doi:10.1136/jmedgenet-2012-101284

Smith, U. M., Consugar, M., Tee, L. J., McKee, B. M., Maina, E. N., Whelan, S., et al. (2006). The Transmembrane Protein Meckelin (MKS3) Is Mutated in Meckel-Gruber Syndrome and the Wpk Rat. *Nat. Genet.* 38 (2), 191–196. doi:10.1038/ng1713

Srour, M., Schwartzentruber, J., Hamdan, F. F., Ospina, L. H., Patry, L., Labuda, D., et al. (2012). Mutations in C5ORF42 Cause Joubert Syndrome in the French Canadian Population. *Am. J. Hum. Genet.* 90 (4), 693–700. doi:10.1016/j.ajhg.2012.02.011

Suzuki, T., Miyake, N., Tsurusaki, Y., Okamoto, N., Alkindy, A., Inaba, A., et al. (2016). Molecular Genetic Analysis of 30 Families with Joubert Syndrome. *Clin. Genet.* 90 (6), 526–535. doi:10.1111/cge.12836

Szymanska, K., Berry, I., Logan, C. V., Cousins, S. R., Lindsay, H., Jafri, H., et al. (2012). Founder Mutations and Genotype-Phenotype Correlations in Meckel-Gruber Syndrome and Associated Ciliopathies. *Cilia* 1 (1), 18. doi:10.1186/2046-2530-1-18

Vilboux, T., Doherty, D. A., Glass, I. A., Parisi, M. A., Phelps, I. G., Cullinane, A. R., et al. (2017). Molecular Genetic Findings and Clinical Correlations in 100 Patients with Joubert Syndrome and Related Disorders Prospectively Evaluated at a Single center. *Genet. Med.* 19 (8), 875–882. doi:10.1038/gim.2016.204

Wu, Q., Yang, S., Wang, C., Shi, H., Ren, S., Jiao, Z., et al. (2020). Ultrasonographic Manifestation and Genetic Analysis of a Fetus with Nephronophthisis Type 2. *Zhonghua Yi Xue Yi Chuan Xue Za Zhi* 37 (5), 559–562. doi:10.3760/cma.j.issn.1003-9406.2020.05.016

Zhang, W., Taylor, S. P., Ennis, H. A., Forlenza, K. N., Duran, I., Li, B., et al. (2018). Expanding the Genetic Architecture and Phenotypic Spectrum in the Skeletal Ciliopathies. *Hum. Mutat.* 39 (1), 152–166. doi:10.1002/humu.23362

Zhu, H., Chen, W., Ren, H., Zhang, Y., Niu, Y., Wu, D., et al. (2021). Non-classic Splicing Mutation in the CPLANE1 (C5orf42) Gene Cause Joubert Syndrome in a Fetus with Severe Craniocerebral Dysplasia. *Eur. J. Med. Genet.* 64 (6), 104212. doi:10.1016/j.ejmg.2021.104212


**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.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

*Copyright © 2021 Liu, Wang, Jin, Shao and Pan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*

frontiers in Genetics

ORIGINAL RESEARCH
published: 25 October 2021
doi: 10.3389/fgene.2021.685123

# A Protocol for Preconceptional Screening of Consanguineous Couples Using Whole Exome Sequencing

*Carolina Maria de Araújo dos Santos[1], Ana Helena Heller[1], Heloisa Barbosa Pena[1] and Sérgio Danilo Junho Pena[1,2,3]**

*[1] GENE – Núcleo de Genética Médica, Belo Horizonte, Brazil, [2] Laboratório de Genômica Clínica, Faculdade de Medicina da Universidade Federal de Minas Gerais, Belo Horizonte, Brazil, [3] Departamento de Bioquímica e Imunologia, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil*

Genetic studies performed in consanguineous couples suggest that the reproductive risk that distinguish them from other couples in the general population is related to autosomal recessive (AR) diseases. This risk is scattered among the thousands of known and potential AR diseases. Thus, for effective preconceptional screening of consanguineous couples it is necessary a test that encompasses the largest number of genes possible. For that reason, we decided to create a protocol based on whole exome sequencing (WES). We sequenced completely the exomes of 39 consanguineous couples at high coverage (∼100×). Applying bioinformatics filters, we could detect genetic variants that were simultaneously present in both members of the couple in all genes listed in the Clinical Genomics Database as causally related to AR diseases. Shared variants were then assessed for pathogenicity. For non-truncating variants (missense and in-frame indels) we considered as pathogenic or likely pathogenic only the variants included as such in the ClinVar database. Shared truncating variants (frameshift, non-sense, and canonical splice variants) were considered likely pathogenic when loss-of-function was a known mechanism of disease. The 39 consanguineous cases included two couples with a coefficient of genetic relationship (CGR) of 0.25, 26 couples with a CGR of 0.125, three couples with a CGR of 0.0625 and eight couples with a CGR of 0.03125. In 21 of the 39 couples (53.8%) we ascertained sharing of heterozygosity for at least one variant considered pathogenic or likely pathogenic for an AR disease. In eight couples we found sharing of heterozygosity for at least two pathogenic variants. Once the specific pathogenic variant was identified, it became possible for the couple to undergo prenatal diagnosis or, if desired, preimplantation genetic diagnosis (PGD) involving *in vitro* fertilization and embryo screening. In conclusion, our results demonstrate that preconceptional screening by WES is a useful new procedure that should be incorporated in the genetic counseling of all consanguineous couples.

**Keywords: consanguinity, preconceptional genetic diagnosis, carrier screening, whole exome sequencing, genetic counseling**

OPEN ACCESS

***Edited by:***
*Kathleen M. Gorman, Temple Street Children's University Hospital, Ireland*

***Reviewed by:***
*Paul Lasko, McGill University, Canada*
*Maria Lalioti, Yale University, United States*

***Correspondence:**
*Sérgio Danilo Junho Pena*
*spena@dcc.ufmg.br*

***Specialty section:***
*This article was submitted to Genetics of Common and Rare Diseases, a section of the journal Frontiers in Genetics*

***Received:*** *24 March 2021*
***Accepted:*** *05 August 2021*
***Published:*** *25 October 2021*

***Citation:***
*Santos CMA, Heller AH, Pena HB and Pena SDJ (2021) A Protocol for Preconceptional Screening of Consanguineous Couples Using Whole Exome Sequencing. Front. Genet. 12:685123. doi: 10.3389/fgene.2021.685123*

## INTRODUCTION

Autosomal recessive (AR) inherited disorders can be a major cause of morbidity and mortality (Bundey and Alam, 1993). The occurrence of these disorders can significantly increase in the offspring of consanguineous couples. Approximately 1.1 billion people live currently in countries where consanguineous marriages are customary, and among them, one in every three marriages is between cousins (Hamamy et al., 2011). Consequently, in many locations, consanguineous couples and their offspring represent a significant proportion of any genetic counselor's case load. The identification of the risks for these couples should enable them to consider reproductive choices to prevent the birth of affected children, including prenatal diagnosis and preimplantation genetic testing.

In theory, genetic counseling could offer an acceptable approach to reduce the burden of recessively inherited disease. However, genetic counseling alone can only present risk estimates and does not offer to the consanguineous couple the possibility of taking practical measures to avoid genetic disease in their future children, since knowledge of the exact gene and variant are necessary for that. Moreover, there seems to exist a lack of standardization of genetic services. A questionnaire sent to certified genetics counselors and medical geneticists in the United States has revealed a wide variation in the risk figures quoted to consanguineous couples about their risk of having offspring with birth defects and intellectual deficiency (Bennett et al., 1999).

Advances in screening couples for heterozygosity over the past decades offer now the possibility of testing populations for all known severe recessive genetic disorders (Antonarakis, 2019). However, to offer effective pre-conceptive screening for consanguineous couples it is not necessary to identify all the pathogenic variants for which each member of a couple is a carrier, being sufficient to identify only which pathogenic variants are shared by them.

Fridman et al. (2021) studied 6,447 exome sequences of healthy, genetically unrelated Europeans from Holland and Estonia and calculated that almost all individuals (>85%) carry at least one pathogenic or likely pathogenic (PLP) variant, with an average of at least 1.3 PLPs for a severe AR disorder and 2.2 PLPs for any AR disorder. Their data did not allow estimation of the upper bound, but they felt it unlikely that the number of PLPs would exceed 8 per individual. If novel AR genes that have not yet been discovered contribute less PLPs than those that have been discovered recently, then the upper bound for the estimate would be more in the range of 4–5 PLPs per individual. They estimate first-cousin consanguineous couples to be at 16 times higher risk to conceive a child with an AR disorder compared to non-consanguineous couples. This translates to 3,400 newborns with a severe AR disorder per 100,000 births for first cousins (3.4%). These estimates are compatible with the calculation that couples who are first cousins show an extra 2–3% increased risk of having children with AR genetic problems (Hamamy et al., 2011). As expected, the risks gradually decreased for more distant relationships and the risk for third cousins was similar to that for non-consanguineous couples (Fridman et al., 2021).

It is important to keep in mind that this risk is diffuse, scattered among the thousands of potential AR diseases. Thus, to prospectively detect consanguineous couples at increased risk, it is necessary to carry out a genetic test with high coverage and high sensitivity, which specifically examines whether both members of the couple are simultaneously carriers of the same pathogenic or probably pathogenic variant, in heterozygosis. The ideal tests for this purpose appear to be whole genome sequencing (WGS) and whole exome sequencing (WES). Several articles have been written providing proof of concept for the use of WES in preconceptional screening for AR diseases (Makrythanasis et al., 2014; Teeuw et al., 2014; Sallevelt et al., 2017; Kirk et al., 2019; Monies et al., 2019).

Challenges remain regarding how to identify which shared genetic variants are capable of causing disease in homozygosity. It is important to avoid ambiguous results and incidental findings by keeping the analysis absolutely focused on the only element that distinguishes consanguineous couples from other couples: the increased risk of producing sons or daughters with an AR disease due to homozygosis. It is also essential to work only with diseases that have a known molecular basis, in order to be able to implement practical preventive measures.

We designed and wish to present a protocol for preconceptional screening of consanguineous couples using WES. According to this protocol, we first sequence the exome of the two members of the consanguineous couple, and then use bioinformatics to filter the genetic information to detect heterozygous genetic variants that are common to both. Our targets are the causative genes of all AR diseases obtained through a survey updated monthly in the Clinical Genomic Database of the National Human Genome Research Institute (2021). Presently, 2,846 genes related to AR diseases (including all AD-AR genes) are listed there.

## MATERIALS AND METHODS

### Patients

We collected buccal swabs from 39 consanguineous couples who presented for genetic counseling at GENE – Núcleo de Genética Médica de Minas Gerais in Belo Horizonte, Brazil. DNA was extracted using a modified salting out procedure (Miller et al., 1988).

The Research Ethics Committee of the Hospital das Clínicas of the Universidade Federal de Minas Gerais approved the study protocol. Informed consent was obtained according to current ethical and legal guidelines. The study was conducted in accordance with the Declaration of Helsinki.

### Whole Exome Sequencing

After DNA extraction and pertinent technical procedures, exome enrichment by the Agilent SureSelect XT V6 was performed at Theragen Bio (Seongnam-si, Gyeonggi-do, Republic of Korea), with a capture greater than 50 Mb. The library was sequenced in an Illumina NovaSeq 6000 instrument following the manufacturer's specifications and resulted in over 92 million readings, with more than 14 Gigabases of sequence data, allowing

an average coverage greater than 100×. Signal processing and base identification (base calling) was performed with the FastQC software, followed by the alignment of the tested exome to the reference human genome (hg19) using the BWA software. Variants were determined by the GATK Unified Genotyper software following parameters specified by the Broad Institute, which developed the software.

## Bioinformatics Analysis

Sequenced genes were filtered for rare variants (allele frequency < 0.01) utilizing databases such as 1000 Genomes Phase 3, NHLBI Exome Sequencing Project (ESP6500), Single Nucleotide Polymorphism database (dbSNP141) and gnomAD database, using the Mendel, MD software developed in-house (Cardenas et al., 2017) and the ENLIS Genome Research software (Enlis Genomics, Berkeley, CA, United States). To analyze the impact of the candidate variants we used the software Alamut Visual version 2.11.0 (Interactive Biosoftware, Paris, France, which showed the alignment of orthologous genes, the gnomAD frequencies and the ClinVar classification. Variants were also analyzed with the Franklin software (The Genoox Platform, 2021). To define what constitutes a pathogenic variant, we used the criteria of the American College of Medical Genetics and Genomics (Richards et al., 2015). Thus, only missense and in-frame indel variants already classified as such in the ClinVar database (ClinVar, 2021) are considered pathogenic or probably pathogenic. Truncating variants that are not in the ClinVar database are considered probably pathogenic when loss-of-function is a known disease mechanism.

# RESULTS AND DISCUSSION

The 39 consanguineous cases included two couples with a coefficient of genetic relationship (CGR) of 0.25, 26 couples with a CGR of 0.125, three couples with a CGR of 0.0625, and eight couples with a CGR of 0.03125 (**Table 1**). In 21 of these 39 couples (53.8%) we ascertained sharing of heterozygosis for at least one variant considered pathogenic for an AR disease (**Table 1**). In eight couples we found sharing of heterozygosity for at least two pathogenic variants.

Except for 25 first cousin couples, the number of the other types of relationship were small. This possibly can explain why the data do not show a decrease in the proportion of positive cases in first cousins once removed (CGR = 0.0625). We should mention that these coefficients of genetic relationships are based on family history and should be considered estimates. A major drawback of pedigree-based calculations is the absolute requirement for a correct pedigree structure, which in practice may be unreliable, or incomplete.

If we limit our analysis to the 25 first cousin couples, we can observe that 13 of them (52%) have one or more shared pathogenic or likely pathogenic (PLP) variant. Although 26 couples are not a large number, this percentage is more than double the rate of 20.9–24.9% at risk couples (ARC) estimated by Fridman et al. (2021) for first cousins, based on simulated consanguineous matings. We can derive mathematically the relationship of the rate of ARCs to the PLPs for a given CGR with the formula:

$$1 - (1 - \mathrm{CGR})^{\mathrm{PLP}} = \mathrm{ARC}$$

For first cousins (CGR = 0.125) and ARCs in the range of 20.9–24.9%, we can use the formula to calculate PLPs in the range of 1.75–2.14, which is compatible with the calculation in the European population of 2.2 PLPs for any AR disorder. However, for the ARC of 0.52 in our sample of first cousin couples we would obtain a larger value of 5.5 PLPs for any AR disease. However, one should not place emphasis on this discrepancy, since we cannot rule out that it was originated from random variation or from ascertainment biases.

If now we focus our attention on the nature of the PLP variants encountered in common in our first cousin group we observe that 11/20 (55%) are truncating loss-of-function (LoF) variants, which is compatible with the data of Fridman et al. (2021) who found that more than half of the PLPs (55.2 and 59.1% in the Dutch and Estonian cohorts) were truncating loss-of-function (LoF) rare variants. However, our data differs from theirs since 64% of our truncating loss-of-function (LoF) variants were described in Clinvar, while none in their study were. Since our numbers are small, this discrepancy could be aleatory.

Once the specific pathogenic variant was identified, it became possible for the couple to undergo prenatal diagnosis or, if desired, preimplantation genetic diagnosis (PGD) involving *in vitro* fertilization and embryo screening.

Our results demonstrate that preconceptional screening by WES is a useful new procedure that should be incorporated in the genetic counseling of all consanguineous couples. However, it is still not perfect for several reasons:

First, the number of AR diseases known is still relatively small. According to the calculations of Bamshad et al. (2019), if we assume that each candidate gene underlies a single Mendelian condition (MC), there are circa 1.5–3 times as many novel genes (4,450–10,467) for MCs yet to be discovered as there are genes (3,519) known already to underlie an MC. If we extrapolate that the same proportion of these genes underlie multiple MCs as is the case for known genes for MCs (i.e., 16% underlie two MCs, 4.7% underlie three, 1.8% underlie four, etc.), we can predict that a minimum of 6,100–14,400 MCs remain to be discovered. And these figures are still an underestimate of the number of unsolved MCs because the authors did not account for the fact that mutant phenotypes for over half (∼12,000) of all protein-coding mouse genes have yet to be assessed.

Second, our ClinVar criteria for identification of disease-causing non-truncating genetic variants are by necessity very conservative, because we cannot risk burdening consanguineous couples with false-positive variants. Even so, the possibility of false-positive finding exists because of the unpredictable contingency of incomplete penetrance (Cooper et al., 2013).

Third, it is important to remember that WES does not detect all genetic variants present in an individual.

**TABLE 1 |** Results of the search for the same pathogenic and likely pathogenic variants in both members of 39 consanguineous couples.

| Degree of consanguinity | CGR | PLP in common | Disease | OMIM | Gene | Variant |
|---|---|---|---|---|---|---|
| Uncle-niece | 0.25 | Yes | Pyruvate kinase deficiency in red blood cells | 266200 | *PKLR* | c.1456C > T p.(Arg486Trp) |
| Double first cousins | 0.25 | Yes | Breast and colorectal cancer, susceptibility to<br>Intellectual deficiency, autosomal recessive 3 Deafness, autosomal recessive 1A | 604373<br>608443<br>220290 | *CHEK2*<br>*CC2D1A*<br>*GJB2* | c.975 + 1G > C<br>c.1357-2A > C<br>c.109G > A p.(Val37Ile) |
| First cousins | 0.125 | Yes | Nemaline myopathy 2, autosomal recessive<br>asphyxiating thoracic dystrophy 3 | 256030<br>613091 | *NEB*<br>*DYNC2H1* | c.1258-2A > G<br>c.6047A > G p.(Tyr2016Cys) |
| 1First cousins | 0.125 | Yes | Dyskeratosis congenita, autosomal recessive 3 | 613988 | *WRAP53* | c.1564dup p.(Ala522Glyfs*8) |
| First cousins | 0.125 | Yes | Glycogen storage disease Ib<br>cranioectodermal dysplasia 1 | 232220<br>218330 | *SLC37A4*<br>*IFT122* | c.935_936del p.(Thr312Serfs*13)<br>c.1301-1G > C |
| First cousins | 0.125 | Yes | Homocystinuria<br>Adrenal hyperplasia, congenital | 236200<br>201910 | *CBS*<br>*CYP21A2* | c.833T > C p.(Ile278Thr)<br>c.955C > T p.(Gln319*) |
| First cousins | 0.125 | Yes | Mitochondrial complex V deficiency, nuclear type 2<br>citrullinemia | 614052<br>215700 | *TMEM70*<br>*ASS1* | c.317-2A > G<br>c.323G > T p.(Arg108Leu) |
| First cousins | 0.125 | Yes | Stargardt disease 1 | 248200 | *ABCA4* | c.2588G > C p.(Gly863Ala) |
| First cousins | 0.125 | Yes | Papillon-lefevre syndrome | 245000 | *CTSC* | c.194_197dup p.(Tyr67Profs*11) |
| First cousins | 0.125 | Yes | Galactosemia | 230400 | *GALT* | c.563A > G p.(Gln188Arg) |
| First cousins | 0.125 | Yes | Spondylocostal dysostosis 6 | 616566 | *RIPPLY2* | c.238A > T p.(Arg80*) |
| First cousins | 0.125 | Yes | Thyroid dyshormonogenesis 6 | 607200 | *DUOX2* | c.2428G > T p.(Glu810*) |
| First cousins | 0.125 | Yes | Sickle cell anemia<br>hyperlipoproteinemia, type I | 603903<br>238600 | *HBB*<br>*LPL* | c.20A > T p.(Glu7Val)<br>c.953A > G p.(Asn318Ser) |
| First cousins | 0.125 | Yes | Microcephaly 1<br>deafness 1A | 251200<br>220290 | *MCPH1*<br>*GJB2* | c.2145G > A p.(Trp715*)<br>c.35del p.(Gly12Valfs*2) |
| First cousins | 0.125 | Yes | Cystathionine beta-synthase deficiency<br>familial adenomatous polyposis 2 | 236200<br>608456 | *CBS*<br>*MUTYH* | c.833T > C p.(Ile278Thr)<br>c.721C > T p.(Arg241Trp) |
| First cousins | 0.125 | No | | | | |
| First cousins | 0.125 | No | | | | |
| First cousins | 0.125 | No | | | | |
| First cousins | 0.125 | No | | | | |
| First cousins | 0.125 | No | | | | |
| First cousins | 0.125 | No | | | | |
| First cousins | 0.125 | No | | | | |
| Half uncle-niece | 0.125 | No | | | | |
| First cousins | 0.125 | No | | | | |
| First cousins | 0.125 | No | | | | |
| First cousins | 0.125 | No | | | | |
| First cousins | 0.125 | No | | | | |
| First cousins | 0.125 | No | | | | |
| First cousins once removed | 0.0625 | Yes | Acyl-Coa Dehydrogenase, very long-chain deficiency | 201475 | *ACADVL* | c.685C > T p.(Arg229*) |
| First cousins once removed | 0.0625 | Yes | Amelogenesis imperfecta, type IJ | 617297 | *ACPT* | c.945dup p.(Glu316*) |
| First cousins once removed | 0.0625 | No | | | | |
| Second cousins | 0.03125 | Yes | Deafness, autosomal recessive 77 | 613079 | *LOXHD1* | c.2047 + 1G > A |
| Second cousins | 0.03125 | Yes | Fetal akinesia deformation sequence 1 | 208150 | *SCN4A* | c.1173del p.(Phe392Serfs*12) |
| Second cousins | 0.03125 | Yes | Joubert Syndrome, Type 13 | 614173 | *TCTN1* | c.1775_1778del p.(Val592_593delins30) |
| Second cousins | 0.03125 | Yes | Inflammatory bowel disease 28, early onset | 613148 | *IL10RA* | c.349C > T p.(Arg117Cys) |
| Second cousins | 0.03125 | No | | | | |
| Second cousins | 0.03125 | No | | | | |
| Second cousins | 0.03125 | No | | | | |
| Second cousins | 0.03125 | No | | | | |

Among the situations in which WES may not be able to identify a pathogenic variant, we can highlight: (a) mutations by deletion or duplication of the entire gene or part of it; microsatellite expansion mutations (dynamic mutations); (c) diseases associated with mutations in two different genes (digenic inheritance); (d) mutations that are not located in an exon (for example, a pathogenic mutation in a promoter or intronic region), (e) mutations in genes whose exons are captured

in low efficiency by the currently available exonic selection kits; (f) mutations in a gene that is very similar to other genes (paralogs) in the human genome; and (h) Mutations caused by the insertion of a transposing element.

Because of these caveats, a negative result in the search for common PLPs in consanguineous couples should not be interpreted as the absence of genetic risks. In fact, couples that do not initiate reproduction immediately are offered the possibility of annual revisions of the exome data, to permit incorporation of new knowledge in human genomics. Also, the finding of a disease-causing variant in a couple does not mean that it is the only genetic risk present. The observation that some consanguineous couples occasionally have more than one shared pathogenic gene means that even couples who have already had a child diagnosed with an AR disease should undergo pre-conceptional screening by WES because of the possibility of other shared pathogenic variants.

## CONCLUSION

Although there is considerable room for progress and inevitable intrinsic uncertainty (Wray and Loo, 2015), our results demonstrate that preconceptional screening by WES is already a viable and useful new procedure that should be incorporated in the genetic counseling routine of all consanguineous couples.

This protocol can also be used to assist family planning for couples who belong to the same ethnic group (Ashkenazi Jews, Sephardic Jews, Mediterraneans, Sub-Saharan Africans, members of endogamous communities, etc.), whose increased risk for group-specific AR diseases is well known.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Research Ethics Committee of the Hospital das Clínicas of the Universidade Federal de Minas Gerais. The patients/participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

CS and AH were responsible for the general conduction of the study, contact with the consanguineous couples, preparing reports, and participation in writing the manuscript. HP was responsible for technical procedures of extraction of DNA, quantitative measurements of purity, and conduction of the study. SP conceived the protocol, analyzed the exome results, and participated in writing the manuscript. All authors contributed to the article and approved the submitted version.

## FUNDING

Funding for this study has been provided by Conselho Nacional de Pesquisas (CNPq) and Fundação de Amparo à Pesquisa de Minas Gerais (FAPEMIG RED-00314-16).

## REFERENCES

Antonarakis, S. E. (2019). Carrier screening for recessive disorders. *Nat. Rev. Genet.* 20, 549–561. doi: 10.1038/s41576-019-0134-2

Bamshad, M. J., Nickerson, D. A., and Chong, J. X. (2019). Mendelian gene discovery: fast and furious with no end in sight. *Am. J. Hum. Genet.* 105, 448–455. doi: 10.1016/j.ajhg.2019.07.011

Bennett, R. L., Hudgins, L., Smith, C. O., and Motulsky, A. G. (1999). Inconsistencies in genetic counseling and screening for consanguineous couples and their offspring: the need for practice guidelines. *Genet. Med.* 1, 286–292. doi: 10.1097/00125817-199909000-00007

Bundey, S., and Alam, H. (1993). A five-year prospective study of the health of children in different ethnic groups, with particular reference to the effect of inbreeding. *Eur. J. Hum. Genet.* 1, 206–219. doi: 10.1159/000472414

Cardenas, G C C L R., Linhares, N. D., Ferreira, R. L., and Pena, S. D. J. (2017). Mendel,MD: a user-friendly open-source web tool for analyzing WES and WGS in the diagnosis of patients with Mendelian disorders. *PLoS Comput. Biol.* 13:e1005520. doi: 10.1371/journal.pcbi.1005520

ClinVar (2021). *ClinVar Database*. Available online at: https://www.ncbi.nlm.nih.gov/clinvar/ (accessed March 15, 2021).

Cooper, D. N., Krawczak, M., Polychronakos, C., Tyler-Smith, C., and Kehrer-Sawatzki, H. (2013). Where genotype is not predictive of phenotype: towards an understanding of the molecular basis of reduced penetrance in human inherited disease. *Hum. Genet.* 132, 1077–1130. doi: 10.1007/s00439-013-1331-2

Fridman, H., Yntema, H. G., Mägi, R., Andreson, R., Metspalu, A., Mezzavila, M., et al. (2021). The landscape of autosomal-recessive pathogenic variants in European populations reveals phenotype-specific effects. *Am. J. Hum. Genet.* 12, S0002–S9297. doi: 10.1016/j.ajhg.2021.03.004

Hamamy, H., Antonarakis, S. E., Cavalli-Sforza, L. L., Temtamy, S., Romeo, G., Kate, L. P., et al. (2011). Consanguineous marriages, pearls and perils: Geneva international consanguinity workshop report. *Genet. Med.* 13, 841–847. doi: 10.1097/GIM.0b013e318217477f

Kirk, E. P., Barlow-Stewart, K., Selvanathan, A., Josephi-Taylor, S., Worgan, L., Rajagopalan, S., et al. (2019). Beyond the panel: preconception screening in consanguineous couples using the TruSight One "clinical exome". *Genet. Med.* 21, 608–612. doi: 10.1038/s41436-018-0082-9

Makrythanasis, P., Nelis, M., Santoni, F. A., Guipponi, M., Vannier, A., and Béna, F. (2014). Diagnostic exome sequencing to elucidate the genetic basis of likely recessive disorders in consanguineous families. *Hum. Mutat.* 35, 1203–1210. doi: 10.1002/humu.22617

Miller, S. A., Dykes, D. D., and Polesky, H. F. (1988). A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res.* 16:1215. doi: 10.1093/nar/16.3.1215

Monies, D., Abouelhoda, M., Assoum, M., Moghrabi, N., Rafiullah, R., Almontashiri, N., et al. (2019). Lessons learned from large-scale, first-tier clinical exome sequencing in a highly consanguineous population. *Am. J. Hum. Genet.* 104, 1182–1201. doi: 10.1016/j.ajhg.2019.04.011

National Human Genome Research Institute (2021). *Clinical Genomic Database of the National Human Genome Research Institute*. Available online at: https://research.nhgri.nih.gov/CGD (accessed March 15, 2021).

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., et al. (2015). ACMG laboratory quality assurance committee. 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

Sallevelt, S. C. E. H., de Koning, B., Szklarczyk, R., Paulussen, A. D. C., de Die-Smulders, C. E. M., and Smeets, H. J. M. (2017). A comprehensive strategy for exome-based preconception carrier screening. *Genet. Med.* 19, 583–592. doi: 10.1038/gim.2016.153

Teeuw, M., Waisfisz, Q., Zwijnenburg, P. J., Sistermans, E. A., Weiss, M. M., Henneman, L., et al. (2014). First steps in exploring prospective exome sequencing of consanguineous couples. *Eur. J. Med. Genet.* 57, 613–616. doi: 10.1016/j.ejmg.2014.09.003

The Genoox Platform (2021). *Franklin Software*. Available online at: http://franklin.genoox.com (accessed March 15, 2021).

Wray, C. M., and Loo, L. K. (2015). The diagnosis, prognosis, and treatment of medical uncertainty. *J. Grad. Med. Educ.* 7, 523–527. doi: 10.4300/JGME-D-14-00638.1

**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.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

*Copyright © 2021 Santos, Heller, Pena and Pena. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*

frontiers in Genetics

ORIGINAL RESEARCH
published: 15 November 2021
doi: 10.3389/fgene.2021.765433

# Genetic and Phenotypic Variability in Chinese Patients With Branchio-Oto-Renal or Branchio-Oto Syndrome

*Haifeng Feng*[1†], *Hongen Xu*[2,3†], *Bei Chen*[1†], *Shuping Sun*[1], *Rongqun Zhai*[1], *Beiping Zeng*[3], *Wenxue Tang*[2,3,4*] and *Wei Lu*[1*]

[1]Department of Otorhinolaryngology-Head and Neck Surgery, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, China, [2]Center for Applied Precision Medicine, The Second Affiliated Hospital of Zhengzhou University, Zhengzhou, China, [3]Precision Medicine Center, Academy of Medical Science, Zhengzhou University, Zhengzhou, China, [4]Henan Institute of Medical and Pharmaceutical Sciences, Zhengzhou University, Zhengzhou, China

OPEN ACCESS

**Edited by:**
Natália Duarte Linhares, Genuity Science, Ireland

**Reviewed by:**
Jinsei Jung, Yonsei University, South Korea
Qifei Li, Boston Children's Hospital and Harvard Medical School, United States

***Correspondence:**
Wenxue Tang
twx@zzu.edu.cn
Wei Lu
luwei611@zzu.edu.cn

[†]These authors have contributed equally to this work and share the first authorship

**Specialty section:**
This article was submitted to Genetics of Common and Rare Diseases, a section of the journal Frontiers in Genetics

**Received:** 27 August 2021
**Accepted:** 30 September 2021
**Published:** 15 November 2021

**Citation:**
Feng H, Xu H, Chen B, Sun S, Zhai R, Zeng B, Tang W and Lu W (2021) Genetic and Phenotypic Variability in Chinese Patients With Branchio-Oto-Renal or Branchio-Oto Syndrome. Front. Genet. 12:765433. doi: 10.3389/fgene.2021.765433

**Background:** Branchio-oto-renal syndrome (BOR) and branchio-oto syndrome (BOS) are rare autosomal dominant disorders defined by varying combinations of branchial, otic, and renal anomalies. Here, we characterized the clinical features and genetic etiology of BOR/BOS in several Chinese families and then explored the genotypes and phenotypes of BOR/BOS-related genes, as well as the outcomes of auditory rehabilitation in different modalities.

**Materials and Methods:** Probands and all affected family members underwent detailed clinical examinations. Their DNA was subjected to whole-exome sequencing to explore the underlying molecular etiology of BOR/BOS; candidate variants were validated using Sanger sequencing and interpreted in accordance with the American College of Medical Genetics guidelines. In addition, a literature review concerning *EYA1* and *SIX1* alterations was performed to explore the genotypes and phenotypes of BOR/BOS-related genes.

**Results:** Genetic testing identified the novel deletion (c.1425delC, p(Asp476Thrfs*4); NM_000,503.6), a nonsense variant (c.889C > T, p(Arg297*)), and two splicing variants in the *EYA1* gene (c.1050+1G > T and c.1140+1G > A); it also identified one novel missense variant in the *SIX1* gene (c.316G > A, p(Val106Met); NM_005,982.4). All cases exhibited a degree of phenotypic variability between or within families. Middle ear surgeries for improving bone-conduction component hearing loss had unsuccessful outcomes; cochlear implantation (CI) contributed to hearing gains.

**Conclusion:** This is the first report of BOR/BOS caused by the *SIX1* variant in China. Our findings increase the numbers of known *EYA1* and *SIX1* variants. They also emphasize the usefulness of genetic testing in the diagnosis and prevention of BOR/BOS while demonstrating that CI for auditory rehabilitation is a feasible option in some BOR/BOS patients.

**Keywords: EYA1 gene, SIX1 gene, branchio-oto-renal syndrome, whole-exome sequencing, hearing rehabilitation**

## INTRODUCTION

Branchio-oto-renal syndrome-1 (BOR1; OMIM#113650), also known as Melnick–Fraser syndrome, is a rare autosomal dominant affected family members disease with an incidence of approximately one in 40,000; it affects 2% of profoundly deaf children (Melnick et al., 1975; Fraser et al., 1980). BOR exhibits a variable spectrum of clinical manifestations that are mainly characterized by the presence of branchial cleft fistulae or cysts, preauricular pits, ear malformations, and hearing loss, along with renal malformations of varying severities (Fraser et al., 1978). In some instances, patients exhibit symptoms similar to those of BOR, with the exception of renal anomalies; they are diagnosed with branchio-oto syndrome-1 (BOS1; OMIM#602588) or branchio-oto syndrome-3 (BOS3; OMIM#608389). The clinical diagnosis of BOR/BOS follows a set of criteria proposed by Chang et al. (2004). Diagnosis of BOR/BOS can be made with at least three major criteria, two major and at least two minor criteria, or one major criterion with at least one first-degree affected family member.

BOR/BOS has marked genetic heterogeneity, and the exact pathogenesis remains unknown in more than half of affected patients. Three genes are currently considered to be associated with this condition: *EYA1* (OMIM #601653), *SIX1* (OMIM #601205), and *SIX5* (OMIM #600963) (Abdelhak et al., 1997; Ruf et al., 2004; Hoskins et al., 2007). Pathogenic variants in the *EYA1* gene, the human homolog of the *Drosophila* "eyes absent" gene, were recognized as a major genetic cause of BOR/BOS; they affect 40% of BOR/BOS patients (Chang et al., 2004). The *SIX1* gene, also known as sine oculis homeobox homolog 1, encodes a transcription factor (Six1) that functions as a DNA-binding protein in combination with Eya1, leading to 3.0–4.5% of BOR/BOS cases (Ruf et al., 2004; Wang et al., 2018). The role of the *SIX5* gene, the sine oculis homeobox homolog 5, in BOR/BOS is controversial; few variants have been reported (Hoskins et al., 2007; Krug et al., 2011). The *EYA1* and *SIX1* genes are co-expressed in the developing kidney and ear, beginning during the emergence of basal plates (Xu et al., 1999). Eya1 is epistatic to Six1, and its function is dependent on interaction with Six1. Eya1 does not have direct DNA-binding ability. Instead, it functions as a transcription co-activator and interacts with Six1, thus providing a molecular mechanism for activation of specific target genes that modulate precursor cell proliferation, survival, and differentiation during multiple types of organogenesis (Li et al., 2003; Ruf et al., 2004). In the absence of this interaction, the transcriptional activation of downstream targets required for the development of the branchial, otic, and renal systems is diminished. Functional studies have confirmed that genetically defective *EYA1* and *SIX1* mice exhibit symptoms similar to BOR/BOS (Xu et al., 1999; Ando et al., 2005).

BOR/BOS is a common form of syndromic hearing loss. Hearing impairment is the most common clinical feature, present in 98.5% of affected individuals; other common features include preauricular pits or tags (83.6%), branchial fistulae or cysts (68.5%), renal anomalies (38.2%), and external auditory canal stenosis (31.5%) (Chang et al., 2004). The forms of hearing loss can be mixed (50%), conductive (30%), or sensorineural (20%), ranging in severity from mild to profound (Abdelhak et al., 1997). Most patients with BOR/BOS present with morphological abnormalities of the middle and inner ear (Chen et al., 1995; Propst et al., 2005), prompting surgeons to explore auditory rehabilitation modalities. Some studies have used middle ear exploratory tympanotomy and ossicular reconstruction to improve mixed and conductive hearing difficulties. However, the postoperative hearing gains have been unsatisfactory (Propst et al., 2005; Song et al., 2013). With the rise of auditory implant technology, cochlear implantation (CI) has provided satisfactory results in some patients with syndromic deafness (e.g., BOR/BOS) (Bajaj et al., 2012).

Here, we screened for *EYA1*, *SIX1*, and *SIX5* variants in patients with BOR or BOS and then investigated the affected individuals' clinical manifestations and auditory rehabilitation outcomes. In addition, we analyzed the genotypes and phenotypes of BOR/BOS caused by *EYA1* variants through a literature review.

## MATERIALS AND METHODS

### Patient Recruitment and Clinical Examinations

In this study, five individuals clinically diagnosed with BOR or BOS were recruited at the Affiliated Hospital of Zhengzhou University. They underwent family history inquiries and detailed physical examinations. Audiological assessment comprised pure-tone audiometry to estimate the extent of hearing loss. Objective audiometry was used for pediatric patients, including auditory brainstem response, auditory steady-state response, and distortion product otoacoustic emissions. Radiological work-up comprised temporal bone thin-section computed tomography and/or magnetic resonance imaging to analyze the middle and inner ear morphologies. Serum creatinine, urea, and renal ultrasonography analyses were performed to screen for renal abnormalities. Written informed consent was obtained from all participating individuals or their guardians prior to enrollment in the study. The project was approved by the Affiliated Hospital of Zhengzhou University (reference number: 2018008), and all procedures were performed in compliance with the Declaration of Helsinki.

### Genetic Examinations and Sanger Sequencing

Peripheral venous blood of the affected individuals was collected for genetic sequencing if available. The procedures for DNA extraction, fragmentation, library construction, targeted enrichment, and sequencing were identical to the approaches used in previous studies (Pan et al., 2020). After sequencing adaptors and inferior reads had been eliminated from raw data, clean reads were mapped to the human reference genome (version GRCh37) using Burrows-Wheeler Aligner (version 0.7.17-r1188). Duplicate reads were flagged by Sambamba (version 0.6.6) (Tarasov et al., 2015). Single-base variations and small insertions or deletions were investigated with the

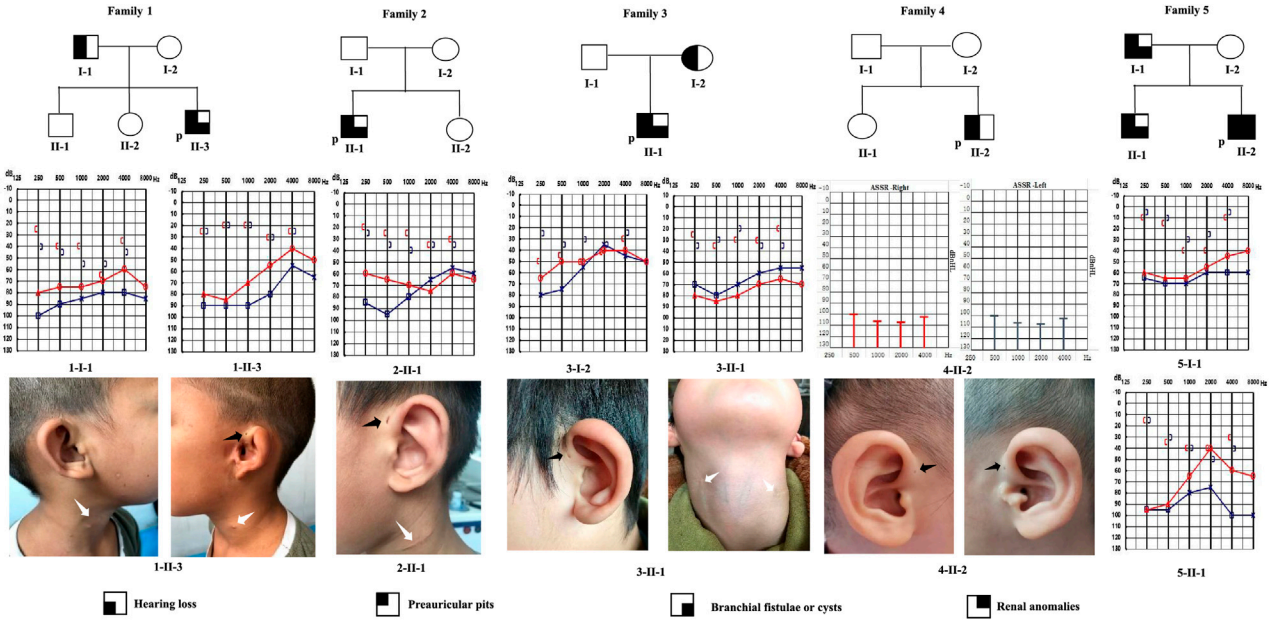



**FIGURE 1 |** Pedigrees of five families with BOR or BOS. Audiograms (pure-tone audiometry or auditory steady-state response) and medical photographs of patients are shown. P: proband; black arrow: preauricular pits; white arrow: branchial cleft fistulae or scar.

Genome Analysis Toolkit version four HaplotypeCaller (DePristo et al., 2011). Variant annotation, filtering, and interpretation were performed as described previously (Oza et al., 2018; Pan et al., 2020). To validate candidate variants detected by whole-exome sequencing, polymerase chain reaction amplification and Sanger sequencing were performed. Amplified polymerase chain reaction products were purified by a polymerase chain reaction purification kit (LifeSciences, Hangzhou, China) and then sequenced using the SeqStudio Genetic Analyzer (Applied Biosystems/Life Technologies, Carlsbad, CA, United States). Variant nomenclature was based on *EYA1* canonical transcript NM_000,503.6 and *SIX1* canonical transcript NM_005,982.4.

## Literature Review and Statistical Tests

Studies spanning 1975–2021 were retrieved using NCBI PubMed and the Human Gene Mutation database with "*EYA1*," "*SIX1*," and "branchio-oto-renal (BOR) syndrome" as the keywords. First, the *EYA1* and *SIX1* alterations were summarized. Phenotypic analysis was then performed, focusing on *EYA1* alterations that occurred in East Asian populations. When appropriate, testing for difference in proportions was carried out using either the chi-square or Fisher's exact test. All tests were two-sided, and *p*-values lower than 0.05 were considered significant.

# RESULTS

## Clinical Characteristics

We identified nine patients from five families with BOR or BOS, using the criteria established by Chang et al. (2004). The genealogies of the enrolled families are illustrated in **Figure 1**. Detailed phenotypic features are shown in **Table 1**. Eight patients were diagnosed with BOS; one patient was diagnosed with BOR. Of the eight patients with BOS, five showed the typical triad of BOS (preauricular pits, branchial fistulae, hearing loss); the remaining three had preauricular pits and hearing loss, without branchial fistulae or cysts. Proband 1-II-3 also had auricle deformity accompanied by preauricular tags and external auditory canal stenosis in the left ear; proband 4-II-2 had mild bilateral microtia and left preauricular tags. Renal ultrasonography showed no abnormalities in the kidneys of eight patients with respect to size, architecture, and origination; no other positive findings were noted. Proband 5-II-2, a 3-year-old, was diagnosed with BOR at 2 months old. Medical records showed that the patient had the abovementioned triad and bilateral auricle deformity. Renal ultrasonography showed left-sided renal hypoplasia and hydronephrosis, along with pyelo-ureteral separation. For radiological work-up, seven patients underwent temporal bone computed tomography. All seven patients exhibited various abnormal configurations of the middle and/or inner ear, such as deformed ossicular chain, hypoplastic cochlea, dysplastic semicircular canals, dilated internal auditory canals, or enlarged vestibular aqueduct.

## Bioinformatics Analysis Identifies *EYA1* and *SIX1* Variants

Probands from five families underwent whole-exome sequencing, yielding 17.69, 13.35, 12.42, 11.65, and 9.75 Gbp of raw data. The mapping rate of sequencing reads to the human reference genome was >99%. Mean sequencing depths for the targeted region were

**TABLE 1 |** Detailed clinical features and genetic analysis of five families.

| Family | Patient | B | P | D, right/ left | R | CT scans | DNA | Protein | Diagnosis | ACMG evidence | ACMG classification |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Proband | + | + | +, m/m | – | NEAC, OA, SC, CH, DIAC | EYA1:c.1050+1G > T | – | BOS | PVS1, PM2, PP3, PP4 | Pathogenic |
| | Father | – | + | +, m/m | – | NA | EYA1:c.1050+1G > T | – | BOS | | |
| 2 | Proband | + | + | +, m/m | – | OA, CH, DIAC, EV, SC, EVA | EYA1:c.1140+1G > A | – | BOS | PVS1, PM2, PP3, PP4 | Pathogenic |
| 3 | Proband | + | + | +, c/c | – | OC | EYA1:c.889C > T | p.(Arg297*) | BOS | PVS1, PM2, PP3, PP4 | Pathogenic |
| | Mother | – | + | +, s/m | – | OA, CH, DIAC, EVA | EYA1:c.889C > T | p.(Arg297*) | BOS | | |
| 4 | Proband | – | + | +, m/m | – | OA, SC, CH, DIAC | SIX1:c.316G > A | p.(Val106Met) | BOS | PM2, PM5, PM6, PP3, PP4 | Likely pathogenic |
| 5 | Proband | + | + | +, NA | + | NA | EYA1:c.1425delC | p.(Asp476Thrfs*4) | BOR | PVS1, PM2, PP3, PP4 | Pathogenic |
| | Brother | + | + | +, m/m | – | NEAC, OA | EYA1:c.1425delC | p.(Asp476Thrfs*4) | BOS | | |
| | Father | + | + | +, m/m | – | OA, EV, CH | NA | NA | BOS | | |

EYA1, *canonical transcript NM_000503.6;* SIX1, *canonical transcript NM_005,982.4; D, hearing loss; P, preauricular pits; B, branchial fistulae or cysts; R, renal anomalies; m, mixed; c, conductive; s, sensorineural; OA, ossicular anomaly; EVA, enlarged vestibular aqueduct; CH, cochlear hypoplasia; NEAC, narrowed external auditory canal; DIAC, dilated internal auditory canal; EV, enlarged vestibule; SC, abnormal semicircular canal; NA, not available.*

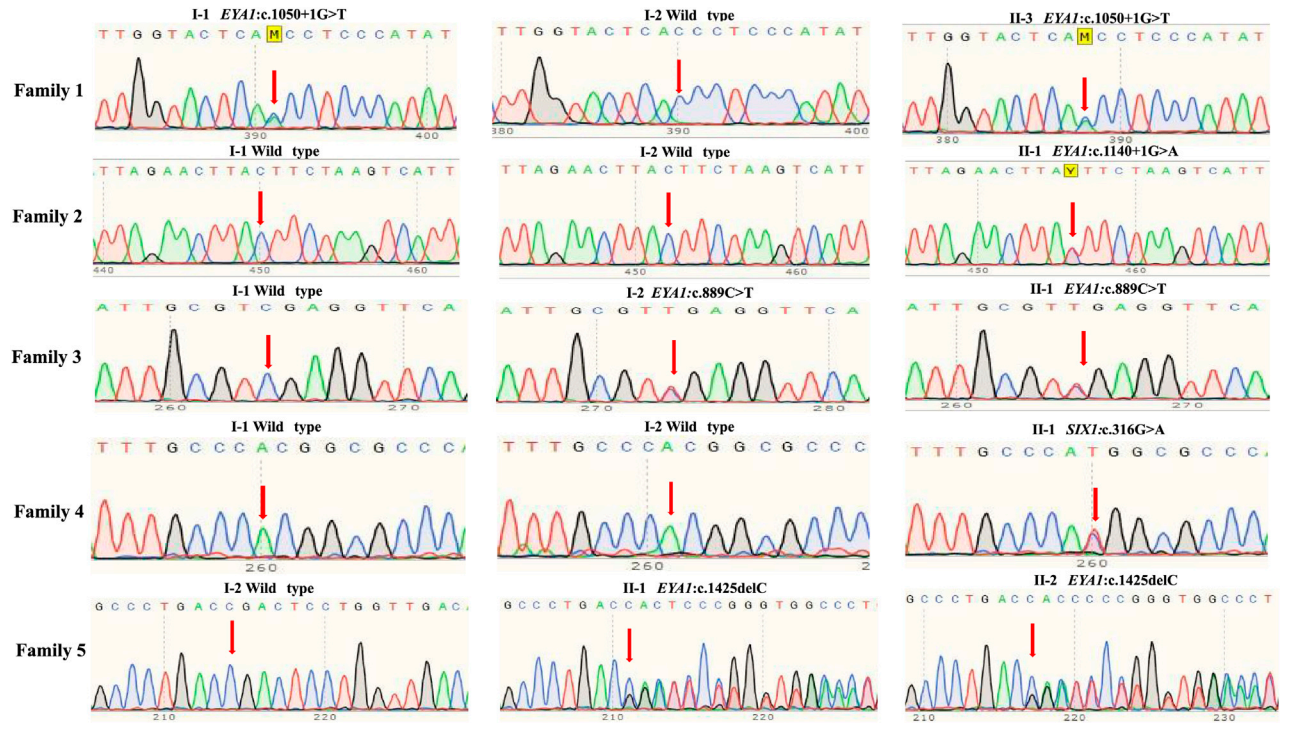



**FIGURE 2 |** Sanger sequencing of five unrelated families with BOR or BOS. Normal family members harbored no mutations in *EYA1* or *SIX1* gene.

approximately 100-fold; more than 95% of the regions were covered by at least 20-fold. We investigated the coverage statistics for genes known to be associated with BOR or BOS: *EYA1*, *SIX1*, and *SIX5*. Regions with poor coverage for whole-exome sequencing (less than 20-fold) were subjected to additional polymerase chain reaction and Sanger sequencing, improving the reliability of sequencing results. After variant calling and annotation had been performed, standalone benign variants were filtered out based on the criteria specified by the Clinical Genome Resource Sequence Variant Interpretation Working Group (Ghosh et al., 2018). Subsequently, we screened in the five probands for variants in the *EYA1*, *SIX1*, and *SIX5* genes. We identified four candidate variants in the *EYA1* gene (c.1050+1G > T; c.1140+1G > A; c.889C > T, p.(Arg297*); c.1425del,

p.(Asp476Thrfs*4)) and one candidate variant in the *SIX1* gene (c.316G > A, p.(Val106Met)). No *SIX5* variants were detected in any of the affected individuals.

To confirm candidate variants and test for co-segregation, we performed Sanger sequencing (**Figure 2**). The *SIX1*:c.316G > A and *EYA1*:c.1140+1G > A variants in the probands were absent from both of their parents, which confirmed that these were *de novo* variants. The *EYA1*:c.1050+1G > T and *EYA1*:c.889C > T variants were confirmed to originate in the affected father and mother, respectively. We could not confirm whether the *EYA1*: c.1425del variant was from the patient's affected father because peripheral blood was not available; however, this variant was also identified in the patient's affected brother. The c.1425del variant of *EYA1* and c.316G > A variant of *SIX1* were not found in public databases such as gnomAD, ClinVar, and the Human Gene Mutation database, indicating that these comprised novel variants. The novel c.1425del variant in exon 15 of *EYA1* changed the arginine at codon 476 to threonine, and the C-to-T transversion at 889 in exon 10 led to the creation of a stop codon. Both were predicted to cause premature truncation of the protein. c.1050+1G > T and c.1140+1G > A variants of *EYA1* were located at canonical splice sites. The splicing effects of both variants were further predicted by SpliceAI (Jaganathan et al., 2019) and dbscSNV (Jian et al., 2014). Both variants were expected to cause aberrant splicing with donor loss scores >0.99 (SpliceAI) and ADA and RF scores >0.9 (dbscSNV). Three *EYA1* variants (c.1050+1G > T, c.1140+1G > A, c.889C > T) were previously reported in several patients with BOR or BOS from different countries (Rickard et al., 2000; Fukuda et al., 2001; Orten et al., 2008; Song et al., 2013; Unzaki et al., 2018). The c.316G > A variant of *SIX1* led to a valine-to-methionine substitution, co-segregating with the phenotype in this Chinese BOS family, and its REVEL score is 0.836. Based on the American College of Medical Genetics sequence variant interpretation guidelines, four variants (c.1050+1G > T, c.1140+1G > A, c.889C > T, and c.1425del) were classified as pathogenic; the remaining variant (c.316G > A) was classified as likely pathogenic. The American College of Medical Genetics evidence for variant interpretation is shown in **Table 1**.

## Auditory Rehabilitation

All affected individuals had hearing loss of varying magnitudes and presented with various forms (mixed in 15 ears, conductive in two ears, and sensorineural in one ear). In proband 1-II-3, air-bone gap persistence led to exploratory tympanotomy in the left ear before the final diagnosis. Stapedotomy and stapes prosthesis insertion were performed on the ipsilateral side because malformation, malposition, and fixation of the stapes were found. However, the air-bone gap was not narrowed. The soft-banded bone-conduction hearing aids (SOPHNON-1) trial showed over 30 dB of hearing gains. In proband 5-I-1, computed tomography revealed malformed ossicles that were surrounded by granulation tissue in the right ear. The patient was fitted with an ossicular replacement prosthesis after inflammation had subsided. Middle ear surgeries did not yield the expected results, according to self-reporting data. Proband 2-II-1 (age 7) exhibited mixed hearing loss for 3 years, which was aggravated by respiratory infections. The patient was diagnosed with otitis media with effusion; tympanostomy tube insertion and exploratory tympanotomy were performed. Ossiculoplasty was not performed because of the poor surgical outcome; the patient was then fitted for air-conduction hearing aids. Proband 4-II-2 exhibited bilateral profound deafness, which was not improved by hearing aids. Otologists made the decision to undergo CI at 1 year old in conjunction with the patient's parents. Postoperative follow-up showed that the CI surgery was successful and had provided considerable hearing improvement (**Supplementary Figure S1**).

## Genotype and Phenotype Analysis

To explore the mutational spectrum of BOR/BOS-related genes, we summarized the reported variants of the *EYA1* and *SIX1* genes shown in **Figure 3**. Variants in the *EYA1* gene were found in various forms; frameshift was the most common type, followed by nonsense, splicing, large deletion, and missense (**Figure 4A**). In the *SIX1* gene, the most commonly reported variant was missense (12/15) (**Figure 3**). The results exhibited genetic variability. Subsequently, we analyzed the phenotypic characteristics of BOR/BOS patients with the *EYA1* variant in East Asian populations (**Table 2** and **Figure 4B**). The results indicated that hearing loss is the most common symptom, with an estimated prevalence of 93.42%, followed by preauricular pits (85.52%), branchial fistulae or cysts (66.45%), and renal anomalies (32.85%). A small number of cases have been reported in China (**Table 2**); the morbidities were hearing loss (91.67%), preauricular pits (86.11%), branchial fistulae or cysts (63.89%), and renal anomalies (13.88%) (**Figure 4B**). The incidence of each phenotype in Chinese patients was further compared with those in Japanese and Korean patients. These proportions in hearing loss, preauricular pits, and branchial fistulae were not significantly different from that observed in patients from Japan and Korea ($p = 0.904$, $p = 1.000$, $p = 0.816$). The proportion of renal anomalies was statistically different ($p = 0.019$), and cases with renal phenotypes in China seem to be scarce, compared with Japan ($p = 0.009$) and Korea ($p = 0.038$), indicating that patients may present mainly with BOS in China.

# DISCUSSION

This study described nine affected individuals in five Han Chinese families with BOR or BOS. Of the nine patients, all had hearing loss (9/9) and preauricular pits (9/9), five had branchial fistulae or cysts (5/9), and only one had renal abnormalities (1/9), indicating that BOS is the major form. Genetic analysis confirmed five causative heterozygous variants in these unrelated families: c.1425delC (frameshift), c.889C > T (nonsense), c.1050+1G > T (splicing), and c.1140+1G > A (splicing) of *EYA1*, as well as c.316 > A (missense) of *SIX1*. To our knowledge, *EYA1*:c.1425delC and *SIX1*:c.316 > A have not been previously reported; c.316G > A is also the first reported *SIX1* variant in China. Our findings expand the mutational spectrum of BOR/BOS-

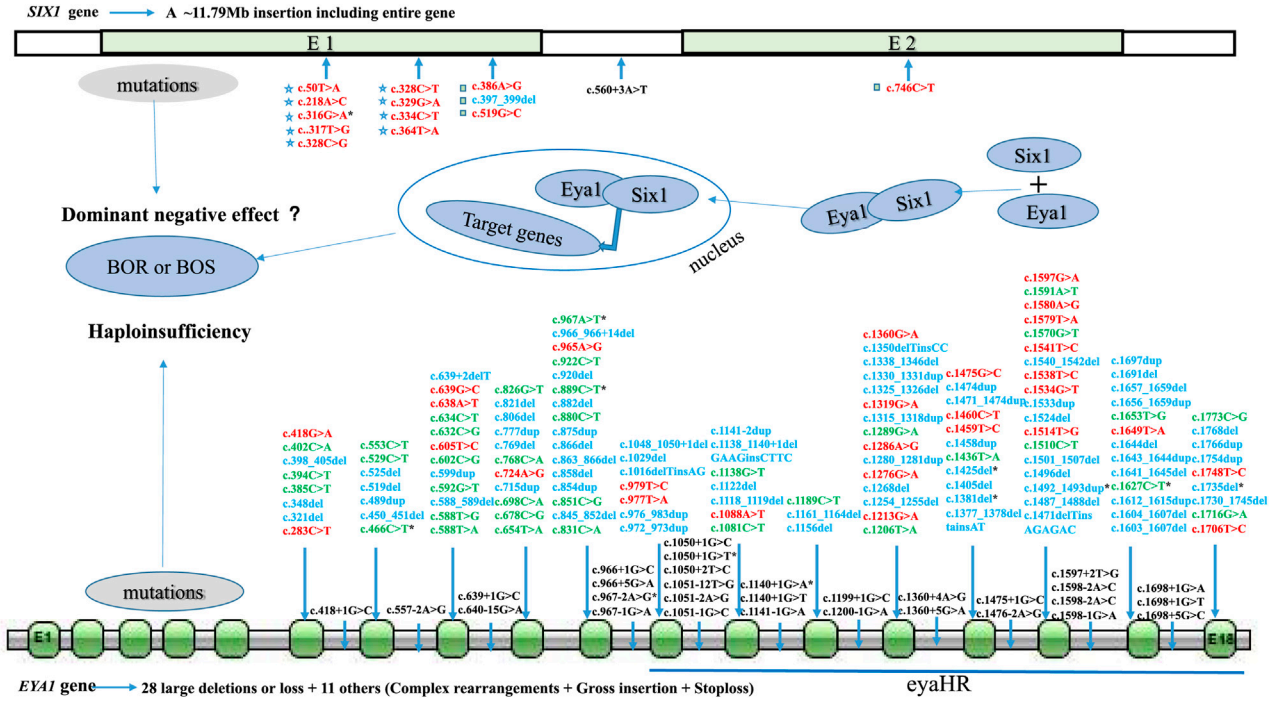



**FIGURE 3 |** Mutational spectrum of BOR/BOS-related genes (*EYA1* and *SIX1*). **E**: exon; red: missense; green: nonsense; blue: frameshift; black: splicing site; pentagram: SIX domain of Six1 protein; square: homeodomain of Six1 protein; asterisk: variants in Chinese patients.

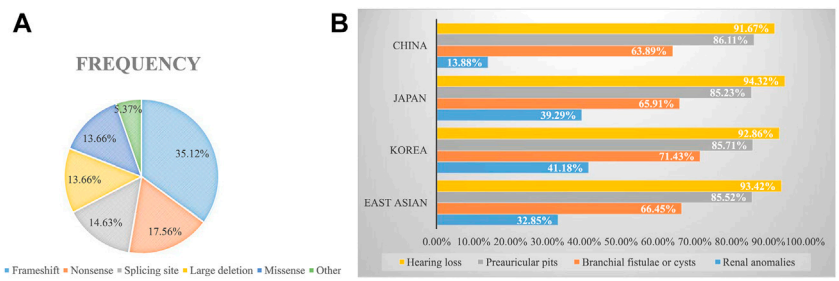



**FIGURE 4 | (A)** Frequencies of variant types in the *EYA1* gene. **(B)** Estimated prevalence of clinical phenotypes among BOR/BOS patients with *EYA1* variant.

related genes and demonstrate genetic and phenotypic variability in Chinese BOR/BOS patients.

*EYA1* and *SIX1* are key genes for mammalian organogenesis; mutations in these genes result in multiorgan malformation that can affect the branchial, ear, and renal systems. In 1997, Abdelhak and others reported novel variants in the *EYA1* gene in several families demonstrating typical manifestations of BOR/BOS (Abdelhak et al., 1997). Since then, more than 200 variants have been identified (**Figure 3**). The *EYA1* gene, located at 8q13.3, is a member of the *EYA* (*EYA1–4*) family; it contains 18 exons encoding a dual-function transcription factor. Similar to other EYA family members, Eya1 possesses a divergent N-terminal transactivation domain and a C-terminal EYA dephosphorylation region (eyaHR) encoded by 11–18 exons (Abdelhak et al., 1997; Li et al., 2003). Known variants in *EYA*1 mainly involve exons 6–18, especially exons 11–18, at least for now (**Figure 3**). These variants mainly cause premature truncation or aberrant splicing of the protein, resulting in the loss of eyaHR function. Screened *EYA1* variants in this study were predicted to interfere with this critical function. The protein Six1 encoded by the *SIX1* gene, a co-factor of Eya1, contains two evolutionarily conserved domains: a SIX domain that interacts with its co-factors and a homeodomain with DNA-binding ability (Ruf et al., 2004). Of the defects in the *SIX1* gene, Six1 functional alterations led to the failed formation of the Six1–Eya1 complex and the DNA–Six1–Eya1 complex (Ruf et al., 2004). Thus far, 14 *SIX1* variants are reportedly associated with BOR/BOS, eight of which

**TABLE 2 |** Genotypes and phenotypes of patients with BOR/BOS caused by *EYA1* variants in East Asian populations.

| Country | DNA change | Location | Protein change* | Phenotype (B, D, P, R) | Family history | References |
|---|---|---|---|---|---|---|
| China | c.466C > T | Exon 7 | p.(Gln156*) | B (2/3), D (2/3), P (3/3) | F | Wang et al. (2012) |
| — | c.889C > T | Exon 10 | p.(Arg297*) | B (1/2), D (2/2), P (1/2) | F | This study |
| — | c.967–2A > G | Intron 10 | Splicing site | B (3/3), D (3/3), P (3/3) | F | Chen et al. (2019) |
| — | c.967A > T | Exon 11 | p.(Arg323*) | B (3/6), D (6/6), P (5/6), R (1/6) | F | Wang et al. (2018) |
| — | c.1050+1G > T | Intron 11 | Splicing site | B (1/2), D (2/2), P (2/2) | F | This study |
| — | c.1140+1G > A | Intron 12 | Splicing site | B, D, P | S | This study |
| — | c.1381del | Exon 15 | p.(Arg461Glyfs*7) | B, P, R | S | Li et al. (2018) |
| — | c.1425del | Exon 15 | p.(Asp476Thrfs*4) | B (2/3), D (3/3), P (3/3), R (1/3) | F | This study |
| — | c.1492_1493dup | Exon 16 | p.(Ile499Phefs*33) | D (4/4), P (4/4) | F | Chen et al. (2019) |
| — | c.1627C>T | Exon 17 | p.(Gln543*) | B (4/4), D (4/4), P (4/4) | F | Han et al. (2021) |
| — | c.1735del | Exon 18 | p.(Asp579fs*60) | B (1/3), D (2/3), P (2/3) | F | Wang et al. (2012) |
| — | Entire deletion | Exon 1–18 | Loss protein | B (3/3), D (3/3), P (1/3), R (1/3) | F | Men et al. (2020) |
| — | Entire deletion | Exon 1–18 | Loss protein | B, D, P, R | S | Chen et al. (2014) |
| Japan | c.418+1G > C | Intron 6 | Splicing site | B, D,R | S | Unzaki et al. (2018) |
| — | c.588T > A | Exon 8 | p.(Tyr196*) | D, P, R | S | Okada et al. (2006) |
| — | c.588T > G | Exon 8 | p.(Tyr196*) | D (2/2), P (2/2) | F | Ideura et al. (2019) |
| — | c.632C > G | Exon 8 | p.(Ser211*) | B (2/4), D (3/4), P (4/4), R (1/4) | F | Uno et al. (2004) |
| — | c.634C > T | Exon 8 | p.(Gln212*) | B (2/2), D (2/2), P (2/2), R (1/2) | F | Unzaki et al. (2018) |
| — | c.678C > G | Exon 9 | p.(Tyr226*) | B (2/3), D (3/3), P (3/3), R (1/3) | F | Usami et al. (1999) |
| — | c.698C > A | Exon 9 | p.(Ser233*) | B, D, P | S | Unzaki et al. (2018) |
| — | c.724A > G | Exon 9 | p.(Ser242Gly) | D (2/2), P (1/2) | F | Yashima et al. (2003) |
| — | c.880C > T | Exon 10 | p.(Arg294*) | B, D, P, R | S | Unzaki et al. (2018) |
| — | c.889C > T | Exon 10 | p.(Arg297*) | B (11/11), D (11/11), P (11/11) | F | Fukuda et al. (2001) |
| — | c.922C > T | Exon 10 | p.(Arg308*) | B, D, R | S | Unzaki et al. (2018) |
| — | c.1050+2T > C | Intron 11 | Splicing site | D, P, R | S | Unzaki et al. (2018) |
| — | c.1051–2A > G | Intron 11 | Splicing site | B (2/3), D (3/3), P (3/3), R (1/3) | F | Okada et al. (2006) |
| — | c.1051G > T | Exon 12 | p.(Asp351Tyr) | B (1/3), D (3/3), P (3/3), R (1/3) | F | Okada et al. (2006) |
| — | c.1122delA | Exon 12 | p.(Leu374Phefs*6) | B (2/2), D (2/2), P (2/2), R (1/2) | F | Unzaki et al. (2018) |
| — | c.1140+1G > A | Intron 12 | Splicing site | B (3/3), D (3/3), P (3/3), R (2/3) | F | Unzaki et al. (2018) |
| — | c.1161_1164del | Exon 13 | p.(Ile387Metfs*12) | B (1/2), D (2/2), P (1/2), R (1/2) | F | Unzaki et al. (2018) |
| — | c.1206T > A | Exon 13 | p.(Tyr402*) | D, P, R | S | Okada et al. (2006) |
| — | c.1254_1255del | Exon 14 | p.(Cys419Phefs*32) | B, D, P, R (NA) | S | Ideura et al. (2019) |
| — | c.1276G > A | Exon 14 | p.(Gly426Ser) | B, D, R | S | Namba et al. (2001) |
| — | c.1286A > G | Exon 14 | p.(Asp429Gly) | B, D, P | S | Namba et al. (2001) |
| — | c.1289G > A | Exon 14 | p.(Trp430*) | B (2/2), D (2/2), P (2/2), R (2/2) | F | Unzaki et al. (2018) |
| — | c.1319G > A | Exon 14 | p.(Arg440Gln) | B (5/8), D (7/8), P (6/8), R (4/8) | F/S | Unzaki et al. (2018) |
| — | c.1501_1507del | Exon 16 | p.(Thr501Leufs*15) | B (3/4), D (4/4), P (2/4) | F | Shimasaki et al. (2004) |
| — | c.1598-1G > A | Intron 16 | Splicing site | B, D, P, R | S | Okada et al. (2006) |
| — | c.1643_1644dup | Exon 17 | p.(Val549Lysfs*7) | B, D, P, R | S | Unzaki et al. (2018) |
| — | c.1730_1745del | Exon 18 | p.(His577Profs*57) | D (3/3), P (3/3), R (2/3) | F | Unzaki et al. (2018) |
| — | c.1766dup | Exon 18 | p.(Glu590Glyfs*42) | B (5/5), D (4/5), P (5/5), R (1/5) | F | Matsunaga et al. (2007) |
| — | Partial deletion | Exon 4–7 | Truncated protein | D, P, R | S | Morisada et al. (2010) |
| — | Partial deletion | Exon 10–18 | Truncated protein | B (2/2), D (2/2), P (2/2) | F | Unzaki et al. (2018) |
| — | Partial deletion | Exon 2–3 | Truncated protein | D (3/3), P (1/3), R (1/3) | F | Unzaki et al. (2018) |
| — | Partial deletion | Exon 2–12 | Truncated protein | B, D, P | S | Unzaki et al. (2018) |
| — | Partial deletion | Exon 12 | Truncated protein | D, P, R | S | Unzaki et al. (2018) |
| — | Partial deletion | Exon 17 | Truncated protein | B (4/4), D (3/4), P (4/4), R (2/4) | F | Unzaki et al. (2018) |
| — | Partial deletion | Exon 17 | Truncated protein | D (2/2), R (1/2) | F | Unzaki et al. (2018) |
| — | Entire deletion | Exon 1–18 | Loss protein | B (1/3), D (3/3), P (3/3), R (NA) | F | Ideura et al. (2019) |
| Korea | c.321del | Exon 6 | p.(Ala107fs*133) | B (2/2), D (2/2), P (2/2), R (1/2) | F | Lee et al. (2009) |
| — | c.418G > A | Exon 6 | p.(Gly140Ser) | B, D, P | S | Kim et al. (2014) |
| — | c.529C > T | Exon 7 | p.(Gln177*) | B (10/10), D (9/10), P (10/10), R (4/5) | F | Lee et al. (2007) |
| — | c.699+5G > A | Intron 10 | Splicing site | D, R (NA) | S | Song et al. (2013) |
| — | c.965A > G | Exon 10 | p.(Glu322Gly) | D (4/4), P (4/4), R (NA) | F | Song et al. (2013) |
| — | c.967–2A > G | Intron 10 | Splicing site | B (3/4), D (3/4), P (4/4) | F | Kwon et al. (2009) |
| — | c.1140+1G > A | Intron 12 | Splicing site | D, P, R (NA) | S | Song et al. (2013) |
| — | c.1474dup | Exon 15 | p.(Arg492Profs*40) | B (3/3), D (3/3), R (2/3) | F | Kim et al. (2005) |
| — | c.1598–2A > C | Intron 16 | Splicing site | B, D, P | S | Song et al. (2013) |
| — | Entire deletion | Exon 1–18 | Loss protein | D, P | S | Song et al. (2013) |

EYA1, *canonical transcript NM_000,503.6;* SIX1, *canonical transcript NM_005,982.4; D, hearing loss; P, preauricular pits; B, branchial fistulae or cysts; R, renal anomalies; NA, not available or not done; F, family history; S, scatter.*

are in the SIX domain (**Figure 3**). We also identified a novel missense variant (c.316G > A) in the *SIX1* gene; we speculate that this variant interferes with Six1 binding to Eya1, which has been previously reported for variants in the SIX domain (Ruf et al., 2004; Kochhar et al., 2008). Loss of function of either gene in a mouse model led to hearing loss and dysmorphic or missing kidneys, along with developmental abnormalities in other organs (Xu et al., 1999; Li et al., 2003). Our patients harboring *EYA1* or *SIX1* alterations also had hearing loss, ear deformities, or kidney problems.

BOR/BOS patients demonstrate intrafamilial and interfamilial phenotypic variability in the clinical setting, suggesting a lack of genotype–phenotype correlation. In the present study, proband 5-II-2 harboring c.1425delC exhibited all major symptoms of BOR, while patients 5-II-1 and 5-I-1 only presented with the triad of BOS, demonstrating coexistence of BOR and BOS in this family. The presence of interfamilial phenotypic variability has also been confirmed. A patient from China with the same c.1140+1G > A variant demonstrated no renal involvement; patients from Korea reportedly lack branchial fistulae; and Japanese patients have four main symptoms of BOR (**Table 2**). Notably, most patients carrying the same or different variants at the same position, such as c.1050+1G > T (Orten et al., 2008), c.1050+1G > C (Henriksen et al., 2004), or c.1050+2T > C (Unzaki et al., 2018), showed renal anomalies; such anomalies were not observed in our patients with the c.1050+1G > T variant. This variability is also present among patients with *SIX1* variants. Patients with c.328C > T variant exhibited distinct phenotypes in several unrelated families (Kochhar et al., 2008). In addition, c.316G > A and c.317T > G led to the substitution of amino acids at the same position to methionine and glycine, respectively (Kochhar et al., 2008). Two individuals exhibited similar symptoms of BOS. Specifically, the same or a similar variant can cause variable clinical phenotypes in BOR/BOS, while different variants can cause similar clinical phenotypes.

The mechanism by which mutations produce phenotypic variability is undefined in BOR/BOS. Known *EYA1* variants mainly comprise loss-of-function mutations; such variants typically imply haploinsufficiency through reduced gene dosage and expression (**Figure 3**) (Zhang et al., 2004). These results suggest that the inconsistency of phenotypes is partially influenced by the dosage effect. The Eya1 protein activates target genes controlling the development of the branchial arch, ear, or kidney only when a specific threshold is exceeded (Wang et al., 2018). Otherwise, environmental factors and genetic modifiers might also modify the phenotypes of BOR/BOS. Notably, we found that reported *SIX1* variants mainly comprise missense (12/15). In the future, further exploration of the role of the dominant-negative effect (Shah et al., 2020) is also warranted.

BOS appears to be the main manifestation in China. In 1995, a study indicated that the estimated prevalence of renal abnormality was 67% in 21 patients with BOR in the United States; renal agenesis occurred most often, followed by hypoplasia, renal dysplasia, ureteral-pelvic junction obstruction, calyceal cyst/diverticulum, caliectasis, and hydronephrosis (Chen et al., 1995). In 2004, Chang et al. reported the renal abnormality frequency of 38% based on analyses of 40 families, but no detailed renal phenotypes were described (Chang et al., 2004). In France, Kurg et al. also identified 53% of the prevalence of renal anomalies in BOR patients harboring *EYA1* variants; there is a wide range of abnormalities, namely, renal hypoplasia, multicystic kidney dysplasia, agenesis, abnormal pyelo-ureteral junction, and kidney malrotation, to name the most common (Krug et al., 2011). In our study, only one patient presented with left-sided renal hypoplasia, hydronephrosis, and pyelo-ureteral separation (1/9). We then calculated the frequency of renal malformations in East Asian populations through a literature review (**Figure 4B**). The frequencies were 14% in China, 39% in Japan, and 41% in Korea. Comparatively, the morbidity of renal anomalies appears to be lower and statistically different, suggesting that patients may primarily present with BOS in China. The elucidation of population differences in renal anomalies requires further exploration.

The clinical management of BOR/BOS requires individual management and multidisciplinary collaboration because of the multi-systemic symptoms. Deafness, the most prevalent manifestation of BOR/BOS, requires close attention and early treatment. Patients with BOR/BOS exhibit multiple structural malformations of the middle ear, especially ossicular malformations. Surgeons perform ossicular reconstruction to repair mixed or conductive hearing loss (Cremers et al., 1981), but the results are not always satisfactory. Our therapeutic findings for probands 1-II-3 and 5-I-1 were consistent with previous results. Thus, we presume that middle ear surgery is not optimal for patients with BOR/BOS. Hearing aids are beneficial to most patients, but this treatment should be individually tailored. Proband 1-II-3 received soft-banded bone-conduction hearing aids, while proband 2-II-1 wore air-conduction hearing aids after fitting. Their hearing abilities were substantially improved, which facilitated the fulfillment of learning and communication needs. Miyagawa et al. performed bone-anchored hearing aid implantation on a patient with BOS, which led to an improved hearing threshold (Miyagawa et al., 2015). Therefore, bone-conduction hearing device implantation may be considered when cranial bone thickness is at least 4 mm. CI may be a good option when hearing aids do not substantially improve hearing. Kameswaran et al. performed CI on a BOR patient with multiple inner ear malformations (Kameswaran et al., 2007); Bajaj et al. analyzed the surgical effect of CI in syndromic children, including children with BOR/BOS (Bajaj et al., 2012); their results indicated that CI was feasible in patients with BOR/BOS. In the present study, proband 5-II-3 underwent CI in the right ear after a comprehensive evaluation of the preoperative findings. CI can be successful through careful intraoperative operation, and the postoperative hearing ability is satisfactory with rehabilitation.

In conclusion, the symptoms of BOR or BOS in this study were attributed to *EYA1* or *SIX1* alterations. BOR/BOS exhibited some genetic and phenotypic variability. The outcomes of auditory rehabilitation reiterated that middle ear surgeries are generally unsatisfactory in patients with BOR/BOS; CI may be a feasible option when patients cannot benefit from hearing aids. Genetic testing

contributes to the diagnosis and future genetic consultation; it has vital roles in therapy and intervention.

## DATA AVAILABILITY STATEMENT

All data supporting the findings of this study are available on request from the corresponding author. The pathogenic variants have been submitted to ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/) with the accession number (VCV001202644, VCV001202645, VCV000854287, VCV000429912, VCV001202646).

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Medical Ethics Committee of The Affiliated Hospital of Zhengzhou University (Approval No. 2018008). Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

## AUTHOR CONTRIBUTIONS

Study design: WL and WT. Patient phenotypic analysis and genetic counseling: HF, BC, SS, and RZ. Targeted sequencing, Sanger sequencing, and variant interpretation: BZ and HX. Writing and reviewing the original draft of the manuscript: HF, HX, and WL. All authors have read and approved the final manuscript.

## FUNDING

The study is funded by the Collaborative Innovation Project of Zhengzhou (Zhengzhou University) (Grant no. 18XTZX12004) to WT and the Key Scientific Research Projects of Colleges and Universities in Henan Province (Grant no. 22A320026) to WL.

## ACKNOWLEDGMENTS

We sincerely thank all participants in this study for their collaboration and cooperation! The National Supercomputing Center in Zhengzhou supported the bioinformatics analysis.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.765433/full#supplementary-material

**Supplementary Data Sheet S1** | **(A)** The mapping after CI surgery (Map). **(B)** Neural response telemetry (NRT). **(C)** Behavioral audiometry.

## REFERENCES

Abdelhak, S., Kalatzis, V., Heilig, R., Compain, S., Samson, D., Vincent, C., et al. (1997). A Human Homologue of the Drosophila Eyes Absent Gene Underlies Branchio-Oto-Renal (BOR) Syndrome and Identifies a Novel Gene Family. *Nat. Genet.* 15 (2), 157–164. doi:10.1038/ng0297-157

Ando, Z.-I., Sato, S., Ikeda, K., and Kawakami, K. (2005). Slc12a2 Is a Direct Target of Two Closely Related Homeobox Proteins, Six1 and Six4. *Febs j* 272 (12), 3026–3041. doi:10.1111/j.1742-4658.2005.04716.x

Bajaj, Y., Gibbins, N., Fawkes, K., Hartley, B., Jephson, C., Jonas, N., et al. (2012). Surgical Aspects of Cochlear Implantation in Syndromic Children. *Cochlear Implants Int.* 13 (3), 163–167. doi:10.1179/1754762811y.0000000020

Chang, E. H., Menezes, M., Meyer, N. C., Cucci, R. A., Vervoort, V. S., Schwartz, C. E., et al. (2004). Branchio-Oto-Renal Syndrome: The Mutation Spectrum inEYA1and its Phenotypic Consequences. *Hum. Mutat.* 23 (6), 582–589. doi:10.1002/humu.20048

Chen, A., Francis, M., Ni, L., Cremers, C. W. R. J., Kimberling, W. J., Sato, Y., et al. (1995). Phenotypic Manifestations of Branchiootorenal Syndrome. *Am. J. Med. Genet.* 58 (4), 365–370. doi:10.1002/ajmg.1320580413

Chen, P., Liu, H., Lin, Y., Xu, J., Zhu, W., Wu, H., et al. (2019). EYA1 Mutations Leads to Branchio-Oto Syndrome in Two Chinese Han Deaf Families. *Int. J. Pediatr. Otorhinolaryngol.* 123, 141–145. doi:10.1016/j.ijporl.2019.05.006

Chen, X., Wang, J., Mitchell, E., Guo, J., Wang, L., Zhang, Y., et al. (2014). Recurrent 8q13.2-13.3 Microdeletions Associated with Branchio-Oto-Renal Syndrome Are Mediated by Human Endogenous Retroviral (HERV) Sequence Blocks. *BMC Med. Genet.* 15, 90. doi:10.1186/s12881-014-0090-9

Cremers, C. W. R. J., Thijssen, H. O. M., Fischer, A. J. E. M., and Marres, E. H. M. A. (1981). Otological Aspects of the Earpit-Deafness Syndrome. *ORL J. Otorhinolaryngol. Relat. Spec.* 43 (4), 223–239. doi:10.1159/000275541

DePristo, M. A., Banks, E., Poplin, R., Garimella, K. V., Maguire, J. R., Hartl, C., et al. (2011). A Framework for Variation Discovery and Genotyping Using Next-Generation DNA Sequencing Data. *Nat. Genet.* 43 (5), 491–498. doi:10.1038/ng.806

Fraser, F. C., Ling, D., Clogg, D., Nogrady, B., and Gorlin, R. J. (1978). Genetic Aspects of the BOR Syndrome-Branchial Fistulas, Ear Pits, Hearing Loss, and Renal Anomalies. *Am. J. Med. Genet.* 2 (3), 241–252. doi:10.1002/ajmg.1320020305

Fraser, F. C., Sproule, J. R., Halal, F., and Optiz, J. M. (1980). Frequency of the Branchio-Oto-Renal (BOR) Syndrome in Children with Profound Hearing Loss. *Am. J. Med. Genet.* 7 (3), 341–349. doi:10.1002/ajmg.1320070316

Fukuda, S., Kuroda, T., Chida, E., Shimizu, R., Usami, S.-I., Koda, E., et al. (2001). A Family Affected by Branchio-Oto Syndrome with EYA1 Mutations. *Auris Nasus Larynx* 28 (Suppl. l), S7–S11. doi:10.1016/s0385-8146(01)00082-7

Ghosh, R., Harrison, S. M., Rehm, H. L., Plon, S. E., and Biesecker, L. G. (2018). Updated Recommendation for the Benign Stand-Alone ACMG/AMP Criterion. *Hum. Mutat.* 39 (11), 1525–1530. doi:10.1002/humu.23642

Han, R., Xia, Y., Liu, Z., Wu, S., Ye, E., Duan, L., et al. (2021). A Mutation of EYA1 Gene in a Chinese Han Family with Branchio-Oto Syndrome. *Medicine (Baltimore)* 100 (25), e24691. doi:10.1097/md.0000000000024691

Henriksen, A. M., Tümer, Z., Tommerup, N., Tranebjærg, L., and Larsen, L. A. (2004). Identification of a NovelEYA1Splice-Site Mutation in a Danish Branchio-Oto-Renal Syndrome Family. *Genet. Test.* 8 (4), 404–406. doi:10.1089/gte.2004.8.404

Hoskins, B. E., Cramer, C. H., Silvius, D., Zou, D., Raymond, R. M., Orten, D. J., et al. (2007). Transcription Factor SIX5 Is Mutated in Patients with Branchio-Oto-Renal Syndrome. *Am. J. Hum. Genet.* 80 (4), 800–804. doi:10.1086/513322

Ideura, M., Nishio, S.-y., Moteki, H., Takumi, Y., Miyagawa, M., Sato, T., et al. (2019). Comprehensive Analysis of Syndromic Hearing Loss Patients in Japan. *Sci. Rep.* 9 (1), 11976. doi:10.1038/s41598-019-47141-4

Jaganathan, K., Kyriazopoulou Panagiotopoulou, S., McRae, J. F., Darbandi, S. F., Knowles, D., Li, Y. I., et al. (2019). Predicting Splicing from Primary Sequence with Deep Learning. *Cell* 176 (3), 535–548. e524. doi:10.1016/j.cell.2018.12.015

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 (22), 13534–13544. doi:10.1093/nar/gku1206

Kameswaran, M., Kumar, R. S. A., Murali, S., Raghunandhan, S., and Karthikeyan, K. (2007). Cochlear Implantation in Branchio-Oto-Renal Syndrome - A Surgical Challenge. *Indian J. Otolaryngol. Head Neck S* 59 (3), 280–283. doi:10.1007/s12070-007-0081-7

Kim, H. R., Song, M. H., Kim, M.-A., Kim, Y.-R., Lee, K.-Y., Sonn, J. K., et al. (2014). Identification of a Novel Nonsynonymous Mutation of EYA1 Disrupting Splice Site in a Korean Patient with BOR Syndrome. *Mol. Biol. Rep.* 41 (7), 4321–4327. doi:10.1007/s11033-014-3303-6

Kim, S. H., Shin, J.-H., Yeo, C.-K., Chang, S. H., Park, S.-Y., Cho, E. H., et al. (2005). Identification of a Novel Mutation in the EYA1 Gene in a Korean Family with Branchio-Oto-Renal (BOR) Syndrome. *Int. J. Pediatr. Otorhinolaryngol.* 69 (8), 1123–1128. doi:10.1016/j.ijporl.2005.03.003

Kochhar, A., Orten, D. J., Sorensen, J. L., Fischer, S. M., Cremers, C. W. R. J., Kimberling, W. J., et al. (2008). SIX1mutation Screening in 247 Branchio-Oto-Renal Syndrome Families: A Recurrent Missense Mutation Associated with BOR. *Hum. Mutat.* 29 (4), 565. doi:10.1002/humu.20714

Krug, P., Morinière, V., Marlin, S., Koubi, V., Gabriel, H. D., Colin, E., et al. (2011). Mutation Screening of the EYA1, SIX1, and SIX5 Genes in a Large Cohort of Patients Harboring Branchio-Oto-Renal Syndrome Calls into Question the Pathogenic Role of SIX5 Mutations. *Hum. Mutat.* 32 (2), 183–190. doi:10.1002/humu.21402

Kwon, M.-J., Boo, S. H., Kwon, M.-J., Boo, S. H., Kim, H.-J., Cho, Y.-S., et al. (2009). A Novel Splice Site Mutation in theEYA1gene in a Korean Family with Branchio-Oto (BO) Syndrome. *Acta Oto-Laryngologica* 129 (6), 688–693. doi:10.1080/00016480802342432

Lee, J. D., Kim, S. C., Koh, Y. W., Lee, H. J., Choi, S. Y., and Kim, U. K. (2009). A Novel Frameshift Mutation in the EYA1 Gene in a Korean Family with Branchio-Oto-Renal Syndrome. *Ann. Clin. Lab. Sci.* 39 (3), 303–306.

Lee, K. Y., Kim, S., Kim, U. K., Ki, C.-S., and Lee, S. H. (2007). Novel EYA1 Mutation in a Korean Branchio-Oto-Renal Syndrome Family. *Int. J. Pediatr. Otorhinolaryngol.* 71 (1), 169–174. doi:10.1016/j.ijporl.2006.08.023

Li, G., Shen, Q., Sun, L., Liu, H., An, Y., and Xu, H. (2018). A De Novo and Novel Mutation in the EYA1 Gene in a Chinese Child with Branchio-Oto-Renal Syndrome. *Intractable Rare Dis. Res.* 7 (1), 42–45. doi:10.5582/irdr.2017.01075

Li, X., Ohgi, K. A., Zhang, J., Krones, A., Bush, K. T., Glass, C. K., et al. (2003). Eya Protein Phosphatase Activity Regulates Six1-Dach-Eya Transcriptional Effects in Mammalian Organogenesis. *Nature* 426 (6964), 247–254. doi:10.1038/nature02083

Matsunaga, T., Okada, M., Usami, S.-I., and Okuyama, T. (2007). Phenotypic Consequences in a Japanese Family Having Branchio-Oto-Renal Syndrome with a Novel Frameshift Mutation in the geneEYA1. *Acta Oto-Laryngologica* 127 (1), 98–104. doi:10.1080/00016480500527185

Melnick, M., Bixler, D., Silk, K., Yune, H., and Nance, W. E. (1975). Autosomal Dominant Branchiootorenal Dysplasia. *Birth Defects Orig. Artic Ser.* 11 (5), 121–128.

Men, M., Li, W., Chen, H., Wu, J., Feng, Y., Guo, H., et al. (2020). Identification of a Novel CNV at 8q13 in a Family with Branchio-Oto-Renal Syndrome and Epilepsy. *The Laryngoscope* 130 (2), 526–532. doi:10.1002/lary.27941

Miyagawa, M., Nishio, S.-y., Hattori, M., Takumi, Y., and Usami, S.-i. (2015). Germinal Mosaicism in a Family with BO Syndrome. *Ann. Otol. Rhinol. Laryngol.* 124 (Suppl. 1), 118s–122s. doi:10.1177/0003489415575062

Morisada, N., Rendtorff, N. D., Nozu, K., Morishita, T., Miyakawa, T., Matsumoto, T., et al. (2010). Branchio-Oto-Renal Syndrome Caused by Partial EYA1 Deletion Due to LINE-1 Insertion. *Pediatr. Nephrol.* 25 (7), 1343–1348. doi:10.1007/s00467-010-1445-x

Namba, A., Abe, S., Shinkawa, H., Kimberling, W. J., and Usami, S. (2001). Genetic Features of Hearing Loss Associated with Ear Anomalies: PDS and EYA1 Mutation Analysis. *J. Hum. Genet.* 46 (9), 518–521. doi:10.1007/s100380170033

Okada, M., Fujimaru, R., Morimoto, N., Satomura, K., Kaku, Y., Tsuzuki, K., et al. (2006). EYA1 and SIX1 Gene Mutations in Japanese Patients with Branchio-Oto-Renal (BOR) Syndrome and Related Conditions. *Pediatr. Nephrol.* 21 (4), 475–481. doi:10.1007/s00467-006-0041-6

Orten, D. J., Fischer, S. M., Sorensen, J. L., Radhakrishna, U., Cremers, C. W. R. J., Marres, H. A. M., et al. (2008). Branchio-oto-renal Syndrome (BOR): Novel Mutations in theEYA1gene, and a Review of the Mutational Genetics of BOR. *Hum. Mutat.* 29 (4), 537–544. doi:10.1002/humu.20691

Oza, A. M., DiStefano, M. T., Hemphill, S. E., Cushman, B. J., Grant, A. R., Siegert, R. K., et al. (2018). Expert Specification of the ACMG/AMP Variant Interpretation Guidelines for Genetic Hearing Loss. *Hum. Mutat.* 39 (11), 1593–1613. doi:10.1002/humu.23630

Pan, Z., Xu, H., Tian, Y., Liu, D., Liu, H., Li, R., et al. (2020). Perrault Syndrome: Clinical Report and Retrospective Analysis. *Mol. Genet. Genomic Med.* 8 (10), e1445. doi:10.1002/mgg3.1445

Propst, E. J., Blaser, S., Gordon, K. A., Harrison, R. V., and Papsin, B. C. (2005). Temporal Bone Findings on Computed Tomography Imaging in Branchio-Oto-Renal Syndrome. *The Laryngoscope* 115 (10), 1855–1862. doi:10.1097/01.mlg.0000177032.98386.20

Rickard, S., Boxer, M., Trompeter, R., and Bitner-Glindzicz, M. (2000). Importance of Clinical Evaluation and Molecular Testing in the Branchio-Oto-Renal (BOR) Syndrome and Overlapping Phenotypes. *J. Med. Genet.* 37 (8), 623–627. doi:10.1136/jmg.37.8.623

Ruf, R. G., Xu, P.-X., Silvius, D., Otto, E. A., Beekmann, F., Muerb, U. T., et al. (2004). SIX1 Mutations Cause Branchio-Oto-Renal Syndrome by Disruption of EYA1-SIX1-DNA Complexes. *Proc. Natl. Acad. Sci.* 101 (21), 8090–8095. doi:10.1073/pnas.0308475101

Shah, A. M., Krohn, P., Baxi, A. B., Tavares, A. L. P., Sullivan, C. H., Chillakuru, Y. R., et al. (2020). Six1 Proteins with Human Branchio-Oto-Renal Mutations Differentially Affect Cranial Gene Expression and Otic Development. *Dis. Model. Mech.* 13 (3), dmm043489. doi:10.1242/dmm.043489

Shimasaki, N., Watanabe, K., Hara, M., and Kosaki, K. (2004). EYA1 Mutation in a Newborn Female Presenting with Cardiofacial Syndrome. *Pediatr. Cardiol.* 25 (4), 411–413. doi:10.1007/s00246-003-0271-3

Song, M. H., Kwon, T.-J., Kim, H. R., Jeon, J. H., Baek, J.-I., Lee, W.-S., et al. (2013). Mutational Analysis of EYA1, SIX1 and SIX5 Genes and Strategies for Management of Hearing Loss in Patients with BOR/BO Syndrome. *PLoS One* 8 (6), e67236. doi:10.1371/journal.pone.0067236

Tarasov, A., Vilella, A. J., Cuppen, E., Nijman, I. J., and Prins, P. (2015). Sambamba: Fast Processing of NGS Alignment Formats. *Bioinformatics* 31 (12), 2032–2034. doi:10.1093/bioinformatics/btv098

Uno, T., Sawada, M., Kurotaki, T., and Shinomiya, N. (2004). EYA1 Gene Nonsense Mutation in a Japanese Family with Branchio-Oto-Renal Syndrome. *Pediatr. Int.* 46 (5), 615–617. doi:10.1111/j.1442-200x.2004.01935.x

Unzaki, A., Morisada, N., Nozu, K., Ye, M. J., Ito, S., Matsunaga, T., et al. (2018). Clinically Diverse Phenotypes and Genotypes of Patients with Branchio-Oto-Renal Syndrome. *J. Hum. Genet.* 63 (5), 647–656. doi:10.1038/s10038-018-0429-8

Usami, S., Abe, S., Shinkawa, H., Deffenbacher, K., Kumar, S., and Kimberling, W. J. (1999). EYA1 Nonsense Mutation in a Japanese Branchio-Oto-Renal Syndrome Family. *J. Hum. Genet.* 44 (4), 261–265. doi:10.1007/s100380050156

Wang, S.-H., Wu, C.-C., Lu, Y.-C., Lin, Y.-H., Su, Y.-N., Hwu, W.-L., et al. (2012). Mutation Screening of the EYA1, SIX1, and SIX5 Genes in an East Asian Cohort with Branchio-Oto-Renal Syndrome. *The Laryngoscope* 122 (5), 1130–1136. doi:10.1002/lary.23217

Wang, Y.-G., Sun, S.-P., Qiu, Y.-L., Xing, Q.-H., and Lu, W. (2018). A Novel Mutation in EYA1 in a Chinese Family with Branchio-Oto-Renal Syndrome. *BMC Med. Genet.* 19 (1), 139. doi:10.1186/s12881-018-0653-2

Xu, P.-X., Adams, J., Peters, H., Brown, M. C., Heaney, S., and Maas, R. (1999). Eya1-deficient Mice Lack Ears and Kidneys and Show Abnormal Apoptosis of Organ Primordia. *Nat. Genet.* 23 (1), 113–117. doi:10.1038/12722

Yashima, T., Noguchi, Y., Ishikawa, K., Mizusawa, H., and Kitamura, K. (2003). Mutation of theEYA1Gene in Patients with Branchio-Oto Syndrome. *Acta Oto-Laryngologica* 123 (2), 279–282. doi:10.1080/0036554021000028103

Zhang, Y., Knosp, B. M., Maconochie, M., Friedman, R. A., and Smith, R. J. H. (2004). A Comparative Study of Eya1 and Eya4 Protein Function and its Implication in Branchio-Oto-Renal Syndrome and DFNA10. *J. Assoc. Res. Otolaryngol.* 5 (3), 295–304. doi:10.1007/s10162-004-4044-3

**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.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

*Copyright © 2021 Feng, Xu, Chen, Sun, Zhai, Zeng, Tang and Lu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*

frontiers in Genetics

BRIEF RESEARCH REPORT
published: 17 November 2021
doi: 10.3389/fgene.2021.776831

# Missense Mutations of Codon 116 in the *SOD1* Gene Cause Rapid Progressive Familial ALS and Predict Short Viability With PMA Phenotype

*Xinmei Wen[1], Wenjia Zhu[1], Nan L. Xia[2], Qianwen Li[3,4], Li Di[1], Shu Zhang[1], Hai Chen[1], Yan Lu[1], Min Wang[1], Min Xu[1], Suobin Wang[1], Xin-Ming Shen[5], Jie Lu[3,4]* and Yuwei Da[1]**

[1]Department of Neurology, Xuanwu Hospital, Capital Medical University, Beijing, China, [2]Department of Rheumatology and Immunology, The Affiliated Drum Tower Hospital of Nanjing University Medical School, Nanjing, China, [3]Department of Radiology and Nuclear Medicine, Xuanwu Hospital, Capital Medical University, Beijing, China, [4]Beijing Key Laboratory of Magnetic Resonance Imaging and Brain Informatics, Beijing, China, [5]Department of Neurology and Neuromuscular Research Laboratory, Mayo Clinic, Rochester, MN, United States

OPEN ACCESS

**Edited by:**
*Alfredo Brusco,*
*University of Turin, Italy*

**Reviewed by:**
*Velina Guergueltcheva,*
*Sofiamed University Hospital, Bulgaria*
*Adriano Chio,*
*University of Turin, Italy*

***Correspondence:**
*Yuwei Da*
*dayuwei100@hotmail.com*
*Jie Lu*
*imaginglu@hotmail.com*

**Specialty section:**
*This article was submitted to Genetics of Common and Rare Diseases, a section of the journal Frontiers in Genetics*

**Received:** *14 September 2021*
**Accepted:** *14 October 2021*
**Published:** *17 November 2021*

**Citation:**
*Wen X, Zhu W, Xia NL, Li Q, Di L, Zhang S, Chen H, Lu Y, Wang M, Xu M, Wang S, Shen X-M, Lu J and Da Y (2021) Missense Mutations of Codon 116 in the SOD1 Gene Cause Rapid Progressive Familial ALS and Predict Short Viability With PMA Phenotype. Front. Genet. 12:776831. doi: 10.3389/fgene.2021.776831*

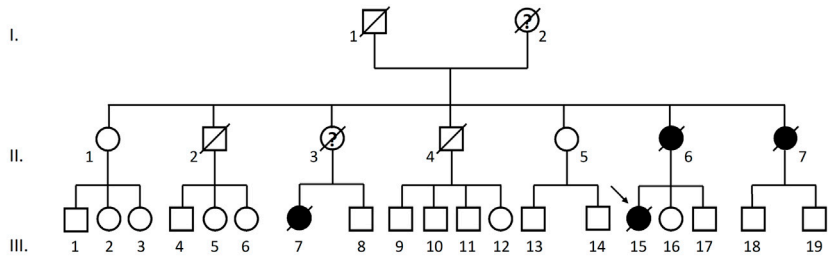
Amyotrophic lateral sclerosis (ALS) is the most common form of motor neuron disease, characterized by a great variety of both clinical presentations and genetic causes. Previous studies had identified two different missense mutations in *SOD1* (p.R116C and p.R116G) causing familial ALS. In this study, we report a novel heterozygous missense mutation in the *SOD1* gene (p.R116S) in a family with inherited ALS manifested as fast-deteriorating pure lower motor neuron symptoms. The patient displayed similar clinical picture and prognostic value to previous reported cases with different R116 substitution mutations. Modeling of all R116 substitutions in the resolved SOD1 protein structure revealed a shared mechanism with destroyed hydrogen bonds between R116 and other two residues, which might lead to protein unfolding and oligomer formation, ultimately conferring neurotoxicity.

**Keywords: amyotrophic lateral sclerosis, SOD1, lower motor neuron, progressive muscular atrophy, rapid progression**

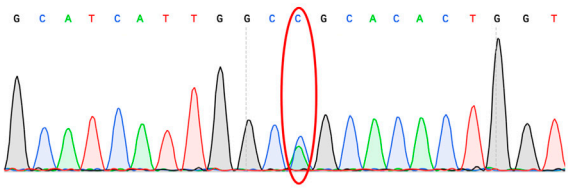
## INTRODUCTION

Amyotrophic lateral sclerosis (ALS) is a relentless neurodegenerative disorder and the most common adult-onset motor neuron disease (MND), which causes muscle weakness and atrophy. It is progressively fatal, and the majority of patients die within 3 years after onset of symptom (Turner et al., 2013). It is considered highly heterogeneous in nature, which displays a great variety of both clinical presentations and underlying pathogenic mechanisms. About 5%–10% of ALS cases are familial (fALS). The genetic spectrums of ALS vary among different populations. In European and American populations, the most common mutations are *C9orf72* intronic repeat expansions, whereas in Chinese and other Asian populations, *SOD1* mutations are most frequent (Zou et al., 2017; Liu et al., 2018). So far, more than 180 distinct *SOD1* mutations have been identified as genetic causes for ALS (Mathis et al., 2019) (https://alsod.ac.uk/). Mutant SOD1 proteins cause neurodegeneration mainly through a gain of toxicity pathway. Other possible mechanisms also include increased protein instability, protein aggregation, and probabilities of fibrillization (Wright et al., 2019).

**A**




**B**




**C**


**D**

| | 102 | | | | | | | | | | | | | | | 116 | | | | | | | | | | | | | | | 131 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Homo sapiens** | D | S | V | I | S | L | S | G | D | H | C | I | I | G | **R** | T | L | V | V | H | E | K | A | D | D | L | G | K | G | G |
| **Macaca mulatta** | D | S | V | I | S | L | S | G | D | H | S | I | I | G | **R** | T | L | V | V | H | E | K | A | D | D | L | G | K | G | G |
| **Rattus norvegicus** | D | R | V | I | S | L | S | G | E | H | S | I | I | G | **R** | T | M | V | V | H | E | K | Q | D | D | L | G | K | G | G |
| **Mus musculus** | D | R | V | I | S | L | S | G | E | H | S | I | I | G | **R** | T | M | V | V | H | E | K | Q | D | D | L | G | K | G | G |
| **Danio rerio** | D | A | M | L | T | L | S | G | Q | H | S | I | I | G | **R** | T | M | V | I | H | E | K | E | D | D | L | G | K | G | G |
| **Saccharomyces cerevisiae** | D | S | L | I | K | L | I | G | P | T | S | V | V | G | **R** | S | V | V | I | H | A | G | Q | D | D | L | G | K | G | D |
| **Gallus gallus** | D | S | V | I | S | L | T | G | P | H | C | I | I | G | **R** | T | M | V | V | H | A | K | S | D | D | L | G | R | G | G |
| **Drosophila melanogaster** | D | S | K | I | T | L | F | G | A | D | S | I | I | G | **R** | T | V | V | V | H | A | D | A | D | D | L | G | Q | G | G |
| **Trichinella pseudospiralis** | D | D | Q | I | K | L | T | G | E | H | S | I | I | G | **R** | T | M | V | V | H | I | Q | E | D | D | L | G | K | G | G |
| **Schizosaccharomyces pombemi** | D | S | V | I | S | L | F | G | A | N | S | I | I | G | **R** | T | I | V | I | H | A | G | E | D | D | L | G | K | G | T |

**E**

| | SIFT | SIFT Tolerability | Provean | PolyPhen-2 Score | PolyPhen-2 | Mutation Taster | Solubility | ΔΔG kcal/mol | Grantham Score |
|---|---|---|---|---|---|---|---|---|---|
| **WT** | - | - | - | - | - | - | 1.702837 | - | - |
| **R116G** | Deleterious | 0 | -6.919 | 1.0000 | Probably Damaging | Disease Causing | 1.577059 | 1.99 | 125 |
| **R116C** | Deleterious | 0 | -7.908 | 1.0000 | Probably Damaging | Disease Causing | 1.3748 | 0.58 | 180 |
| **R116S** | Deleterious | 0 | -5.931 | 1.0000 | Probably Damaging | Disease Causing | 1.562741 | 1.62 | 110 |

**FIGURE 1 | (A)** Pedigree of the family. Square = male; circle = female; diagonal black line = deceased individual; black filled symbol = clinically proven affected individual; empty symbol = clinically healthy relative; question mark = suspected; arrow = proband. **(B)** DTI reconstruction of the bilateral cortical spinal tracts of the proband. RCST, right cortical spinal tract. LCST, left cortical spinal tract. **(C)** Confirmation by Sanger sequencing showing point mutation (red oval). **(D)** Sequence alignment analysis of SOD1 in different species. Amino acid at position 116 is highlighted in red. **(E)** Functional prediction of mutations at 116th amino acid in SOD1.

The genotype–phenotype correlation has been described for a few *SOD1* mutations for disease durations but not for clinical features. p.A5V and p.G42S (previously denoted as p.A4V and p.G41S, respectively) *SOD1* substitutions have been consistently associated with a fast progressive phenotype, while patients with a p.H47R (previously denoted as p.H46R) mutation showed a more benign phenotype and a much longer life expectancy (Battistini et al., 2010; Takazawa et al., 2010). However, these mutations are marked by symptoms representing a varying degree of lower and/or upper motor neuron involvement of spinal, bulbar, and cortical regions. The genetic and phenotypic heterogeneities have been suggested to play critical roles in determining disease prognosis and in designing clinical trials (Manera et al., 2019). Whether there is a correlation between mutation status and clinical traits remains to be answered.

Here, we report a novel missense mutation in the *SOD1* gene (p.R116S), causing rapid deteriorating lower motor neuron symptoms in a fALS. The function of the mutant protein was explored to try to find a correlation between perturbed protein functions and clinical severity.

# METHODS

## Pedigree

The proband was from a Chinese family of Han ethnicity with 28 known family members from three generations (**Figure 1A**). Overall, four affected family members have been diagnosed with MND. Two members from this family (proband and patient II-6) were seen at the ALS/MND Clinic, Xuanwu Hospital, Beijing, China, and further information on additional family members (II-7, III-7) was provided by the proband afterward. Disease onset was defined as first reported symptoms of weakness, dysarthria, or dysphagia. Patients underwent extensive laboratory studies to rule out other causes of neuropathy. The study was approved by the local medical ethical committee, with all participants providing written informed consent.

## Genetic Analysis

DNA was extracted from peripheral blood samples using a standard phenol–chloroform method. Whole-exome sequencing was used to detect gene mutations. DNA libraries were prepared with KAPA Library Preparation Kit and enriched using SeqCap hybridization probes following the manufacturer's instructions (Roche, Basel, Switzerland). Captured DNA libraries were subsequently sequenced on the Illumina NovaSeq platform. Low-quality reads and adapter sequences were first excluded. Single-nucleotide variants (SNVs) and insertions and deletions (INDELs) were filtered using GATK. All called variants were annotated based on several public databases, including the 1,000 Genomes Project, Human Gene Mutation Database Professional, gnomAD, and China National Gene Bank (CNGB). The pathogenicity of the identified variants was analyzed mainly based on criteria following the American College of Medical Genetics and Genomics (ACMG) guidelines. Candidate mutation detected in exome sequencing was confirmed by Sanger sequencing.

## Protein Function Analysis and Structure Modeling

The sequence-based assessment of the potential pathogenicity of each variant was assessed by the SIFT web server (https://sift.bii.a-star.edu.sg), Provean Protein web server (http://provean.jcvi.org/genome_submit_2.php, PolyPhen-2 web server (http://genetics.bwh.harvard.edu/pph2/), CamSol web server (https://www-cohsoftware.ch.cam.ac.uk), and PopMuSiC web server (https://soft.dezyme.com). The protein crystal structure of the wild-type human SOD1 was retrieved from the Protein Data Bank repository (PDB code 2C9V), *in silico* mutagenesis was performed for the three mutations discussed in this study using Pymol (The PyMOL Molecular Graphics System, Schrödinger, LLC, http://www.pymol.org), and the Find function was used to identify potential hydrogen bonds.

# RESULTS

## Clinical Findings

The proband was a 50-year-old female presented with a 2.5-month history of progressive muscle weakness. She first experienced a heavy feeling in her right leg when climbing stairs. It worsened over the next 2 weeks that she had trouble standing up from a squatting position without support. The weakness continued to progress and spread to her left leg in the following 2 weeks. By then, she noticed muscle wasting in her right thigh and experienced significant difficulty climbing stairs as well as trouble getting up from a sitting even with support. She soon showed problems reaching upward using her right arm. Neurological examination revealed mild to moderate weakness of proximal muscle groups in the lower limbs of both sides and the right upper limb. Muscle atrophy of the right thigh was noted. Deep tendon reflexes were decreased to normal with down-going plantar reflex on both sides. Other pathological reflexes were not elicited. There were no abnormal findings on sensory system examination. There were no cognitive or mood changes. Pulmonary function test indicated restrictive deficit. Electromyography showed in both upper and lower limbs as well as thoracic paraspinal muscles a neurogenic pattern. Sensory and motor nerve conduction velocities were within normal ranges. $^{18}$F-FDG PET did not detect any hypometabolism at the motor and premotor cortices or frontal–temporal regions. 3 Tesla magnetic resonance imaging (MRI) of the brain and spinal cord were unremarkable. Diffusion tensor imaging (DTI) reconstruction of the bilateral cortical–spinal tracts (CSTs) was performed, and no overt deficits were observed (**Figure 1C**). Together, a suspected diagnosis of progressive muscular atrophy (PMA) was made, and based on the familial history of MND, her blood sample was also sent out for WES and genetic analysis.

The proband received standard treatment with Edaravone and Riluzole and was evaluated every 3–6 months. Unfortunately, due to the COVID-19 pandemic, her follow-ups were only conducted

*via* Telemedicine; therefore, detailed neurological check-ups were not possible. Her initial ALSFRS-R score was 39/48, which declined to 22/48 at her 3-month follow-up when she developed bulbar symptoms. Her ALSFRS-R score further decreased to 16/48 at her 5-month follow-up. One year after the disease onset, she was completely bedridden and required full-time non-invasive positive pressure ventilation (NIPVV). By then, her ALSFRS was 8/48. Six months later, she was admitted to the hospital and intubated due to respiratory failure and died 1 month after.

The proband's mother (patient II-6) was diagnosed of MND at age 67 and died 1 year later. Her symptoms started with weakness and muscle wasting of the lower limbs. Later, she experienced weakness in upper limbs and complained of shortness of breath. The symptoms were more pronounced on the right side. Upon neurological examination, substantial proximal > distal bilateral muscle weakness of her lower limbs and mild weakness of her right upper limb were found. Atrophies of multiple muscle groups were also observed. Deep tendon reflexes were diminished as well. No pathological reflexes were detected, and both cranial nerves and sensory modalities were intact. There was no cerebellar sign. The pulmonary function test indicated restrictive deficit. Needle EMG showed chronic and active denervation in bulbar and spinal regions.

Patient III-7 was diagnosed of MND at age 52 and died of respiratory failure 1 year later. Her initial symptom included weakness in the upper limb. Another patient (II-7) in this family was also diagnosed of MND with lower-limb onset when she was 62 years old and died at age 63. Both underwent detailed laboratory studies that did not reveal any other cause of diffuse denervation. Both the mother (II-3) and the grandmother (I-2) of patient III-7 were suspected mutation carriers who died of other causes over the age of 70 years.

## Genetic Analysis

The DNA sample of patient III-15 was subjected to whole-exome sequencing (WES) to get a comprehensive profile of genetic variants. A novel heterozygous c.346C>A mutation (NM_000454.5, p.R116S, previously denoted as p.R115S) in the *SOD1* gene, was revealed (**Figure 1B**). Due to the death of all other affected members, segregation of the mutations could not be confirmed in the family presented here. However, this mutation was absent in the databases of healthy controls, including Genome Aggregation Database, the Exome Sequencing Project (ESP), and Exome Aggregation Consortium (ExAC) databases. The mutation altered a highly conserved residue and was predicted to be deleterious by both SIFT and PROVEAN and ranked as Probably Damaging and Disease Causing by PolyPhen-2 and Mutation Taster, respectively (**Figure 1E**).

## Prediction on Mutated Protein Function

*SOD1* R116S mutation-induced alterations in protein functions were explored using a battery of algorithms. Mutation at this site resulted in a polar to non-polar amino acid change, implying a significant change in amino acid chemistry (Grantham score of 110) and affecting a highly conserved residue (**Figures 1D,E**). The impact of R116S mutation on protein solubility was predicted using the CamSol Intrinsic web server, which suggested that the R116S SOD1 mutant was less soluble compared to the wild-type protein. Additionally, free energy change upon mutation was predicted using PopMuSiC, and an arginine-to-serine change at codon 116 was predicted to have a calculated $\Delta\Delta G$ of 1.62, suggesting a destabilizing effect.

The potential effects of the mutation on SOD1 protein stabilization were further investigated by protein modeling. SOD1 protein is a homodimer composed of two β-barrel monomers with copper and zinc in the active site. Codon 116 is located at the dimer contact site, being part of the interlocking Greek key loop connecting two four-stranded anti-parallel β-sheets facing each other (**Figure 2A**) (Wright et al., 2019). Interatomic interactions between the side chain of mutated residue and residues in the vicinity were analyzed, and hydrogen bonds were represented in the form of pseudo bonds using Pymol. Arginine at position 116 was shown to form hydrogen bonds with neighboring amino acids, particularly glutamic acid 50 and cysteine 112. R116S was found to disrupt hydrogen bonding with E50 and C112, suggesting alterations in native protein conformation and structure destabilization, thereby serving as a basis for neurotoxicity (**Figures 2B,D**).

# DISCUSSION

We found a Chinese fALS family carrying a novel R116S change in *SOD1*. All four patients displayed an aggressive disease course with disease duration less than 2 years from initial symptom to death. Molecular modeling and bioinformatics analysis on mutant SOD1 protein suggested that loss of hydrogen bonds at the dimer interface, together with decreases in protein solubility and stability, which might explain the aggressive nature associated with the mutation.

The mean age at onset was 58 years in the family. The proband and her mother were characterized by a rapid disease progression with prominent lower motor neuron symptoms. Physical examinations showed no evidence of upper motor neuron signs, and there were no CSTs deficits on DTI in the proband, suggesting a progressive spinal muscular atrophy (PMA) phenotype of MND. It was suggested that around 10% of MND patients clinically presented as PMA (Van den Berg-Vos et al., 2009). Genetic causes might partially determine the motor neuron/clinical phenotypes. Some specific *SOD1* missense mutations have been associated with a consistent clinical presentation. Here, we summarized a group of *SOD1* mutations that have been found to cause a predominant lower motor neuron syndrome (**Table 1**). Most of the mutations belong to exon 4 and are located at β-strand structures of SOD1 protein (**Figure 2A**; **Table 1**). Regardless of the mutation status, these patients were more likely to have a limb onset with bulbar involvement during disease course and were also associated with a rapid disease progression. Consistent with clinical features, pathological findings showed severe spinal lower motor neuronal loss with minimal involvement of upper

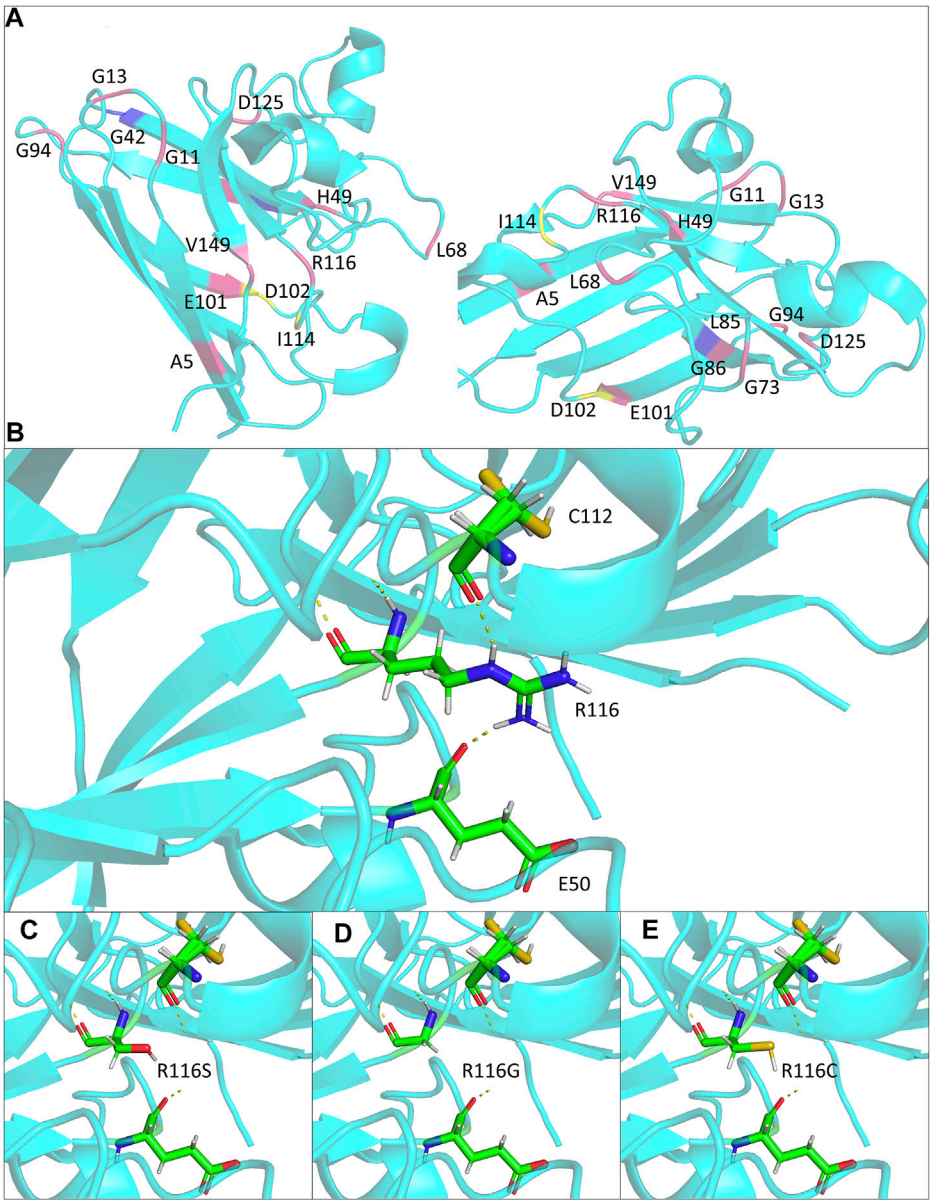



**FIGURE 2 |** Structural modeling of the ALS-associated SOD1 R116 mutants in the study. **(A)** Monomer views of the SOD1 homodimer (PDB: 2c9v). Mutations discussed in **Table 1** are highlighted and annotated. **(B–E)** Close lookup views of the hydrogen bonds (yellow dash line) for R116 and mutants R116S, R116G, and R116C, respectively.

motor neurons of motor cortex and cortical spinal tracts (**Table 1**). Of notice, different missense mutations of the same site at the *SOD1* gene could lead to different disease phenotypes (i.e., A5V, A5T, previously denoted as A4V, A4T) (**Table 1**). It is reasonable to speculate that the relative position, property of the side chain of the mutated residue, and nature of the adjacent amino acid of the mutated residues were likely to affect the stability and conformation of SOD1 protein, thus conferring toxicity on mutant proteins.

To our knowledge, the R116S mutation has never been reported so far (http://alsod.iop.kcl.ac.uk). However, R116G was found to be the most common *SOD1* mutation in German fALS patients, which accounts for up to 44% of *SOD1* associated ALS patients and was suggested to arise from a common founder (Niemann et al., 2004; Rabe et al., 2010). A different point mutation R116C was identified in one Italian sporadic ALS patient (Tortelli et al., 2013). Consistently, these two mutations are also associated with rapid disease progression and predominant lower motor neuron symptoms (**Table 2**). R116 mutations of *SOD1* in three distinct ethnic groups caused the similar severe phenotype, indicating that R116 is critical for the protein function. Additionally, it has been suggested that protein aggregation is favored by mutations that bring the net charge of the protein closer to neutrality (Lindberg et al., 2005). Thus, the

**TABLE 1 |** Clinical table of *SOD1* fALS patients with predominant lower motor neuron symptoms, color-coded by symptom group. The reference protein sequence for mutations included in the table is NP_000445.1.

| Mutation | Exon | Structure | Site of onset | Bulbar sign | UMN sign | Weakness pattern | Disease course | CNS pathology | Primary references |
|---|---|---|---|---|---|---|---|---|---|
| A5V | 1 | β1-strand | Limb Bulbar | Yes | Mild-absent | Distal Bulbar | Rapid (1 ± 0.5 years) | Severe loss LMNs, absent to mild UMN abnormalities | Andersen et al. (1997); Cudkowicz et al. (1998) |
| A5T | 1 | β1-strand | Limb | Yes | Mild-absent | Proximal | Rapid | Severe LMN loss, mild CST damage, no UMN loss | (Takahashi et al., 1994; Aksoy et al., 2003) |
| G11V | 1 | β1-strand | Limb Bulbar | Yes | Mild-absent | Distal Bulbar | Rapid (14.7 ± 5.3 m) | N/A | Kim et al. (2007) |
| G13R | 1 | Loop I | Limb | Yes | Absent | Distal | Slow | N/A | Penco et al. (1999) |
| G42S[a] | 2 | β4-strand | Limb | Absent | Absent | N/A | Rapid (<15 m) | N/A | Subramony et al. (2011) |
| | | | Limb | Yes | Mild | N/A | Rapid (11.6 ± 1.7 m) | N/A | Rainero et al. (1994) |
| H49G | 2 | β4-strand | N/A | N/A | Absent | N/A | Rapid (8 m) | N/A | Enayat et al. (1995) |
| L68P | 3 | Loop IV | Limb | Absent | Absent | Distal | Slow | N/A | del Grande et al. (2011) |
| G73C | 3 | Loop IV | Limb | Yes | Absent | Proximal | Slow (53 m) | N/A | Stewart et al. (2006) |
| L85V[a] | 4 | β5-strand | Limb | Yes | Absent | Distal | Rapid (<1.5 years) | N/A | Aoki et al. (1995) |
| | | | Limb | Yes | Mild | N/A | Moderate (4.8 ± 2.7 years) | N/A | Ceroni et al. (1999) |
| G86S | 4 | β5-strand | Limb | Absent | Absent | Distal | Rapid (15–18 m) | N/A | Takazawa et al. (2010) |
| G94C | 4 | Loop V | Limb | Absent | Absent | Distal | Slow (153.0 ± 46.1 m) | Severe LMN loss, minimal involvement of CST | Regal et al. (2006) |
| E101K | 4 | β6-strand | Limb | Yes | Absent | N/A | Slow (10-20 years) | TDP-43 negative neuropathology | Rabe et al. (2010) |
| D102N[b] | 4 | β6-strand | Limb | Mild | Absent | Distal | Rapid (~28 m) | Severe loss of LMNs, relatively well preserved UMNs, undamaged pyramidal tracts | Cervenakova et al. (2000) |
| I114T[b] | 4 | Loop VI | Limb | N/A | Absent | Distal | Rapid (~3 years) | Severe LMN loss, well preserved UMNs and CSTs | Rouleau et al. (1996) |
| R116G | 4 | β7-strand | Limb | Yes | Absent | Proximal | Rapid (2-3 years) | N/A | Rabe et al. (2010) |
| R116S | 4 | β7-strand | Limb | Yes | Absent | Proximal | Rapid (<2 years) | N/A | our study |
| D125G | 5 | Loop VII | Limb | Yes | Absent | Proximal | Moderate (2–5 years) | N/A | Ricci et al. (2019) |
| V149I | 5 | β8-strand | Limb Bulbar | Yes | Absent | Distal Bulbar | Rapid (1.8 ± 0.5 years) | N/A | Abe et al. (1996) |

[a]*The reported mutation has been reported to show predominant LMN symptoms but with different accompanying symptoms in separate studies.*
[b]*The reported mutation has been reported to manifest as classical ALS phenotype in separate studies.*
*N/A, not available.*

substitution of a positively charged amino acid (arginine) with a neutral one (glycine, cysteine, or serine) would probably promote protein aggregation and lead to neurotoxicity.

Within the resolved protein structure of SOD1, R116 forms hydrogen bonds with E50 and C112, respectively, both of which were reported to play important roles in SOD1 protein stability and function. Specifically, E50 locates in a loop structure, whose alterations could cause β-sheet partial unfolding, rendering it more prone to aggregation (Ding and Dokholyan, 2008). On the other hand, C112 is a primary target for redox attacks, the oxidation of which results in conformational changes of SOD1 protein and formation of oligomers, ultimately leading to axon transport deficits and motor neuron death (Bosco et al., 2010; Nagano et al., 2015). Notably, all R116 substitution mutations reported to date, including R116S, R116C, and R116G, are predicted to abolish the two hydrogen bonds and likely to affect the relative positions of E50 and C112 (**Figures 2C–E**), which in turn might cause β-sheet partial unfolding and might make C112 more accessible for oxidation, leading to mutant SOD1 misfolding and toxicity.

Taken together, we report a *SOD1* mutation site causing the most consistent phenotype: ubiquitous limb onset characterized by prominent LMN presentation with rapid progression, adding to the list of ALS causing mutations bearing a genotype–phenotype correlation. Additionally, it displays a prognostic value of *SOD1* R116 missense mutation.

**TABLE 2 |** Clinical features of previously reported patients carrying *SOD1* mutations of R116. The reference cDNA and protein sequences for mutations included in the table are NM_000454.5 and NP_000445.1.

| Family/ Individual | Family history | Ethnicity | Patient I.D. | Mutation cDNA (NM_000454.5) | Amino acid change (NP_000445.1) | Gender | Age of onset | Age of death | Avg. age of onset | Site of onset | Bulbar symptom | Proximal vs. distal | Disease duration | Affected members | Generations (n) | References |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | fALS | German | Mother<br>Son | c.346C>G | p.R116G | F<br>M | 66<br>43 | 68<br>45 | 60 | LL | Y<br>Y | Proximal<br>Proximal | 2-3 y | 6 | 3 | Rabe et al. (2010) |
| 2 | fALS | German | N/A | c.346C>G | p.R116G | N/A | 67 | N/A | | LL | Y | Proximal | | 2 | N/A | Rabe et al. (2010) |
| 3 | fALS | German | N/A | c.346C>G | p.R116G | N/A | 64 | N/A | | LL | Y | Proximal | | 2 | N/A | Rabe et al. (2010) |
| 4 | fALS | German | III-ALS3.7<br>II-2<br>II-3 | c.346C>G | p.R116G | N/A<br>F<br>F | 42<br>66<br>48 | N/A<br>68<br>51 | | LL | Y | Proximal | | 5 | 3 | Kostrzewa et al. (1994); Niemann et al. (2004); Rabe et al., 2010 |
| 5 | fALS | German | ALS73.1 | c.346C>G | p.R116G | N/A | 52 | N/A | | LL | Y | Proximal | | 2 | 5 | Niemann et al. (2004) |
| 6 | fALS | German | ALS129.1 | c.346C>G | p.R116G | N/A | 60 | N/A | | LL | Y | Proximal | | 5 | 4 | Niemann et al. (2004) |
| 7 | fALS | German | ALS179.1 | c.346C>G | p.R116G | N/A | 74 | N/A | | LL | Y | Proximal | | 3 | 3 | Niemann et al. (2004) |
| 8 | fALS | German | N/A | c.346C>G | p.R116G | N/A | 56 | N/A | | LL | Y | Proximal | | 3 | N/A | Rabe et al. (2010) |
| 9 | fALS | German | N/A | c.346C>G | p.R116G | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | Muller et al. (2018) |
| 10 | fALS | Chinese | III-15<br>III-7<br>II-6<br>II-7 | c.346C>A | p.R116S | F<br>F<br>F<br>F | 50<br>52<br>67<br>63 | 51<br>53<br>68<br>64 | 51<br>65 | LL<br>UL<br>LL<br>LL | Y<br>Y<br>Y<br>Y | Proximal | 20 mon<br><1 y<br>14 mon<br><1 y | 4 | 2 | our study |
| 11 | sALS | Italian | N/A | c.346C>T | p.R116C | M | 72 | N/A | N/A | UL | Y | N/A | >2 y | N/A | N/A | Tortelli et al. (2013) |
| 12 | N/A | Japanese | patient 21 | c.346C>G | p.R116G | M | 66 | N/A | N/A | LL | N/A | N/A | >22 months | N/A | N/A | Sato et al. (2005) |
| 13 | N/A | American | N/A | N/A | p.R116G | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | Andersen et al. (2003) |

*LL, lower limb; UL, upper limb; N/A, not available; Y, yes.*

## DATA AVAILABILITY STATEMENT

The datasets for this article are not publicly available due to concerns regarding participant/patient anonymity. Requests to access the datasets should be directed to the corresponding author.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Xuanwu Hospital Medical Ethical Committee. The patients/participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

Conceptualization: XW and YD; patient evaluation: XW, WZ, LD, SZ, and YD; electrophysiology evaluation: YL and HC; protein structure modeling: XW and NX; imaging acquisition and analysis: QL and JL; writing: XW and YD; resources: MW, MX, SW; and supervision: X-MS, JL, and YD. All authors read, discussed, and approved the manuscript as submitted.

## FUNDING

This work was funded by the National Natural Science Foundation of China (81801255, 82001352), National Key R&D Program of China Precision Medicine Project (No. 2017YFC0907700), Beijing Hospitals Authority Youth Program (QML20190803), Capital Health Development Research Project (2018-4-2015), Beijing Municipal Administration of Hospitals' Ascent Plan (Code: DFL20180802), and the Natural Science Foundation of Jiangsu (Grant BK20210017).

## ACKNOWLEDGMENTS

We thank our patients and their family members for participating in this study. We thank members of the Division of Neuromuscular Disease at Department of Neurology, Xuanwu Hospital, for the suggestions and critical evaluation of this work.

## REFERENCES

Abe, K., Aoki, M., Ikeda, M., Watanabe, M., Hirai, S., and Itoyama, Y. (1996). Clinical Characteristics of Familial Amyotrophic Lateral Sclerosis with CuZn Superoxide Dismutase Gene Mutations. *J. Neurol. Sci.* 136, 108–116. doi:10.1016/0022-510x(95)00314-r

Aksoy, H., Dean, G., Elian, M., Deng, H. X., Deng, G., Juneja, T., et al. (2003). A4T Mutation in the SOD1 Gene Causing Familial Amyotrophic Lateral Sclerosis. *Neuroepidemiology* 22, 235–238. doi:10.1159/000070564

Andersen, P. M., Nilsson, P., Keränen, M. L., Forsgren, L., Hägglund, J., Karlsborg, M., et al. (1997). Phenotypic Heterogeneity in Motor Neuron Disease Patients with CuZn-Superoxide Dismutase Mutations in Scandinavia. *Brain* 120 (Pt 10), 1723–1737. doi:10.1093/brain/120.10.1723

Andersen, P. M., Sims, K. B., Xin, W. W., Kiely, R., O'neill, G., Ravits, J., et al. (2003). Sixteen Novel Mutations in the Cu/Zn Superoxide Dismutase Gene in Amyotrophic Lateral Sclerosis: a Decade of Discoveries, Defects and Disputes. *Amyotroph. Lateral Scler. Other Motor Neuron Disord.* 4, 62–73. doi:10.1080/14660820310011700

Aoki, M., Abe, K., Houi, K., Ogasawara, M., Matsubara, Y., Kobayashi, T., et al. (1995). Variance of Age at Onset in a Japanese Family with Amyotrophic Lateral Sclerosis Associated with a Novel Cu/Zn Superoxide Dismutase Mutation. *Ann. Neurol.* 37, 676–679. doi:10.1002/ana.410370518

Battistini, S., Ricci, C., Giannini, F., Calzavara, S., Greco, G., Del Corona, A., et al. (2010). G41SSOD1mutation: A Common Ancestor for Six ALS Italian Families with an Aggressive Phenotype. *Amyotroph. Lateral Scler.* 11, 210–215. doi:10.3109/17482960902995592

Bosco, D. A., Morfini, G., Karabacak, N. M., Song, Y., Gros-Louis, F., Pasinelli, P., et al. (2010). Wild-type and Mutant SOD1 Share an Aberrant Conformation and a Common Pathogenic Pathway in ALS. *Nat. Neurosci.* 13, 1396–1403. doi:10.1038/nn.2660

Ceroni, M., Malaspina, A., Poloni, T. E., Alimonti, D., Rognoni, F., Habgood, J., et al. (1999). Clustering of ALS Patients in central Italy Due to the Occurrence of the L84F SOD1 Gene Mutation. *Neurology* 53, 1064. doi:10.1212/wnl.53.5.1064

Cervenakova, L., Protas, I. I., Hirano, A., Votiakov, V. I., Nedzved, M. K., Kolomiets, N. D., et al. (2000). Progressive Muscular Atrophy Variant of Familial Amyotrophic Lateral Sclerosis (PMA/ALS). *J. Neurol. Sci.* 177, 124–130. doi:10.1016/s0022-510x(00)00350-6

Cudkowicz, M. E., Mckenna-Yasek, D., Chen, C., Hedley-Whyte, E. T., and Brown, R. H., Jr. (1998). Limited Corticospinal Tract Involvement in Amyotrophic Lateral Sclerosis Subjects with the A4V Mutation in the Copper/zinc Superoxide Dismutase Gene. *Ann. Neurol.* 43, 703–710. doi:10.1002/ana.410430604

Del Grande, A., Luigetti, M., Conte, A., Mancuso, I., Lattante, S., Marangi, G., et al. (2011). A Novel L67P SOD1 Mutation in an Italian ALS Patient. *Amyotroph. Lateral Scler.* 12, 150–152. doi:10.3109/17482968.2011.551939

Ding, F., and Dokholyan, N. V. (2008). Dynamical Roles of Metal Ions and the Disulfide Bond in Cu, Zn Superoxide Dismutase Folding and Aggregation. *Proc. Natl. Acad. Sci.* 105, 19696–19701. doi:10.1073/pnas.0803266105

Enayat, Z. E., Orrell, R. W., Claus, A., Ludolph, A., Bachus, R., Brockmuller, J., et al. (1995). Two Novel Mutations in the Gene for Copper Zinc Superoxide Dismutase in UK Families with Amyotrophic Lateral Sclerosis. *Hum. Mol. Genet.* 4, 1239–1240. doi:10.1093/hmg/4.7.1239

Kim, W., Kim, J. S., Lee, K. S., Gwoun, Y. J., Kim, J. M., and Lee, K. H. (2007). Anticipation and Phenotypic Heterogeneity in Korean Familial Amyotrophic Lateral Sclerosis with Superoxide Dismutase 1 Gene Mutation. *J. Clin. Neurol.* 3, 38–44. doi:10.3988/jcn.2007.3.1.38

Kostrzewa, M., Burck-Lehmann, U., and Muller, U. (1994). Autosomal Dominant Amyotrophic Lateral Sclerosis: a Novel Mutation in the Cu/Zn Superoxide Dismutase-1 Gene. *Hum. Mol. Genet.* 3, 2261–2262. doi:10.1093/hmg/3.12.2261

Lindberg, M. J., Bystrom, R., Boknas, N., Andersen, P. M., and Oliveberg, M. (2005). Systematically Perturbed Folding Patterns of Amyotrophic Lateral Sclerosis (ALS)-associated SOD1 Mutants. *Proc. Natl. Acad. Sci. U S A.* 102, 9754–9759. doi:10.1073/pnas.0501957102

Liu, X., He, J., Gao, F. B., Gitler, A. D., and Fan, D. (2018). The Epidemiology and Genetics of Amyotrophic Lateral Sclerosis in China. *Brain Res.* 1693, 121–126. doi:10.1016/j.brainres.2018.02.035

Manera, U., Calvo, A., Daviddi, M., Canosa, A., Vasta, R., Torrieri, M. C., et al. (2019). Regional Spreading of Symptoms at Diagnosis as a Prognostic Marker in Amyotrophic Lateral Sclerosis: a Population-Based Study. *J. Neurol. Neurosurg. Psychiatry* 91 (3), 291–297. doi:10.1136/jnnp-2019-321153

Mathis, S., Goizet, C., Soulages, A., Vallat, J. M., and Masson, G. L. (2019). Genetics of Amyotrophic Lateral Sclerosis: A Review. *J. Neurol. Sci.* 399, 217–226. doi:10.1016/j.jns.2019.02.030

Muller, K., Brenner, D., Weydt, P., Meyer, T., Grehl, T., Petri, S., et al. (2018). Comprehensive Analysis of the Mutation Spectrum in 301 German ALS Families. *J. Neurol. Neurosurg. Psychiatry* 89, 817–827. doi:10.1136/jnnp-2017-317611

Nagano, S., Takahashi, Y., Yamamoto, K., Masutani, H., Fujiwara, N., Urushitani, M., et al. (2015). A Cysteine Residue Affects the Conformational State and Neuronal Toxicity of Mutant SOD1 in Mice: Relevance to the Pathogenesis of ALS. *Hum. Mol. Genet.* 24, 3427–3439. doi:10.1093/hmg/ddv093

Niemann, S., Joos, H., Meyer, T., Vielhaber, S., Reuner, U., Gleichmann, M., et al. (2004). Familial ALS in Germany: Origin of the R115G SOD1 Mutation by a Founder Effect. *J. Neurol. Neurosurg. Psychiatry* 75, 1186–1188. doi:10.1136/jnnp.2003.028324

Penco, S., Schenone, A., Bordo, D., Bolognesi, M., Abbruzzese, M., Bugiani, O., et al. (1999). A SOD1 Gene Mutation in a Patient with Slowly Progressing Familial ALS. *Neurology* 53, 404–406. doi:10.1212/wnl.53.2.404

Rabe, M., Felbecker, A., Waibel, S., Steinbach, P., Winter, P., Muller, U., et al. (2010). The Epidemiology of CuZn-SOD Mutations in Germany: a Study of 217 Families. *J. Neurol.* 257, 1298–1302. doi:10.1007/s00415-010-5512-9

Rainero, I., Pinessi, L., Tsuda, T., Vignocchi, M. G., Vaula, G., Calvi, L., et al. (1994). SOD1 Missense Mutation in an Italian Family with ALS. *Neurology* 44, 347–349. doi:10.1212/wnl.44.2.347

Regal, L., Vanopdenbosch, L., Tilkin, P., Van Den Bosch, L., Thijs, V., Sciot, R., et al. (2006). The G93C Mutation in Superoxide Dismutase 1: Clinicopathologic Phenotype and Prognosis. *Arch. Neurol.* 63, 262–267. doi:10.1001/archneur.63.2.262

Ricci, C., Giannini, F., Intini, E., and Battistini, S. (2019). Genotype-phenotype Correlation and Evidence for a Common Ancestor in Two Italian ALS Patients with the D124G SOD1 Mutation. *Amyotroph. Lateral Scler. Frontotemporal Degener* 20, 611–614. doi:10.1080/21678421.2019.1621345

Rouleau, G. A., Clark, A. W., Rooke, K., Pramatarova, A., Krizus, A., Suchowersky, O., et al. (1996). SOD1 Mutation Is Associated with Accumulation of Neurofilaments in Amyotrophic Lateral Sclerosis. *Ann. Neurol.* 39, 128–131. doi:10.1002/ana.410390119

Sato, T., Nakanishi, T., Yamamoto, Y., Andersen, P. M., Ogawa, Y., Fukada, K., et al. (2005). Rapid Disease Progression Correlates with Instability of Mutant SOD1 in Familial ALS. *Neurology* 65, 1954–1957. doi:10.1212/01.wnl.0000188760.53922.05

Stewart, H. G., Mackenzie, I. R., Eisen, A., Brannstrom, T., Marklund, S. L., and Andersen, P. M. (2006). Clinicopathological Phenotype of ALS with a Novel G72C SOD1 Gene Mutation Mimicking a Myopathy. *Muscle Nerve* 33, 701–706. doi:10.1002/mus.20495

Subramony, S. H., Ashizawa, T., Langford, L., Mckenna, R., Avvaru, B., Siddique, T., et al. (2011). Confirmation of the Severe Phenotypic Effect of Serine at Codon 41 of the Superoxide Dismutase 1 Gene. *Muscle Nerve* 44, 499–502. doi:10.1002/mus.22117

Takahashi, H., Makifuchi, T., Nakano, R., Sato, S., Inuzuka, T., Sakimura, K., et al. (1994). Familial Amyotrophic Lateral Sclerosis with a Mutation in the Cu/Zn Superoxide Dismutase Gene. *Acta Neuropathol.* 88, 185–188. doi:10.1007/bf00294513

Takazawa, T., Ikeda, K., Hirayama, T., Kawabe, K., Nakamura, Y., Ito, H., et al. (2010). Familial Amyotrophic Lateral Sclerosis with a Novel G85S Mutation of Superoxide Dismutase 1 Gene: Clinical Features of Lower Motor Neuron Disease. *Intern. Med.* 49, 183–186. doi:10.2169/internalmedicine.49.2720

Tortelli, R., Conforti, F. L., Cortese, R., D'errico, E., Distaso, E., Mazzei, R., et al. (2013). Amyotrophic Lateral Sclerosis: a New Missense Mutation in the SOD1 Gene. *Neurobiol. Aging* 34, 1709 e3–1709 e5. doi:10.1016/j.neurobiolaging.2012.10.027

Turner, M. R., Hardiman, O., Benatar, M., Brooks, B. R., Chio, A., De Carvalho, M., et al. (2013). Controversies and Priorities in Amyotrophic Lateral Sclerosis. *Lancet Neurol.* 12, 310–322. doi:10.1016/s1474-4422(13)70036-x

Van Den Berg-Vos, R. M., Visser, J., Kalmijn, S., Fischer, K., De Visser, M., De Jong, V., et al. (2009). A Long-Term Prospective Study of the Natural Course of Sporadic Adult-Onset Lower Motor Neuron Syndromes. *Arch. Neurol.* 66, 751–757. doi:10.1001/archneurol.2009.91

Wright, G. S. A., Antonyuk, S. V., and Hasnain, S. S. (2019). The Biophysics of Superoxide Dismutase-1 and Amyotrophic Lateral Sclerosis. *Q. Rev. Biophys.* 52, e12. doi:10.1017/s003358351900012x

Zou, Z. Y., Zhou, Z. R., Che, C. H., Liu, C. Y., He, R. L., and Huang, H. P. (2017). Genetic Epidemiology of Amyotrophic Lateral Sclerosis: a Systematic Review and Meta-Analysis. *J. Neurol. Neurosurg. Psychiatry* 88, 540–549. doi:10.1136/jnnp-2016-315018

**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.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article or claim that may be made by its manufacturer is not guaranteed or endorsed by the publisher.

*Copyright © 2021 Wen, Zhu, Xia, Li, Di, Zhang, Chen, Lu, Wang, Xu, Wang, Shen, Lu and Da. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*

frontiers in Genetics

BRIEF RESEARCH REPORT
published: 26 November 2021
doi: 10.3389/fgene.2021.762987

# Novel Partial Exon 51 Deletion in the Duchenne Muscular Dystrophy Gene Identified *via* Whole Exome Sequencing and Long-Read Whole-Genome Sequencing

*Qianqian Li[1*†], Zhanni Chen[2†], Hui Xiong[2†], Ranran Li[3†], Chenguang Yu[4], Jingjing Meng[1], Panlai Shi[1] and Xiangdong Kong[1*]*

[1]Genetics and Prenatal Diagnosis Center, Department of Obstetrics and Gynecology, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, China, [2]Genokon Institute of Medical Science and Laboratory, Xiamen, China, [3]School of Life Science and Technology, Xinxiang Medical University, Xinxiang, China, [4]Key Laboratory of Molecular Biophysics of the Ministry of Education, Cardio-X Center, College of Life Science and Technology and Center for Human Genome Research, Huazhong University of Science and Technology, Wuhan, China

OPEN ACCESS

**Edited by:**
Natália Duarte Linhares, Genuity Science, Ireland

**Reviewed by:**
Chun Hang Au, Hong Kong Genome Institute (HKGI), Hong Kong SAR, China
Sakthivel Murugan, Medgenome Pvt. Ltd., India

***Correspondence:**
Qianqian Li
337030810@qq.com
Xiangdong Kong
kongxd@263.net

†These authors have contributed equally to this work

**Specialty section:**
This article was submitted to Genetics of Common and Rare Diseases, a section of the journal Frontiers in Genetics

**Received:** 23 August 2021
**Accepted:** 28 October 2021
**Published:** 26 November 2021

**Citation:**
Li Q, Chen Z, Xiong H, Li R, Yu C, Meng J, Shi P and Kong X (2021) Novel Partial Exon 51 Deletion in the Duchenne Muscular Dystrophy Gene Identified via Whole Exome Sequencing and Long-Read Whole-Genome Sequencing. Front. Genet. 12:762987. doi: 10.3389/fgene.2021.762987

Duchenne muscular dystrophy (DMD), one of the most common progressive and severely disabling neuromuscular diseases in children, can be largely attributed to the loss of function of the *DMD* gene on chromosome Xp21.2-p21.1. This paper describes the case of a 10-year-old boy diagnosed with DMD. Whole exome sequencing confirmed the hypothesized large partial exonic deletion of c.7310-11543_7359del (chrX: g.31792260_31803852del) spanning exon 51 and intron 50 in *DMD*. This large deletion was verified to be *de novo* by PCR, and the two breakpoints were further confirmed by Sanger sequencing and long-read whole-genome sequencing. Notably, this partial exonic deletion was the only complex variation in the deep intron regions or intron–exon junction regions in *DMD*. In addition, the case study demonstrates the clinical importance of using multiple molecular genetic testing methods for the diagnosis of rare diseases.

**Keywords: Duchenne muscular dystrophy, partial exonic deletion, breakpoints, whole exome sequencing, long-read whole-genome sequencing**

## INTRODUCTION

Duchenne muscular dystrophy (DMD, MIM# 310200) is inherited in an X-linked recessive manner, occurring in 1/3,600 to 1/6,000 live male births (Bushby et al., 2010). DMD is associated with the *DMD* gene (OMIM* 300377), which consists of 79 exons and is one of the largest genes. *DMD* is located on chromosome Xp21.2-p21.1. About two-thirds of patients inherit DMD maternally, while the remaining cases are a result of *de novo* mutations in *DMD* (Lane et al., 1983). There are many *DMD* variants, including those with deletions (~60%) or duplication (~7%) of one or more exons, small insertions and deletions (INDELs) within an exon (~7%), single nucleotide variants (SNVs) (~20%), and rare mutations such as splice site or intronic mutations (<1%) (Aartsma-Rus et al., 2006) and partial exonic deletions (~5%). In 2018, Liu et al. observed a novel mutation of c.6241_c.6290 + 1109del1159insAC using targeted next-generation sequencing (NGS). This mutation was a 1,159-bp deletion spanning the last 50 bp of exon 43 and the first 1,109 bp of

intron 43, causing the partial deletion of exon 43 of *DMD* (Liu et al., 2018). In 2021, Geng et al. detected intron 44 deletion breakpoints using long-read whole-genome sequencing (LR-WGS). This study was the first to use LR-WGS to explore the possible mechanisms underlying exon deletions starting from intron 44 of *DMD* (Geng et al., 2021). Later in 2021, Chin et al. used LR-WGS to confirm that a 33-year-old G2P1 proband at 26 weeks of pregnancy with a heterozygous 426.1-kb duplication on chromosome Xp21.2 (chX:30936321–31362374, GRCh37/hg19) carried the duplication at chrX:30939526–31362638 as a direct repeat inserted downstream of *DMD* (Chin et al., 2021). In this study, we identified a *de novo* partial exon 51 deletion in *DMD* in a 10-year-old boy diagnosed with DMD by using whole exome sequencing (WES), Sanger sequencing, and LR-WGS.

# MATERIALS AND METHODS

## Ethics Statement

The family provided written informed consent, and this study was approved by the appropriate local institutional review boards on human subject research at the First Affiliated Hospital of Zhengzhou University.

## Quantitative Fluorescent PCR

The genetic relationship of the proband and the parents was confirmed by quantitative fluorescent polymerase chain reaction (QF-PCR) using the Goldeneye DNA ID System 20A Kit (Peoplespot, Beijing, China).

## Multiplex Ligation-dependent Probe Amplification

A multiplex ligation-dependent probe amplification (MLPA) assay was performed using the SALSA MLPA Kit P034/P035 for *DMD* (MRC-Holland, Amsterdam, Netherlands).

## Whole Exome Sequencing

WES was performed using Illumina library construction and capture kits (Illumina, San Diego, California, USA) according to the manufacturer's instructions. Briefly, genomic DNA was fragmented with enrichment Bead-Linked Transposomes (eBLT) in Tagmentation Buffer 1 (TB1) (Illumina), with a target fragment size of 200 to 300 bp. The fragment ends were repaired, and an adenosine residue was added to the 3′ end of the fragments. Adaptors were then ligated to the fragments using the Fast DNA Library Prep Set for Illumina CW3045M (CWBIO Inc., China). Exon-containing libraries were captured using IDT xGen exome baits (Integrated DNA Technologies, Inc., USA), and quality and purification were assessed using Qubit 4.0 (Thermo Fisher Scientific Inc., USA). Also, 150 bp pair-end sequencing was conducted on NovaSeq 6000 for sequencing depths greater than 100×.

After sequencing, paired-end reads were trimmed using Trimmomatic version 0.36 (Bolger et al., 2014) to filter out low-quality reads and adaptors from the dataset. High-quality reads were aligned to the human reference genome GRCh37/hg19 using Burrows–Wheeler Aligner MEM version 0.7.17 (Li and Durbin, 2009), and the duplicates were removed using Picard version 2.9.0 (http://broadinstitute.github.io/picard). Small variants were identified using the Genome Analysis Toolkit (GATK) version 3.8 (McKenna et al., 2010). Copy number variants (CNVs) were detected by comparing the coverage of depth between the target sample and a baseline, which was determined using other male control samples in the same pipeline. In addition, the breakpoints of CNVs were called in the patient's BAM file by detecting soft-clipped and abnormal mapping orientation reads. The script for calling CNVs was run using the R programming language, and details are described in **Supplementary File 1**.

## PCR and Sanger Sequencing

The breakpoints of the partial exonic deletion were confirmed by PCR and Sanger sequencing of DNA samples from the proband, the proband's mother, and a healthy man (a control subject unrelated to this family without any muscular diseases). The primers used for sequencing are listed in **Supplementary Table S1**. Primers were designed using Primer-BLAST and were based on the combination of sequences on both sides of the deletion. The PCR product of the mutant type was approximately 1 kb, while that of the wild type was approximately 12 kb. Lastly, the sequences were aligned to the human reference genome sequence (GRCh37/hg19) using BLAST version 2.9.0 (Camacho et al., 2009) to identify the breakpoints of the large deletion.

## Long-read Whole-genome Sequencing

The genomic DNA extracted from the fresh peripheral blood of the proband was purified with AMPure XP magnetic beads (Beckman Coulter, CA, USA) and electrophoresed to confirm its integrity. The concentration of the DNA was 44.0 ng $\mu l^{-1}$ when measured using Nanodrop 2000 (Thermo, Massachusetts, USA) and 43.4 ng $\mu l^{-1}$ when measured using Qubit 4.0. The DNA input mass for the single library preparation without topping-up or refuel was 2.0398 μg. LR-WGS was carried out without fragmentation using the SQK-LSK109 kit (ONT, Oxford, UK) on MinION (ONT) with R9.4 flow cells (#FLO-MIN106, ONT) according to the manufacturer's instructions (GDE_9603_v109_revW_Aug 14, 2019) for 72 h. Raw fast5 data were basecalled using Guppy GPU (version 4.3.4) with a high-accuracy base calling model. After base calling, the sequencing data were aligned to the GRCh37/hg19 reference genome using minimap2 (version 2.17-r941) with "map-ont" model, and structural variants (SVs) calling was performed using sniffles (version 1.0.11) with the option "-s 1 -r 1,000 -q 20".

# RESULTS

## Clinical Report

The proband was a 10-year-old boy who had developed more slowly than his peers since birth—for example, he was unable to walk until he was approximately 1.5 years old. As time progressed, both his calves became thick and hard (**Figure 1A**). He also struggled with walking, running, ascending stairs, squatting, and

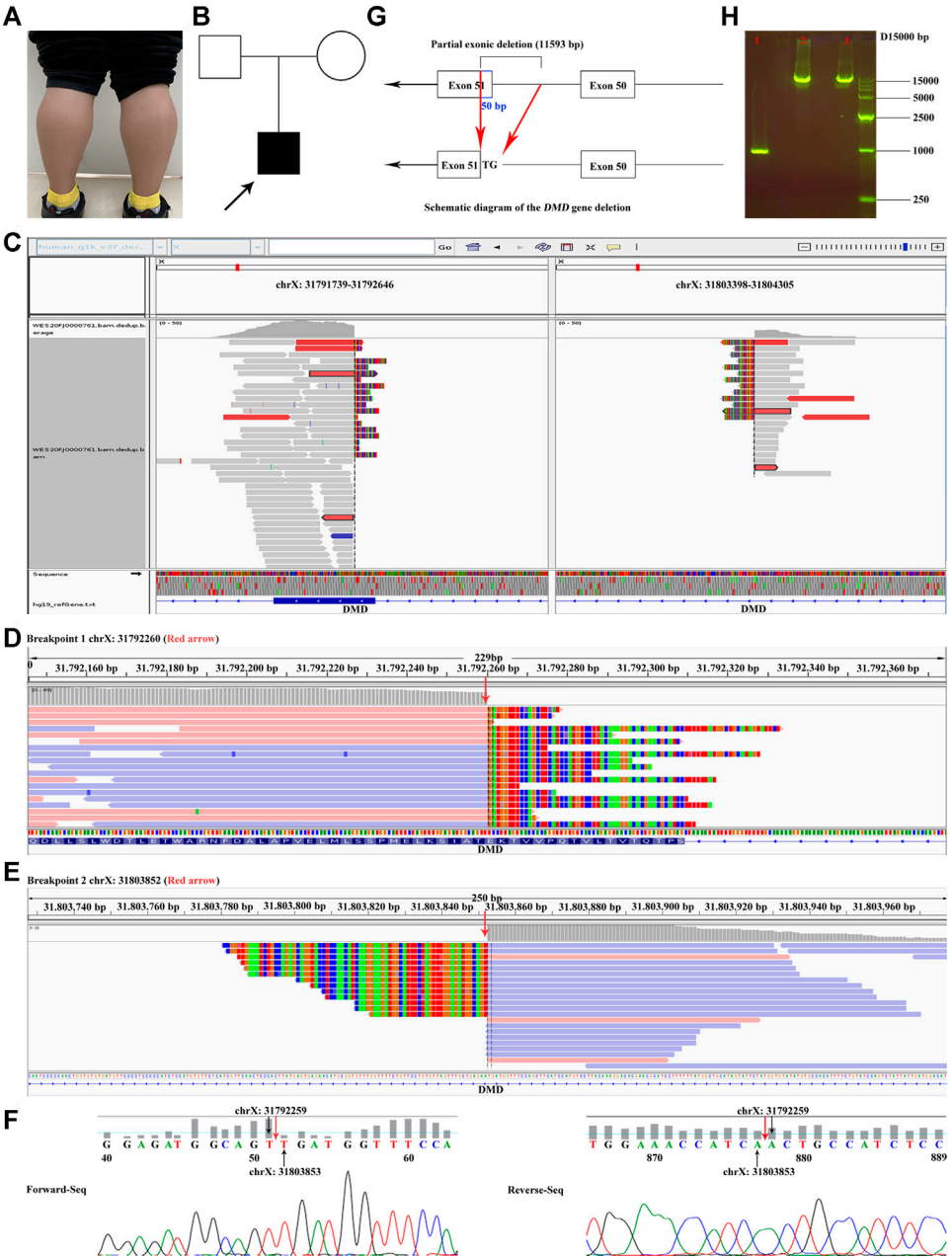



**FIGURE 1 |** Proband clinical features, breakpoints suggested by WES, and breakpoints validation. **(A)** Bilateral gastrocnemius hypertrophy of the proband in November 2020. **(B)** Family tree. **(C)** The overview of the two breakpoints after analyzing the original BAM file of the proband. One breakpoint is in chrX: 31791739–31792646; the other is in chrX:31803398–31804305. **(D, E)** The two breakpoints (red arrow) of *DMD* suggested by CNV and breakpoint analysis of the WES data. **(F)** Validation of the breakpoints by Sanger sequencing. Red arrow represents the positions of the breakpoints. **(G)** Diagrammatic sketch of the partial exonic deletion. **(H)** Results of agarose gel electrophoresis. 1. Proband; 2. Mother; 3. Negative control. The PCR product of the proband was approximately 1 kb, while those of the mother and the negative control were between 10 and 15 kb. **Supplementary Videos**. Various states of the proband: standing up after squatted down, going up and down stairs, and walking.

standing up. Because the proband's family was living in a remote and underserved area, the proband remained undiagnosed until July 2020, when he was diagnosed with DMD in Henan Children's Hospital. Electromyogram results confirmed myogenic damage of the nerves and muscles of both lower limbs and the left upper limb of the proband and the mother (**Supplementary Table S2**). The concentrations of creatine kinase (CK) and creatine kinase isoenzymes (CK-MB) in the venous serum of the proband were 14,984.0 U/L and 803.0 U/L (**Supplementary Table S3**), respectively. The proband had no family history of DMD

(**Figure 1B**). Since the pathogenetic cause was not confirmed in Henan Children's Hospital, the family came to our center in August 2020 to seek additional information about the genetic etiology of DMD.

## Deletions and Duplications in *DMD* of the Proband

MLPA analysis suggested that there was no deletion or duplication of exons in *DMD* (NM_004006.2) of the proband and the mother (**Supplementary Figures S1A,B, S2A,B**). Furthermore, the accuracy of kinship among the proband and the parents was confirmed by QF-PCR (**Supplementary Table S4**; **Supplementary Figure S3**). The quality control of the WES data is summarized in **Supplementary Table S5**. No candidate SNVs or small INDELs in *DMD* were found, and there was no significant variation in exon coverage. However, a 11,593-bp hemizygous deletion (c.7310-11543_7359del, chrX:g.31792260-31803852) spanning exon 51 and intron 50 in *DMD* was identified by CNV and breakpoint analysis of the WES data (**Figure 1C–E**). The 5′ breakpoint was only 50 bp from the 5′ end of exon 51 (**Figure 1G**).

## Additional Confirmation of the Breakpoints

The breakpoints of the large deletion were further confirmed by agarose gel electrophoresis and Sanger sequencing (**Figure 1F**). The results confirmed that the 5′ end 50 bp (chrX: 31792260–31792309) in exon 51 and 3′ end 11,543 bp in intron 50 of *DMD* were hemizygous deletion in the proband. Furthermore, this deletion was *de novo*, since the PCR product of the proband was approximately 1 kb, and those of the mother and the negative control were between 10 and 15 kb (**Figure 1H**).

The quality control of the LR-WGS data is presented in **Supplementary Tables S6–S8 and Supplementary Figure S4**. A total of two reads captured the large deletions in *DMD*, one of ~40 kb (read name: 97d7f30e-0f30-4ddc-b66d-c8c82c0fb221; chrX: 31790731–31792260, chrX:31803853–31842466) and the other of ~2.3 kb (read name: 1e46feed-d90e-44b8-bc47-56eb0495b76d; chrX:31791318–31792259, chrX:31803853–31805193) (**Supplementary Figure S5**, red arrows; **Supplementary VCF File**, line 1,569). The breakpoints of the large deletion were consistent with those identified by WES and Sanger sequencing. The structural variants (SVs; >1 kb) are listed in the **Supplementary VCF File** (GRCh37/hg19).

# DISCUSSION

MLPA is often the first assay performed in DMD case studies, as deletion(s) and duplication(s) of one or more exons account for ~67% of *DMD* variants. If the results of MLPA are negative, targeted NGS or WES will then be performed to further identify the genetic etiology of DMD (Aartsma-Rus et al., 2016). If the results are negative again, there are currently no prescriptive guidelines for further technical implementations. As a result, there are still undiagnosed cases of DMD. Most often, undiagnosed DMD cases are chimerism cases, rare mutation cases (e.g., gene inversion and translocation), and variable splicing mutation cases. Detection of deep intron or rare variant cases using LR-WGS will improve patient management and greatly contribute to the accuracy of prenatal diagnosis for mothers in the future.

In the present study, the MLPA assay was negative, indicating that no deletion or duplication of one or more exons was observed in either the proband or the mother (**Supplementary Figures S1, S2**; red arrow and red box), which may be due to that the probe of exon 51 for *DMD* (NM_004006.2; ligation site: 7,666–7,667; 24 nt adjacent to ligation site: GCTCTGGCAGAT-TTCAACCGGGCT) was located on the upstream of the breakpoints with reference to the direction of the genome (GRCh37/hg19). Considering the high specificity of the clinical phenotype and other biochemical test results (**Figure 1A**; **Supplementary Tables S2,S3**; **Supplementary Videos**), the proband was still diagnosed with DMD. To verify the previous results and explore whether there is gene heterogeneity in the disease, WES was performed for the proband. The results of routine SNV analysis of the WES data were still negative; however, CNV and breakpoint analysis of the WES data suggested a possible out-of-frame deletion (representing most mutational events in patients with severe DMD) of c.7310-11543_7359del (chrX: g.31792260_31803852del) of *DMD* (**Figure 1C–E**). Moreover, this large intronic deletion may affect splicing and exon skipping of *DMD*. In addition, large intronic deletions may result in clinical effects due to splicing and exon skipping. The breakpoints of chrX:31792260 and chrX:31803852 were in exon 51 and intron 50, respectively (**Figure 1G**). This was confirmed by Sanger sequencing (**Figure 1F**). LR-WGS was needed to further explore whether there were more complex variations in the deep intron regions or intron–exon junction regions in *DMD* and to determine if the LR-WGS data collected (**Supplementary Figure S5**; **Supplementary VCF File**) were consistent with those of WES and Sanger sequencing. Lastly, although the large complex deletion of the proband was *de novo* in this study (**Figure 1H**), the prenatal diagnosis of *DMD* for the mother during pregnancy had to be performed since the possibility of gonadal chimerism could not be ruled out.

The genetic diagnosis in this case demonstrates a new application of WES-CNV analysis, and breakpoint analysis verified the accuracy of LR-WGS for one-time analysis of SVs. Although CNV and breakpoint analysis has not regularly been performed in conjunction with WES, CNVs can be potentially detected by WES (Nam et al., 2016). Recent attempts at optimizing the analytic pipeline for CNV detection have brought newfound attention to WES-CNV analysis (Zhao et al., 2020; Zhai et al., 2021).

The major advantage of LR-WGS is that it can directly analyze the SVs of DMD cases that produce negative WES results, without fragmentation of the whole genome of the target sample. The major disadvantage of LR-WGS is its relatively high cost: performing LR-WGS on an ONT sequencer is expensive and cannot be implemented in routine practice. Sanger sequencing can also detect breakpoints. Although Sanger sequencing does not give researchers as much specificity as LR-WGS does, it is much less expensive and more applicable for routine purposes. The use of LR-

WGS for detecting intronic SVs and CNVs cannot be underestimated. In this study, it seems that ONT sequencing was not necessary when Sanger sequencing was sufficient to identify the deletion breakpoints. However, LR-WGS is carried out after Sanger sequencing for the aim described in the previous text. Moreover, this genetic diagnosis may not be suitable for the patient in this study. However, in the clinic, for patients with unremarkable results of MLPA and WES, ONT sequencing might be another selective method to determine the genetic etiology of rare genetic diseases, including DMD. This case may provide new data that will enhance the accessibility of research and development of *DMD* gene treatments for the partial deletion of exon 51.

A major limitation of this study was that RNA sequencing of a muscle biopsy sample from the proband was not performed to assess the overall clinical effect of this large deletion since it was extremely difficult to collect muscle biopsy sample from the proband.

In conclusion, this study reported a *de novo* partial exonic deletion in *DMD*. The discovery of this partial exonic deletion provided a theoretical basis for prenatal gene diagnosis for the mother in this family and will potentially help improve the efficiency of gene diagnosis.

## DATA AVAILABILITY STATEMENT

The datasets for this article are not publicly available due to concerns regarding participant/patient anonymity. Requests to access the datasets should be directed to the corresponding authors.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Declaration of Helsinki and was conducted in accordance with the International Ethical Guidelines for Biomedical Research Involving Human Subjects (CIOMS). 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) for the publication of any potentially identifiable images or data included in this article.

## AUTHOR CONTRIBUTIONS

QL, ZC, HX, RL, and XK designed the research. QL, ZC, HX, and JM carried out the experiments and data analysis. QL, ZC, and HX wrote the article. RL, CY, PS, and XK revised the article.

## FUNDING

National Natural Science Foundation of China (No. 82000321) and Henan Educational Committee Program for Science and Technology Development of Universities (22A310022).

## ACKNOWLEDGMENTS

The authors were very grateful to the family for participating in this study and appreciated Guangzhou Darui Biotechnology Co., Ltd. for the technical support to LR-WGS.

## SUPPLEMENTARY MATERIAL

The supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.762987/full#supplementary-material

## REFERENCES

Aartsma-Rus, A., Ginjaar, I. B., and Bushby, K. (2016). The Importance of Genetic Diagnosis for Duchenne Muscular Dystrophy. *J. Med. Genet.* 53 (3), 145–151. doi:10.1136/jmedgenet-2015-103387

Aartsma-Rus, A., Van Deutekom, J. C. T., Fokkema, I. F., Van Ommen, G.-J. B., and Den Dunnen, J. T. (2006). Entries in the Leiden Duchenne Muscular Dystrophy Mutation Database: an Overview of Mutation Types and Paradoxical Cases that Confirm the reading-frame Rule. *Muscle Nerve* 34 (2), 135–144. doi:10.1002/mus.20586

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a Flexible Trimmer for Illumina Sequence Data. *Bioinformatics* 30 (15), 2114–2120. doi:10.1093/bioinformatics/btu170

Bushby, K., Finkel, R., Birnkrant, D. J., Case, L. E., Clemens, P. R., Cripe, L., et al. (2010). Diagnosis and Management of Duchenne Muscular Dystrophy, Part 1: Diagnosis, and Pharmacological and Psychosocial Management. *Lancet Neurol.* 9 (1), 77–93. doi:10.1016/S1474-4422(09)70271-6

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al. (2009). BLAST+: Architecture and Applications. *BMC Bioinformatics* 10, 421. doi:10.1186/1471-2105-10-421

Chin, H. L., O'Neill, K., Louie, K., Brown, L., Schlade-Bartusiak, K., Eydoux, P., et al. (2021). An Approach to Rapid Characterization of DMD Copy Number Variants for Prenatal Risk Assessment. *Am. J. Med. Genet. A.* 185 (8), 2541–2545. doi:10.1002/ajmg.a.62349

Geng, C., Tong, Y., Zhang, S., Ling, C., Wu, X., Wang, D., et al. (2021). Sequence and Structure Characteristics of 22 Deletion Breakpoints in Intron 44 of the DMD Gene Based on Long-Read Sequencing. *Front. Genet.* 12, 638220. doi:10.3389/fgene.2021.638220

Lane, R. J., Robinow, M., and Roses, A. D. (1983). The Genetic Status of Mothers of Isolated Cases of Duchenne Muscular Dystrophy. *J. Med. Genet.* 20 (1), 1–11. doi:10.1136/jmg.20.1.1

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, C., Deng, H., Yang, C., Li, X., Zhu, Y., Chen, X., et al. (2018). A Resolved Discrepancy between Multiplex PCR and Multiplex Ligation-dependent Probe Amplification by Targeted Next-Generation Sequencing Discloses a Novel Partial Exonic Deletion in the Duchenne Muscular Dystrophy Gene. *J. Clin. Lab. Anal.* 32 (8), e22575. doi:10.1002/jcla.22575

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

Nam, J. Y., Kim, N. K. D., Kim, S. C., Joung, J. G., Xi, R., Lee, S., et al. (2016). Evaluation of Somatic Copy Number Estimation Tools for Whole-

Exome Sequencing Data. *Brief Bioinform* 17 (2), 185–192. doi:10.1093/bib/bbv055

Zhai, Y., Zhang, Z., Shi, P., Martin, D. M., and Kong, X. (2021). Incorporation of Exome-Based CNV Analysis Makes Trio-WES a More Powerful Tool for Clinical Diagnosis in Neurodevelopmental Disorders. *A retrospective study* 42 (8), 990–1004. doi:10.1002/humu.24222

Zhao, L., Han, L., Yuan, X., Gao, K., and Duan, J. (2020). Comparative Study of Whole Exome Sequencing-Based Copy Number Variation Detection Tools. *BMC Bioinform.* 21 (1), 97. doi:10.1186/s12859-020-3421-1

**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.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

*Copyright © 2021 Li, Chen, Xiong, Li, Yu, Meng, Shi and Kong. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*

frontiers in Genetics

ORIGINAL RESEARCH
published: 17 December 2021
doi: 10.3389/fgene.2021.782629

# X Chromosome Inactivation Pattern and Pregnancy Outcome of Female Carriers of Pathogenic Heterozygous X-Linked Deletions

*Yuanyin Zhao[1,2†], Jia Li[1†], Limeng Dai[1], Yongyi Ma[3], Yun Bai[1] and Hong Guo[1,4*]*

[1]Department of Medical Genetics, College of Basic Medical Science, Army Medical University, Chongqing, China, [2]Department of Biochemistry and Molecular Biology, College of Basic Medical Science, Army Medical University, Chongqing, China, [3]Department of Gynaecology and Obstetrics, Southwest Hospital, Army Medical University, Chongqing, China, [4]Department of Gynaecology and Obstetrics, Xinqiao Hospital, Army Medical University, Chongqing, China

OPEN ACCESS

**Edited by:**
Kathleen M. Gorman,
Temple Street Children's University Hospital, Ireland

**Reviewed by:**
Eva Trevisson,
University of Padua, Italy
Olivia J. Veatch,
University of Kansas Medical Center, United States

***Correspondence:**
Hong Guo
guohong02@gmail.com

[†]These authors have contributed equally to this work and share first authorship

**Specialty section:**
This article was submitted to Genetics of Common and Rare Diseases, a section of the journal Frontiers in Genetics

**Received:** 24 September 2021
**Accepted:** 22 November 2021
**Published:** 17 December 2021

**Citation:**
Zhao Y, Li J, Dai L, Ma Y, Bai Y and Guo H (2021) X Chromosome Inactivation Pattern and Pregnancy Outcome of Female Carriers of Pathogenic Heterozygous X-Linked Deletions. Front. Genet. 12:782629. doi: 10.3389/fgene.2021.782629

Prenatal risk assessment of carriers of heterozygous X-linked deletion is a big challenge due to the phenotypic modification induced by X chromosome inactivation (XCI). Herein, we described four Chinese pedigrees with maternal-inherited X-deletions above 1 Mb. The pathogenic evaluation revealed that all X-deletions are harmful to heterozygous carriers; however, the asymptomatic pregnant female carriers in these families tremendously complicate the prognostic assessment of the unborn heterozygous embryos. In this study, we detected the XCI pattern of 11 female carriers of heterozygous X-linked deletions and 4 non-carrier females in these families and performed the first prenatal XCI pattern analysis in a fetal female carrier of heterozygous *PCDH19*-deletion to make risk prediction. In an adult female who lost one copy of the terminal of X chromosome short arm (Xp), a region enriching a large number of XCI escapees, the expression level of representative XCI escape genes was also detected. Pregnancy outcomes of all families were followed up or retrospected. Our research provides clinical evidence that X-deletions above 1 Mb are indeed associated with extremely skewed XCI. The favorable skewed XCI in combination with potential compensatory upregulation of XCI escapees would protect some but not all female carriers with pathogenic X-deletion from severe clinical consequences, mainly depending on the specific genetic contents involved in the deletion region. For *PCDH19*-disorder, the XCI pattern is considered as the decisive factor of phenotype expression, of which prenatal XCI assay using uncultured amniocytes could be a practicable way for risk prediction of this disease. These results provide valuable information about the usage of XCI assay in the prenatal risk assessment of heterozygous X-linked deletions.

**Keywords: heterozygous X-linked deletion, X chromosome inactivation (XCI), prenatal risk assessment, PCDH19-disorder, borjeson-forssman-lehmann syndrome**

## INTRODUCTION

X chromosome comprises 155 million base pairs, contains upward of 1,200 genes, and is covered by a large number of repetitive sequences (Schwartz, 2013). It is well known that aberrant recombination between nearby DNA repeats is the molecular foundation of the chromosome structural abnormalities; hence, the copy number variants (CNV) on the X chromosome have a high

incidence and many of them are associated with a spectrum of developmental disorders (Yuan et al., 2021). Compared to hemizygous males, female heterozygous carriers of pathogenic X-linked CNV usually present highly variable disease penetrance and phenotype expression, making a great difficulty for their prognostic evaluation, a leading cause of which is the phenotypic modification effect of X chromosome inactivation (XCI) (Puck and Willard, 1998; Brown and Robinson, 2000; Lyon, 2002).

XCI is an epigenetic process to equalize the dosage of X-linked genes between XX females and XY males through transcriptionally silencing one of the two X chromosomes in somatic cells of the mammalian female during early embryogenesis (Lyon, 1961). In most females, XCI is a random process. Parental X chromosome has the same probability to be inactivated in each cell at the initial time and then remains inactive in all descendant cells, ultimately giving rise to mosaicism of two roughly equal populations of cells expressing either the paternal or the maternal X chromosome (Lyon, 1961). However, sometimes the XCI is non-random, leading to a cell ratio significantly deviating from 1:1, which is called non-random (skewed) XCI (Minks et al., 2008). Generally, non-random XCI can arise through situations as below: (i) non-random XCI arising by chance, which refers to the stochastic statistical distribution of inactivation in a relatively small cell number when XCI takes place (Minks et al., 2008); (ii) primary non-random XCI, which is a rare condition that refers to the preferential choice of which X chromosome (paternal or maternal) is inactivated at the initial stage. This condition usually results from disruption of the key genomic elements that coordinate the XCI process (Rastan, 1982; Gendrel and Heard, 2011); (iii) secondary/selective non-random XCI, which refers to skewed XCI that takes place during the clonal propagation process of the post-inactivation cells rather than the initial "chromosome-choice" stage. In this situation, initial inactivation of the parental X chromosome is random, while descendant cells expressing one of the X chromosomes would be privileged in their cell growth and/or survival during subsequent mitosis, thus gradually establishing a skewed cell proportion in somatic cells (Gendrel and Heard, 2011). Secondary non-random XCI is most common and always linked to a carrier status of certain deleterious X-linked mutations including unbalanced structural abnormalities (deletions, duplications, and isochromosomes) (Leppig and Disteche, 2001). Cells containing such X chromosomes are always considered to lose growth and/or survival privilege and would be gradually eliminated with continuous mitosis. Thus, unbalanced abnormalities on the X chromosome are assumed to have a less phenotypic impact than that of similar size involving autosomes (Orstavik, 2009).

The epigenetic inactivation can spread the entire X chromosome, but not all X-linked genes are subject to the transcriptional silence. Genes that continue to be expressed from inactivated X are defined as XCI escape genes/escapees. It is estimated that at least 23% of X chromosome genes escape the XCI to some degree (Tukiainen et al., 2017), bi-allelic expression of which is essential to normal female development. Furthermore, XCI escapees are not distributed randomly along the X chromosome; many more are located on the distal short arm of the human X chromosome-Xp (21%) than on the long arm-Xq (3%) (Disteche, 1999). This finding provides an important implication for genetic counseling that structural imbalance on Xp may have greater clinical significance than that on Xq, where the effect could be largely mitigated by XCI.

Although X chromosomes with unbalanced structural abnormalities are believed to be selectively inactivated, XCI pattern analysis is rarely performed in this kind of clinical case. In this study, we collected four pedigrees carrying maternal-inherited pathogenic X-linked deletions above 1 Mb. We analyzed the XCI pattern of peripheral blood lymphocytes in all female carriers and several non-carrier females in these families, conducted the first prenatal XCI assay in a fetal heterozygous carrier of the whole *PCDH19*-deletion using uncultured amniocytes, detected the expression level of representative XCI escape genes in an adult female carrier of Xp-terminal deletion, and followed up or retrospected pregnancy outcomes of all female carriers. Our experiences provide useful references for prenatal risk assessment of heterozygous X-linked deletions and laid a foundation for further study of prognostic valuation of XCI pattern for these cases.

# MATERIALS AND METHODS

## Study Participants

This descriptive study focused on four independent Chinese pedigrees who underwent pre- or post-natal diagnosis (**Figure 1**). All pregnant women in the study got amniocentesis for chromosome micro-array (CMA) testing due to risky indications of non-invasive prenatal test (NIPT). The indication for a postnatal genetic test of pedigree four is a birth defect history in this family. Genetic tests revealed that all pedigrees contain more than one female carrier of heterozygous X-linked deletions. The common characteristic of these X-deletions is that the deletion region is above 1 Mb and contains at least one identified disease-causing gene that primarily affects heterozygous females. The study was approved by the Ethical Committee of Xinqiao Hospital, Chongqing, China.

## Molecular Diagnosis and Variant Classification

The pregnant women in this study underwent NIPT at 14–18 weeks of gestation for routine prenatal examination. The procedure was performed using the Ion Proton Sequencer (Life Technologies) according to the reported document (Yin et al., 2015). Amniocentesis was performed before 24 weeks of gestation following the ISUOG practice guidelines (Ghi et al., 2016) for indication of NIPT. Subsequent CMA using Affymetrix CytoScan HD chips was performed to detect fetal chromosome abnormality. Amniotic fluid fetal DNA was processed by sequential steps of digestion, polymerase chain reaction (PCR), PCR-product check, purification, quantization, fragmentation, and QC gel labeling, hybridization, washing, staining, and scanning according to the manufacturer's instruction. The data

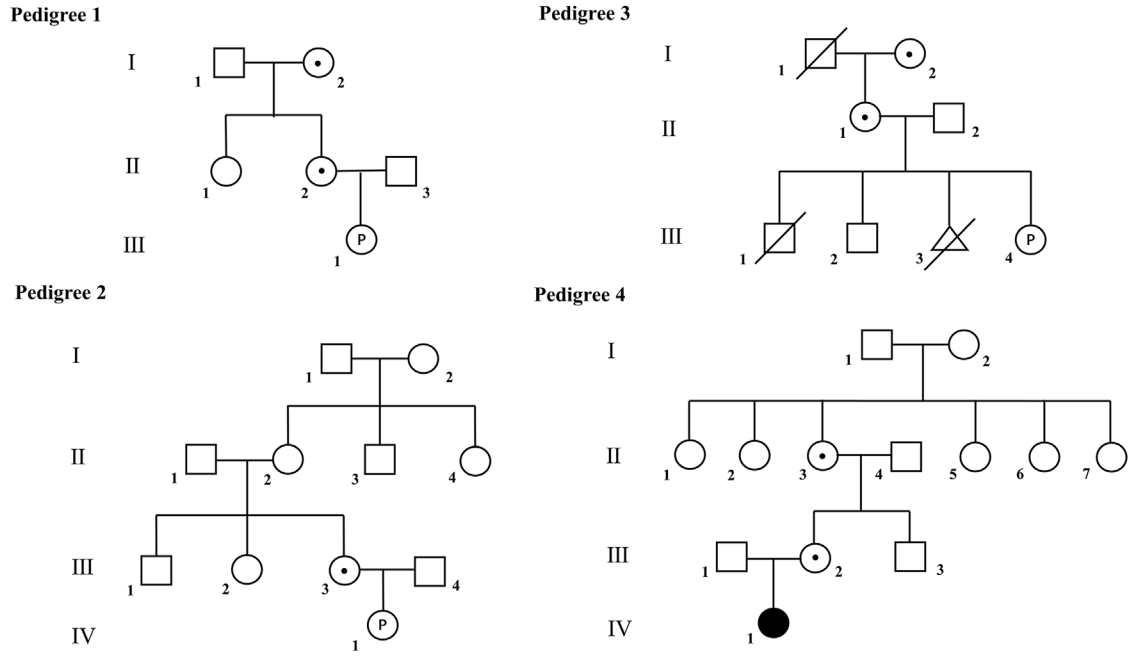



**FIGURE 1 |** Pedigrees of the families recruited in this study. Filled symbol indicates affected individuals; open symbols indicate unaffected individuals; open symbols with dot indicate asymptomatic carriers. Open symbols with letter "P" indicate the unborn fetus.

were analyzed using Affymetrix Chromosome Analysis Suite Software (version4.0). The interpretation of CNVs follows the guideline recommended by the American College of Medical Genomics (ACMG) and the Clinical Genome Resource (ClinGen) (Riggs et al., 2020).

Quantitative real-time PCR (Q-PCR) and peripheral blood karyotyping were used to pedigree verification. X-deletions below 5 Mb taken by pedigree 1 and 2, respectively, were confirmed by Q-PCR. In each case, three pairs of primer were designed to amplify the distinct coding regions of the deletion interval, respectively (**Supplementary Figure S1**). Standard 2-step-Q-PCR was performed using IQ$^{TM}$ SYBR Green Supermix (Bio-Rad, Hercules, CA, United States) on the CFX Connect Real-time PCR Detection System (Bio-Rad, Hercules, CA, United States). Each 20 μl assay was performed in triplicate and contained 20 ng of gDNA, 800 nM each of forward and reverse primers for the reference gene (RNaseP) or for the target genes. Relative quantity is determined by the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001), where a calibrator sample known to have two copies of the test sequence is used as the basis for comparative results. Gene copy number is 2 × the relative quantity. X-deletions above 10 Mb found in pedigree 3 and four were verified using karyotyping. Approximately 2 ml of peripheral blood of the fetus's parents and grandparents were collected to carry out G-banding analysis at a resolution of approximately 320–400 bands following standard laboratory protocols and ISCN 2016.

Trio-whole exome sequencing (WES) was carried out in pedigree 4. Exome targets were captured using the Agilent Inherited Disease Panel (Agilent Technologies, Inc.) and sequenced on the HiSeq 2,500 platform (Illumina, Inc.). Read alignment against the human reference genome builds hg19 was performed using the software NextGENe. Variants present at allele frequency greater than 0.5% in genomic databases (dbSNP138, 1,000 Genomes, ExAC) were disregarded. The software ANNOVAR was used to annotate variants. All relevant variants were visualized with Integrative Genomics Viewer (IGV) to exclude false results.

## X Chromosome Inactivation (XCI) Assay

The 5meCpG-based RP2/AR repeat biplex assay, a PCR assay that detects 5meCpG statuses of short tandem repeats (STR) alleles on the Xp (RP2) and the Xq (AR) in a single reaction, was employed to determine the XCI pattern (Machado et al., 2014). For each sample, 500 ng of gDNA was digested overnight with methylation-sensitive *Hp*aII restriction endonuclease (New England Biolabs, Ipswich, MA) at 37°C. Digested products together with mock-digested products (approximately 50 ng/sample) were used as templates for quantitative fluorescence polymerase chain biplex reactions (QF-PCR). PCR products were separated *via* capillary electrophoresis on an ABI 3130XL genetic analyzer (Applied Biosystems, Foster City, CA, United States). Allele profile and peak areas were analyzed by GeneMapper software (Applied Biosystems, Foster City, CA, United States). The informative allelic statuses were used to calculate the X chromosome inactivation ratio according to a previously published method (Jones, 2014). Interpretation of XCI pattern for a given sample was as follows: random (below 80:20), moderately skewed (80:20 to 90:10), and extremely skewed (above 90:10).

**TABLE 1 |** Summary of the X-linked deletions carried by four pedigrees.

| Pedigree number | Deletion position (hg19) and size | Primary disease-causing genes and related diseases | Inheritance pattern | Pathogenic curation |
|---|---|---|---|---|
| 1 | chrX:96821302-100378384, Xq21.33-Xq22.1 3.55 Mb | *PCDH19* (*300460) related early infantile epileptic encephalopathy (EIEE9, #300088) | Maternal | Pathogenic |
| 2 | chrX:47249368-50896523, Xp11.23-Xp11.22 3.65 Mb | *PORCN* (*300651) related focal dermal hypoplasia (#305600)<br>*EBP* (*300205) related X-linked dominant chondrodysplasia (#302960)<br>*SLC35A2* (*314375)-related congenital disorder of glycosylation, type II m (#300896)<br>*WDR45* (*300526)-related neurodegeneration with brain iron accumulation 5 (#300894) | Maternal | Pathogenic |
| 3 | chrX:178624-18404079, Xp22.33-Xp22.13 18.225 Mb | *SHOX* (*312865) related Leri-Weill dyschondrostosis (LWD, #127300)<br>*HCCS* (*300056) related linear skin defects with multiple congenital anomalies1 (#309801)<br>*OFD1* (*300170) related orofaciodigital syndrome1 (#311200)<br>*NHS* (*300457) related Nance-Horan syndrome (#302350) | Maternal | Pathogenic |
| 4 | chrX:129511205-142742928, Xq26.1-Xq27.3 13.23 Mb | *F9* (*300746) related hemophilia B (#306900)<br>*FRMD7* (*300628) related nystagmus-1 (#310700)<br>*GPC3*(*300037) related Simpson-Golabi-Behmel syndrome, type 1 (#312870)<br>*HPRT1*(*308000) related Lesch-Nyhan syndrome (#300322)<br>*PHF6*(*300414) related Borjeson-Forssman-Lehmann syndrome (#301900)<br>*SLC9A6* (*300231) related Christianson type of X-linked syndromic mental retardation (#300243) | Maternal | Pathogenic |

*Detailed genomic locations and contents (the haploinsufficiency genes primarily affecting female carriers) of the four distinct X-linked deletions were listed in the table. The inheritance patterns of these X-linked deletions were determined by the pedigree analysis presented in supplementary table. The pathogenic curation was made according to ACMG guideline (Riggs et al., 2020).*

## Gene Expression Assay

Given that the deletion region taken by pedigree 3 contains a large number of clinical relative XCI escape genes (Tukiainen et al., 2017), the XCI pattern alone is insufficient to explain the phenotype expression of the carriers, and the expression level of these genes is indispensable for comprehensive risk assessment for this family. Therefore, total RNA of peripheral blood lymphocytes from II-1 of pedigree 3 and four control peoples was extracted and reverse transcribed into cDNA to compare the expression level of representative genes that escape X-inactivation in pseudoautosomal region 1. Expression assay based on quantitative PCR was performed using IQ$^{TM}$ SYBR Green Supermix (Bio-Rad, Hercules, CA, United States) on the iQ-5 Real-time PCR Detection System (Bio-Rad, Hercules, CA, United States) as mentioned above. The expression level of a tested gene was normalized to β-actin mRNA. The relative level of mRNA was determined using the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001). Results are displayed as the mean ± SD from triplicate samples for each group. Significant differences were established by the one-way ANOVA followed by the Tukey-Kramer multiple comparisons post-test using the computer program GraphPad Prism (GraphPad Software Inc. V5.0, San Diego, CA, United States).

## Genetic Counselling and Follow-Up After Birth

All tested family members were told the results of molecular diagnosis and XCI pattern assay. The variant classification and the potential consequences for the fetal carriers or recurrency risk of next pregnancy have been clearly described. Women who choose to continue pregnancy were contacted by telephone to inquire their pregnancy outcomes.

# RESULTS

## Clinical Presentation and Molecular Diagnosis

### Pedigree 1

A 24-year-old pregnant woman (II-2), G1P0, referred for 15q11.2-15q13.1 deletion of the fetus (III-1) was indicated by NIPT. Amniocentesis followed by CMA preclude the suspected microdeletion on chr15 but uncovered a 3.557 Mb heterozygous deletion in the Xq21.33-Xq22.1 region of the X chromosome (chrX: 96821302-100378384, hg19), which involves a total of 11 OMIM genes (**Table 1**). Among them, *PCDH19* gene (MIM: 300460) has been documented as a haploinsufficiency gene (HI gene) in the ClinGene database. Loss-of-function of *PCDH19* can result in a rare early infantile epileptic encephalopathy (MIM: 300088, EIEE9), which almost exclusively occurs in heterozygous females. In the DECIPHER database, this interval embraces a likely-pathogenic document without phenotype description (ID: 300247). There was no evidence to refuse pathogenicity in the Database of Genomic Variants (DGV Gold standard variants). Taken together, this X-linked deletion should be assigned to pathogenic CNV. However, following pedigree analysis (summarized in **Supplementary Table**) demonstrating that

this pathogenic CNV inherited from the fetus's apparently healthy mother (II-2), both the pregnant woman and her mother (I-2) are heterozygous carriers of the same X-deletion (verified by Q-PCR and shown in **Supplementary Figure S1**).

### Pedigree 2

A 28-year-old pregnant woman (III-3), G1P0, underwent amniocentesis because NIPT indicated a 4.00 Mb deletion in the Xp11.23-Xp11.22 region of the fetus's X chromosome. Subsequent CMA performed on amniocytes confirmed that the fetus (IV-1) carried a heterozygous deletion of chrX: 47249368-50896523 (hg19) region. In DGV database, no CNV including this deletion has been recorded. In DECIPHER database, this region contains a pathogenic report described as a heterozygous carrier with intellectual disability (ID:322311). Assessment of genomic contents uncovered that this interval contains 72 OMIM genes, 7 of which have been curated haploinsufficiency. Among the 7 HI genes, 4 genes are relative to X-linked dominant diseases that primarily affect heterozygous females (summarized in **Table 1**). In this case, we already have enough compelling evidence to call this CNV pathogenic. However, evaluation of the inheritance pattern revealed that the deleted allele was inherited from the fetus's mother (demonstrated by CMA and Q-PCR) and the microdeletion is a *de novo* mutation for the apparently unaffected pregnant woman (III-3) because the Q-PCR verification showed that the pregnant woman's parents, sister, and maternal aunt do not carry the deletion (detailed information was shown in **Supplementary Table** and **Supplementary Figure S1**).

### Pedigree 3

A pregnant woman of 33 years old (II-1), G4P3, gave birth to two healthy boys in 2012 and 2015, respectively (the elder brother died in an accident) and terminated the third pregnancy for personal reasons in 2017. In the fourth pregnancy, NIPT and CMA uncovered an 18.225 Mb-heterozygous deletion spanning Xp22.33-Xp22.13 (chrX: 178624-18404079, hg19). There is no record that completely overlapped this interval in the population database. In DECIPHER, the region involves 2 types of syndromes: Leri-Weill dyschondrostosis (LWD) and Steroid sulphatase deficiency and numerous unique loss-pathogenic documents. The clinical phenotypes of these heterozygous cases involved short stature (ID:307967, 363886, 391541), autism and language development disorders (ID:338884), and global development delay (ID: 285776, 285002, 301496). In this region, 11 curated haploinsufficiency genes were covered, 4 of which primarily affect heterozygous women (summarized in **Table 1**). All of the evidence listed above are enough to demonstrate the pathogenic nature of the CNV. However, just like pedigree 1 and 2, the mutant allele is also maternally inherited. Karyotyping displayed that the mother (II-1) and the maternal grandmother (I-2) of the fetus present the same abnormal karyotype: 46, X,del (X) (p22.1) (summarized in **Supplementary Table**), but no severe manifestation was presented by the two carriers besides isolated short stature (both two carriers' height was 150 cm, −1.9 SDS; other features of LWD, such as limb shortening or abnormality of the forearms, were absent).

### Pedigree 4

A 30-year-old woman (III-2) asked for pregnancy counseling because she gave birth to a girl (IV-1) presented with global developmental delay and atrial septal defect in 2014. CMA performed on the girl uncovered a deleterious copy number variant: Xq26.1-q27.3 (129511205-142742928, hg19) × 1. The 13.2 Mb-deletion region contains 62 OMIM genes, 8 of them have been curated as HI genes (partial female-affected HI genes are summarized in **Table 1**). Although a vast majority of potential diseases caused by deletion of this region primarily affect hemizygous males, affected cases have been observed in heterozygous carriers of HI genes mentioned above. Actually, several loss-pathogenic reports involving heterozygous females in this region have been documented in DECIPHER (ID: 424282 and 318740 mainly involving *GPC3* gene; ID: 262468 and 402512 mainly involving *PHF6* gene). Whole or partial deletion of the *PHF6* gene (MIM: 300414) leads to an X-linked disorder called Borjeson-Forssman-Lehmann syndrome (BFLS, MIM: 301900), which usually affects males, but mild to severe symptoms are also present in female carriers. In our case, phenotypic features of the patient highly overlapped with the clinical presentation of previously reported female BFLS patients (Zweier et al., 2014), which includes developmental delay with severely affected speech and symbolic physical features (sparse hair, tapering, and fifth curved fingers/toes, skin hyperpigmentation, and narrow auditory canal, presented in **Supplementary Figure S2**). Subsequent pedigree investigation (summarized in **Supplementary Table**) uncovered that the unaffected mother (III-2) contains the same CNV as her sick daughter (indicated by CMA). The G-banding analysis performed on other family members (**Supplementary Table**) displayed that the maternal grandmother (II-3, unaffected) of the patient owns an abnormal karyotype: 46,X,de l(X) (q24q26). Trio-WES (including III-1, III-2, and IV-1) was also involved to preclude other potential disease-related single nucleotide variants or indels; no suspected variants could well explain the phenotypes of the patient. Taken together, we concluded that the deleterious allele causing BFLS is maternally inherited.

## X Chromosome Inactivation (XCI) Assay

Given XCI is a vital factor to modify the clinical consequences of female carriers of X-linked disorder, the XCI pattern of all female carriers of X-linked deletion and four non-carrier females in these pedigrees was detected using blood lymphocytes. Our results (**Table 2**) displayed that all female carriers present a completely skewed XCI pattern (100:0), while random XCI occurred in all tested non-carrier females. Among the female carriers with favorable XCI patterns against deleted alleles, the five-year-old girl with BFLS syndrome (in pedigree 4) was the only one who has obvious clinical manifestations. In pedigree 4, heterozygous carriers of Xq26.1-q27.3 deletion in successive three generations exhibited the same XCI pattern (extremely skewed XCI pattern −100% towards null allele, shown in **Table 2**) but completely different phenotypes: the mother (III-2) and the

**TABLE 2 |** The XCI patterns of the female relatives in four pedigrees.

| Pedigree | Family member | Blood DNA XCI ratio | | Most inactive chromosome | Origin of X-linked deletion |
|---|---|---|---|---|---|
| | | AR alleles | RP2 alleles | | |
| 1 | I-2‡ | n/i | 100:0 | n/p | n/p |
| | II-2‡ | 100:0 | n/i | Maternal | Maternal |
| | III-1‡ | 100:0 | n/i | Maternal | Maternal |
| 2 | II-2 | 79:21 | 67:33 | — | — |
| | II-4 | 69:31 | 63:37 | — | — |
| | III-2 | n/i | 51:48 | — | — |
| | III-3‡ | 100:0 | 100:0 | Paternal | *De novo* |
| | IV-1‡ | n/i | 100:0 | Maternal | Maternal |
| 3 | I-2‡ | 100:0 | 100:0 | n/p | n/p |
| | II-1‡ | 100:0 | 100:0 | Maternal | Maternal |
| | III-4‡ | 100:0 | 100:0 | Maternal | Maternal |
| 4 | II-1 | n/i | 77:23 | — | — |
| | II-3‡ | n/i | 100:0 | n/p | n/p |
| | III-2‡ | 100:0 | 100:0 | Maternal | Maternal |
| | IV-1‡ | 100:0 | n/i | Maternal | Maternal |

*X chromosome inactivation (XCI) assay based on 5meCpG statuses of STR alleles on the Xp (RP2) and the Xq (AR) were performed in 11 heterozygous carriers of X-linked deletions (indicated by symbol ‡ in column 2) and four non-carrier females in four pedigrees. The informative allelic statuses (n/i: non-informative allele) were used to calculate the X chromosome inactivation ratio (described in "materials and methods"). The XCI ratio below 80:20 is interpreted into a "random inactivation pattern", while the XCI ratio above 90:10 means "extremely skewed inactivation". The most inactivated X chromosome of a given sample was determined by trio-STR genotyping (n/p: the trio-STR genotyping is not performed). The origin of X-deletion of a given sample was determined by Q-PCR or karyotyping presented in supplementary table.*

maternal grandmother II-3) are asymptomatic, while the proband present obvious manifestations of BFLS. This phenomenon implied that the XCI pattern of blood cell could not be a reliable indicator of the clinical consequences of this disease.

Moreover, in pedigree 1, which contains a fetus having a risk of *PCDH19*-deletion disorder, we detected the fetus's XCI pattern in prenatal using uncultured amniocytes. The result revealed that the female fetus presents a 100% XCI skewing against the maternal (origin of deletion) X-chromosome, which implies a positive prognosis. Furthermore, the XCI pattern of amniocytes was consistent with that of neonatal cord blood lymphocytes detected after birth (**Figure 2**), indicating the feasibility of XCI assay in amniocytes.

## The Expression Level of XCI Escapees Located in Deletion Region of Pedigree 3

It is known that the majority of escapees of XCI are located on the Xp, especially highly enriched in the pseudoautosomal region 1 (PAR1). The 18.225 Mb deletion taken by pedigree 3 contains whole PAR1 and adjacent escapee strata, including 44 defined XCI escape genes (Tukiainen et al., 2017), which account for half of the total number of genes in this region. Theoretically, irrespective of the deletions of escape genes that occurred in the active X chromosome (Xa) or inactive X chromosome (Xi), they could be harmful. However, our study revealed that in comparison to four race, age, and sex-matched controls, the copy number alteration at the genomic level of several potential disease-causing escape genes in our female carrier did not result in significant dosage insufficiency at the transcriptional level (**Figure 3**), which implied that potential upregulation of escapees from intact X chromosome could equalize the gene dosage between bi-allelic female and structurally mono-allelic female.

## Genetic Counseling and Clinical Outcome Follow-Ups

For the pedigrees with fetal carriers of X-deletions, comprehensive prognostic evaluation was informed to the pregnant women. After a full consultation, all of them decided to continue their pregnancies. Follow-up so far showed that the baby girl with a risk of EIEE9 in pedigree 1 has not suffered from seizures during the 16 months after birth (the average onset age of EIEE9 is 14 months (Samanta, 2020)); the baby girl harboring a heterozygous deletion related to multisystem anomalies and intellectual disability (in pedigree 2) has no other distinguishing features at 10 months after birth, aside from many café au lait spots distributed to the whole body (a retrospective survey found that similar pigmentation of skin also occurred in her "non-phenotypic" carrier mother); the baby girl (in pedigree 3) with deletion of numerous XCI escapees including *SHOX* presented with normal growth milestones and limb length during the first year after birth.

We assessed the recurrence risk for the couples (III-1, III-2 in pedigree 4) who gave birth to a BFLS girl. In their every pregnancy, there would be a 50% chance to transmit the deleted-*PHF6* allele to their offspring, irrespective of the sex. Males who inherited the deleted allele would 100% present full-blown BFLS phenotypes, and even for female heterozygous, the penetrance is estimated up to 44% (Zhang et al., 2019). More importantly, our study and quite a bit of previous cases showed that favorable skewed XCI is usually insufficient to prevent harmful outcomes for this disease (Crawford et al., 2006; Berland et al., 2011; Zhang et al., 2019). After careful consideration, the couples chose preimplantation genetic

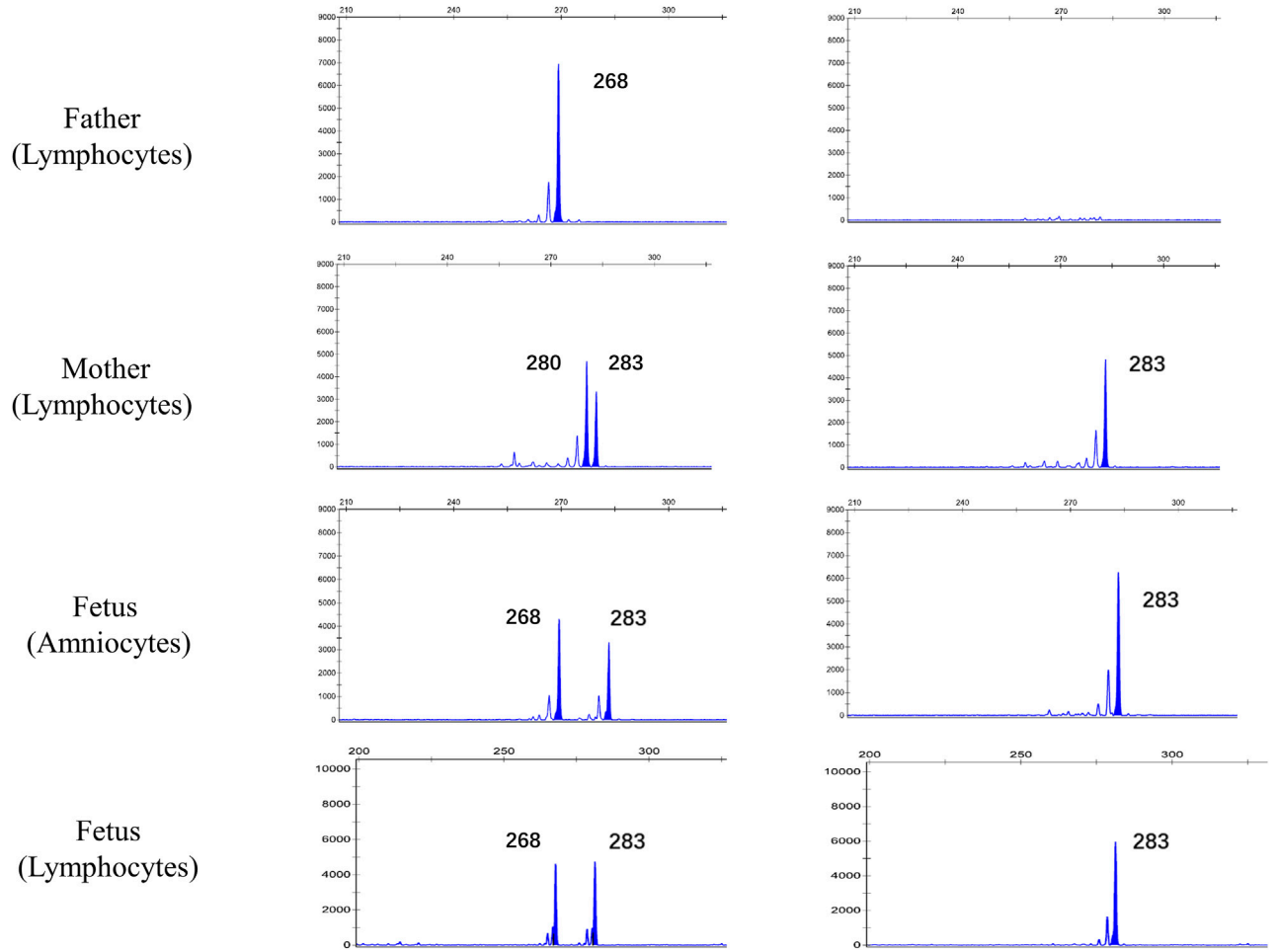



**FIGURE 2 |** Prenatal XCI assay performed in the female fetus of pedigree 1. gDNA extracted from fresh uncultured amniocytes of the fetus and peripheral blood lymphocytes of her parents was used to detect 5meCpG statuses of *AR* and *RP2* alleles. The peaks indicate the informative *AR* alleles. The size of each allele is determined by the length of polymorphic microsatellite in the exon1 of *AR* gene. The allele$^{268}$ and allele$^{283}$ of the fetus is inherited from her father and mother, respectively. After digestion by methylation-sensitive HpaII restriction endonuclease, bi-allelic status of the mother and fetus's AR gene converts to a mono-allelic status, which means that both of them possess completely skewed XCI pattern and the remaining alleles represent the alleles that have been inactivated (methylated). Thus, for the fetus, skewed XCI towards allele$^{283}$ of her, which indicates the maternal-original X chromosome (the one taking microdeletion), has been inactivated. The XCI pattern of amniocytes is perfectly replicated in neonatal cord blood lymphocytes.

diagnosis (PGD) to prevent the transmission of the pathogenic CNV.

## DISCUSSION

In this study, we reported four pedigrees carrying different heterozygous X-linked deletions (>1 Mb), which were discovered incidentally through prenatal screening (their unborn fetuses were indicated as X chromosome deletion by NIPT) or diagnosed during the retrospection of a birth defect history (pedigree 4). According to the ACMG interpretation framework (Riggs et al., 2020), all microdeletions should be assigned to pathogenic CNV for heterozygous carriers. However, the fact that these CNV are inherited from the unaffected mothers complicates the valuation of their clinical implications. Given several X-linked disorders including X-deletion have been reported to be associated with skewed XCI and normal phenotype in heterozygous carriers (Orstavik, 2009), we set out to determine the XCI pattern of female relatives in our pedigrees at first. We found that a total of 11 X-deletion female carriers 100% presented completely skewed XCI pattern against abnormal X-chromosome, while the prevalence of XCI skewing in non-carrier females is zero (4/4 random XCI), which seems highly consistent with the principle that the unbalanced structurally abnormal X chromosome is preferential to be inactivated. However, limited by the small number and types of cases in our study, a definitive conclusion about the

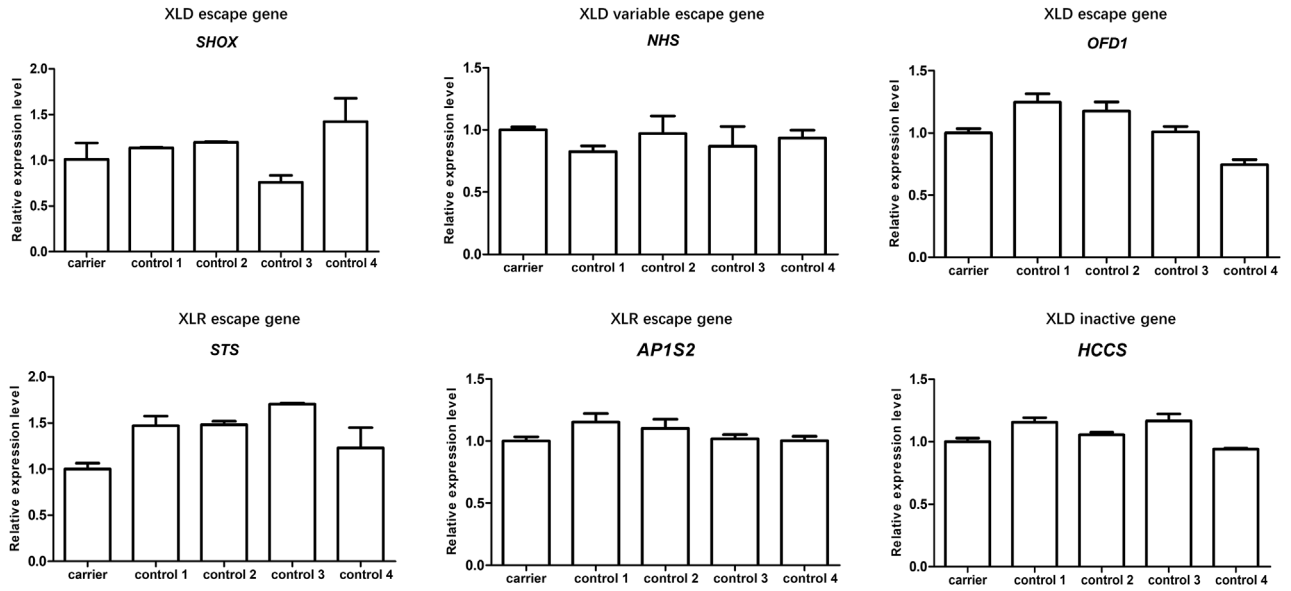



**FIGURE 3 |** The expression levels of representative genes involved in the deletion region of pedigree 3. Total RNA was extracted from peripheral blood lymphocytes of one female carrier of Xp22.33-Xp22.13 deletion (II-1in of pedigree 3) and four race- and age-matched non-carrier female controls. qRT-PCR was used to detect the expression of six representative genes located in the deletion region in these samples. The inheritance mode and reported XCI status of these gene have been annotated in the figure. The relative levels of these genes are normalized to β-actin expression. Data represents the mean of experiment performed in triplicate; there is no significant difference among the tested samples.

relationship between X-structural abnormality and XCI pattern cannot be drawn just based on this observation. For instance, our study did not show whether the X-microdeletion below 1 Mb would also lead to selective XCI skewing and we also noticed that certain X-linked CNV (such as Xp11.23-p11.22 duplication) exhibits an unusual skewed XCI pattern that is in favor of the aberrant X chromosome (Di-Battista et al., 2016). Therefore, we proposed that the skewed XCI associated with X chromosome structural abnormality highly depends on the size and particular genomic contents of the variant region. Systemic investigations focused on the XCI patterns in various X chromosome structural anomalies will make a better understanding of the correlation between them.

More importantly, even though the favorable skewed XCI against abnormal X chromosome occurred, does that definitely indicate a positive prognosis? Our experiences suggest that the prognostic value of the XCI pattern is inconsistent in disorders with different genetic foundations. For example, in our cases, skewed XCI seems a positive prognostic indicator of *PCDH19*-deletion (3 asymptomatic heterozygous carriers of *PCDH19*-deletion with completely skewed XCI in pedigree 1), due to the special "mosaic pathogenesis" of PC*DH19*-disorder (Dibbens et al., 2011; Dimova et al., 2012; Terracciano et al., 2012; Hoshina et al., 2021). *PCDH19* gene encodes a single transmembrane protein involved in calcium-dependent cell-cell interaction and adhesion. Cellular interference between wild type-*PCDH19* neurons and mutant/deleted-*PCDH19* neurons is the main pathogenic mechanism underlying the disorder (Hoshina et al., 2021). Therefore, the heterozygous females with extremely skewed XCI always remain spared due to a low level of somatic mosaicism of this gene. However, for some other X-linked disorders, such as BFLS induced by *PHF6* gene mutation, our case (in pedigree 4) and numerous previous reports all uncovered that the XCI pattern of blood cell has poor prognostic value for clinical consequences (Crawford et al., 2006; Berland et al., 2011; Zhang et al., 2019). The reasonable explanation could be that unlike *PCDH19*, expression of which peaks in the first postnatal period (Hoshina et al., 2021), *PHF6* is primarily expressed in an early stage of fetal brain development (Jahani-Asl et al., 2016). At that time, the concrete ratio and distribution of the two populations of cells (expressing normal or mutant *PHF6*, respectively) are stochastic and almost impossible to be accurately reflected by an XCI ratio established in peripheral blood cells after birth because the XCI ratios of tissues in different development stages are inconsistent. Generally, the more mitosis experienced, the more skewed the XCI ratio presented due to clonal selection (Minks et al., 2008). Therefore, employing the XCI assay in the risk prediction of certain X-linked disorders must take into account the particular pathogenic mechanism of the disease and the representativity of the alternative tested tissues.

When we discuss the relationship between XCI and phenotypes of X-linked disorder, we cannot ignore the presence of a considerable number of XCI escapees. In pedigree 3, a three-generation family takes a heterozygous deletion located in the most enriched region of XCI escapees (Tukiainen et al., 2017). This "escapee cluster" encompasses several X-linked dominant (XLD) escape genes either with high penetrance (such as *SHOX*, MIM: 312865, haploinsufficiency of which resulted in short stature in varying

degree) or with severe phenotypes (such as *OFD1*, MIM: 311200, mutation of which could be lethal in males). Theoretically, regardless of the XCI pattern, the dosage imbalance of these important XLD escape genes at the genomic level would result in severe phenotypes. Surprisingly, there are no significant clinical symptoms presented in our carriers. Karyotyping has ruled out post-zygotic mosaicism involving the concomitant presence of normal and X-deletion cells; thus, we compared the expression level of six representative genes involved in the deletion between one of our carriers (II-1 of pedigree 3) and four race- and age-matched non-carrier females. The nearly equalized expression of all XCI escapees among them exactly coincides with a scenario reported by Santos-Reboucas et al. recently, who observed a skewed XCI and compensatory upregulation of escape genes in an asymptomatic female with a *de novo* heterozygous deletion (32 Mb) containing Xq-escapee cluster (Santos-Reboucas et al., 2020). Both two cases suggested that large deletion of the X chromosome could induce a potential dosage compensation mechanism to balance the expression of XCI escapees between bi-allelic females and structurally hemizygous females. However, due to few related cases having been reported so far, the universality and molecular mechanism of the dosage compensation effect merit further investigation.

As we mentioned above, the XCI pattern is significantly implicated in the phenotype expression of *PCDH19*-disorder. Nevertheless, the clinical application of XCI assay in the risk prediction of the disease has not been reported. The free cells from amniotic fluid contain multi-organ cells of fetal, such as skin, gastrointestinal, and urinary tract cells, in theory, should be more appropriate for XCI assay than most used blood cells. Moreover, it has been demonstrated that the XCI pattern of amniocytes has been established by the time of routine prenatal diagnosis and used to predict the risk of dystrophinopathy in prenatal (He et al., 2019). Therefore, in this study, we performed the first prenatal XCI assay using uncultured amniocytes to evaluate the prognosis of a fetal heterozygous carrier of *PCDH19*-deletion. Although the validity of this analysis is needed to be assessed by further studies with a large number of cases, our study makes the first step towards the prenatal risk prediction of *PCDH19*-disorder using the XCI pattern.

In conclusion, the female carriers with heterozygous X-linked deletion (>1 Mb) would undergo XCI skewing. The favorable skewed XCI in combination with potential compensatory upregulation of XCI escapees would protect a part of female carriers with pathogenic X-deletion from severe clinical consequences. The prognostic value of the XCI pattern depends on particular genomic contents of the deletion region and the choice of appropriate tested tissue. For certain diseases like loss-of-function of the *PCDH19* gene, the XCI pattern could be a useful prognostic indicator. XCI analysis employing uncultured amniocytes might be a practicable way for prognostic evaluation of *PCDH19*-disorder in prenatal. The most shortcoming of this study is the limited case number and type; thus, clinicians should be cautious when applying these conclusions in clinical practice, while we still believe that our experiences will provide useful references for prenatal risk assessment of heterozygous X-linked deletions and laid a foundation for further systemic studies of prognostic valuation of XCI pattern for these cases.

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation. The CMA and karyotype data analyzed in the study was obtained from Chongqing health center for women and children. According to the Regulations on Management of Human Genetic Resources in China, the domestic prenatal genetic data cannot be copied or uploaded. Requests to access these datasets should be directed to the prenatal diagnostic center.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethical Committee of Army Medical University. 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

HG and YZ were responsible for project conception. JL, YM, and HG were responsible for recruiting subjects and gathering patient history, clinical information, and written informed consent. YZ performed XCI assay and gene expression assay. JL and LD performed Sanger sequencing and data analysis. YM performed amniocentesis. YZ analyzed the data and drafted the manuscript. The manuscript was critically revised by HG and accepted by all co-authors.

## FUNDING

The study was funded by the National Natural Science Foundation of China (No. 82171840) and the Natural Science Foundation Project of Chongqing (cstc2021jcyj-msxmX0281).

## ACKNOWLEDGMENTS

We thank all families and voluntary subjects for their understanding and cooperation. We thank Xintong Zhu, Junwen Wang, Mao Zhang, and Xinying Guan for the technical help and discussion.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.782629/full#supplementary-material

## REFERENCES


Berland, S., Alme, K., Brendehaug, A., Houge, G., and Hovland, R. (2011). PHF6 Deletions May Cause Borjeson-Forssman-Lehmann Syndrome in Females. *Mol. Syndromol.* 1 (6), 294–300. doi:10.1159/000330111

Brown, C., and Robinson, W. (2000). The Causes and Consequences of Random and Non-random X Chromosome Inactivation in Humans. *Clin. Genet.* 58 (5), 353–363. doi:10.1034/j.1399-0004.2000.580504.x

Crawford, J., Lower, K. M., Hennekam, R. C., Van Esch, H., Megarbane, A., Lynch, S. A., et al. (2006). Mutation Screening in Borjeson-Forssman-Lehmann Syndrome: Identification of a Novel De Novo PHF6 Mutation in a Female Patient. *J. Med. Genet.* 43 (3), 238–243. doi:10.1136/jmg.2005.033084

Di-Battista, A., Meloni, V. A., da Silva, M. D., Moysés-Oliveira, M., and Melaragno, M. I. (2016). Unusual X-Chromosome Inactivation Pattern in Patients with Xp11.23-p11.22 Duplication: Report and Review. *Am. J. Med. Genet.* 170 (12), 3271–3275. doi:10.1002/ajmg.a.37888

Dibbens, L. M., Kneen, R., Bayly, M. A., Heron, S. E., Arsov, T., Damiano, J. A., et al. (2011). Recurrence Risk of Epilepsy and Mental Retardation in Females Due to Parental Mosaicism of PCDH19 Mutations. *Neurology* 76 (17), 1514–1519. doi:10.1212/WNL.0b013e318217e7b6

Dimova, P. S., Kirov, A., Todorova, A., Todorov, T., and Mitev, V. (2012). A Novel PCDH19 Mutation Inherited from an Unaffected Mother. *Pediatr. Neurol.* 46 (6), 397–400. doi:10.1016/j.pediatrneurol.2012.03.004

Disteche, C. M. (1999). Escapees on the X Chromosome. *Proc. Natl. Acad. Sci.* 96 (25), 14180–14182. doi:10.1073/pnas.96.25.14180

Gendrel, A.-V., and Heard, E. (2011). Fifty Years of X-Inactivation Research. *Development* 138 (23), 5049–5055. doi:10.1242/dev.068320

Ghi, T., Sotiriadis, A., Calda, P., Da Silva Costa, F., Raine-Fenning, N., Alfirevic, Z., et al. (2016). ISUOG Practice Guidelines: Invasive Procedures for Prenatal Diagnosis. *Ultrasound Obstet. Gynecol.* 48 (2), 256–268. doi:10.1002/uog.15945

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. *Prenatal Diagn.* 39 (8), 603–608. doi:10.1002/pd.5473

Hoshina, N., Johnson-Venkatesh, E. M., Hoshina, M., and Umemori, H. (2021). Female-specific Synaptic Dysfunction and Cognitive Impairment in a Mouse Model of PCDH19 Disorder. *Science* 372 (6539), eaaz3893. doi:10.1126/science.aaz3893

Jahani-Asl, A., Cheng, C., Zhang, C., and Bonni, A. (2016). Pathogenesis of Börjeson-Forssman-Lehmann Syndrome: Insights from PHF6 Function. *Neurobiol. Dis.* 96, 227–235. doi:10.1016/j.nbd.2016.09.011

Jones, J. R. (2014). Nonrandom X Chromosome Inactivation Detection. *Curr. Protoc. Hum. Genet.* 80, Unit 9 7. doi:10.1002/0471142905.hg0907s80

Leppig, K. A., and Disteche, C. M. (2001). Ring X and Other Structural X Chromosome Abnormalities: X Inactivation and Phenotype. *Semin. Reprod. Med.* 19 (2), 147–158. doi:10.1055/s-2001-15395

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method. *Methods* 25 (4), 402–408. doi:10.1006/meth.2001.1262

Lyon, M. F. (1961). Gene Action in the X-Chromosome of the Mouse (*Mus musculus* L.). *Nature* 190, 372–373. doi:10.1038/190372a0

Lyon, M. (2002). X-chromosome Inactivation and Human Genetic Disease. *Acta Paediatr. Suppl.* 91 (439), 107–112. doi:10.1111/j.1651-2227.2002.tb03120.x

Machado, F. B., Machado, F. B., Faria, M. A., Lovatel, V. L., Alves da Silva, A. F., Radic, C. P., et al. (2014). 5meCpG Epigenetic marks Neighboring a Primate-Conserved Core Promoter Short Tandem Repeat Indicate X-Chromosome Inactivation. *PLoS One* 9 (7), e103714. doi:10.1371/journal.pone.0103714

Minks, J., Robinson, W. P., and Brown, C. J. (2008). A Skewed View of X Chromosome Inactivation. *J. Clin. Invest.* 118 (1), 20–23. doi:10.1172/JCI34470

Ørstavik, K. H. (2009). X Chromosome Inactivation in Clinical Practice. *Hum. Genet.* 126 (3), 363–373. doi:10.1007/s00439-009-0670-5

Puck, J. M., and Willard, H. F. (1998). X Inactivation in Females with X-Linked Disease. *N. Engl. J. Med.* 338 (5), 325–328. doi:10.1056/NEJM199801293380611

Rastan, S. (1982). Primary Non-random X-Inactivation Caused by Controlling Elements in the Mouse Demonstrated at the Cellular Level. *Genet. Res.* 40 (2), 139–147. doi:10.1017/s0016672300019017

Riggs, E. R., Andersen, E. F., Cherry, A. M., Kantarci, S., Kearney, H., Patel, A., et al. (2020). Technical Standards for the Interpretation and Reporting of Constitutional Copy-Number Variants: a Joint Consensus Recommendation of the American College of Medical Genetics and Genomics (ACMG) and the Clinical Genome Resource (ClinGen). *Genet. Med.* 22 (2), 245–257. doi:10.1038/s41436-019-0686-8

Samanta, D. (2020). PCDH19-Related Epilepsy Syndrome: A Comprehensive Clinical Review. *Pediatr. Neurol.* 105, 3–9. doi:10.1016/j.pediatrneurol.2019.10.009

Santos-Rebouças, C. B., Boy, R., Vianna, E. Q., Gonçalves, A. P., Piergiorge, R. M., Abdala, B. B., et al. (2020). Skewed X-Chromosome Inactivation and Compensatory Upregulation of Escape Genes Precludes Major Clinical Symptoms in a Female with a Large Xq Deletion. *Front. Genet.* 11, 101. doi:10.3389/fgene.2020.00101

Schwartz, C. E. (2013). *Brenner's Encyclopedia of Genetics*. Second Edition. Elsevier.

Terracciano, A., Specchio, N., Darra, F., Sferra, A., Bernardina, B. D., Vigevano, F., et al. (2012). Somatic Mosaicism of PCDH19 Mutation in a Family with Low-Penetrance EFMR. *Neurogenetics* 13 (4), 341–345. doi:10.1007/s10048-012-0342-9

Tukiainen, T., Villani, A. C., Villani, A.-C., Yen, A., Rivas, M. A., Marshall, J. L., et al. (2017). Landscape of X Chromosome Inactivation across Human Tissues. *Nature* 550 (7675), 244–248. doi:10.1038/nature24265

Yin, A.-h., Peng, C.-f., Zhao, X., Caughey, B. A., Yang, J.-x., Liu, J., et al. (2015). Noninvasive Detection of Fetal Subchromosomal Abnormalities by Semiconductor Sequencing of Maternal Plasma DNA. *Proc. Natl. Acad. Sci. USA* 112 (47), 14670–14675. doi:10.1073/pnas.1518151112

Yuan, H., Shangguan, S., Li, Z., Luo, J., Su, J., Yao, R., et al. (2021). CNV Profiles of Chinese Pediatric Patients with Developmental Disorders. *Genet. Med.* 23 (4), 669–678. doi:10.1038/s41436-020-01048-y

Zhang, X., Fan, Y., Liu, X., Zhu, M.-A., Sun, Y., Yan, H., et al. (2019). A Novel Nonsense Mutation of PHF6 in a Female with Extended Phenotypes of Borjeson-Forssman-Lehmann Syndrome. *J. Clin. Res. Pediatr. Endocrinol.* 11 (4), 419–425. doi:10.4274/jcrpe.galenos.2019.2018.0220

Zweier, C., Rittinger, O., Bader, I., Berland, S., Cole, T., Degenhardt, F., et al. (2014). Females with De Novo Aberrations inPHF6: Clinical Overlap of Borjeson-Forssman-Lehmann with Coffin-Siris Syndrome. *Am. J. Med. Genet.* 166 (3), 290–301. doi:10.1002/ajmg.c.31408


**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.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

*Copyright © 2021 Zhao, Li, Dai, Ma, Bai and Guo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*

frontiers in Genetics

ORIGINAL RESEARCH
published: 18 January 2022
doi: 10.3389/fgene.2021.783455

# A Novel Heterozygous Pathogenic Variation in *CYCS* Gene Cause Autosomal Dominant Non-Syndromic Thrombocytopenia 4 in a Large Chinese Family

*Fengyu Che*[1], *Jiangang Zhao*[2], *Yujuan Zhao*[2], *Zhi Wang*[2], *Liyu Zhang*[1] *and Ying Yang*[1]*

[1]*Shaanxi Institute for Pediatric Diseases, Xi'an Children's Hospital, Xi'an, China,* [2]*Department of Neonatology, Xi'an Children's Hospital, Xi'an, China*

**Aim:** To determine the etiology of a Chinese family with thrombocytopenia by analyzing the clinical features and genetic variation.

**Methods:** Clinical profiles and genomic DNA extracts of the family members were collected for the study. Whole exome sequencing and Sanger sequencing was used to detect the associated genetic variation and verify the family co-segregation respectively. Bioinformatics analysis assessed the pathogenicity of missense mutations.

**Results:** The study reported a 3-generation pedigree including eight family members with thrombocytopenia. The platelet counts of the patients were varied, ranging from 38 to 110 × $10^9$/L (reference range: 150–450 x $10^9$/L). The mean volumes and morphology of the sampled platelet were both normal. The bleeding abnormality and mitochondriopathy were not observed in all the patients. Clinical signs of thrombocytopenia were mild. A novel heterozygous missense variant c.79C > T (p.His27Tyr) was identified in *CYCS* gene associated with autosomal dominant thrombocytopenia.

**Conclusion:** We report the first large family with autosomal dominant non-syndromic thrombocytopenia 4 in a Chinese family, a novel heterozygous missense variant c.79C > T (p.His27Tyr) was identified. The whole exome sequencing is an efficient tool for screening the variants specifically associated with the disease. The finding enriches the mutation spectrum of *CYCS* gene and laid a foundation for future studies on the correlation between genotype and phenotype.

**Keywords: CYCS gene, platelet counts, cytochrome c, mutation, thrombocytopenia**

OPEN ACCESS

***Edited by:***
*Kathleen M. Gorman, Temple Street Children's University Hospital, Ireland*

***Reviewed by:***
*Stefano Castellana, Casa Sollievo della Sofferenza (IRCCS), Italy*
*Alejandro Ferrer, Mayo Clinic, United States*

***Correspondence:**
*Ying Yang*
*yying1930@163.com*

***Specialty section:***
*This article was submitted to Genetics of Common and Rare Diseases, a section of the journal Frontiers in Genetics*

***Received:*** *26 September 2021*
***Accepted:*** *06 December 2021*
***Published:*** *18 January 2022*

***Citation:***
*Che F, Zhao J, Zhao Y, Wang Z, Zhang L and Yang Y (2022) A Novel Heterozygous Pathogenic Variation in CYCS Gene Cause Autosomal Dominant Non-Syndromic Thrombocytopenia 4 in a Large Chinese Family. Front. Genet. 12:783455. doi: 10.3389/fgene.2021.783455*

## 1 INTRODUCTION

Thrombocytopenia is defined as a haematological condition with blood platelet count below 100–150 × $10^9$/L (Greenberg and Kaled, 2013). It is correlated with inherited hematologic diseases, immune diseases (e.g.rheumatoid), radiation/chemotherapy injuries, infection, and drug-related thrombocytopenia (Lee and Lee, 2016; Franchini et al., 2017). At present, more than 20 genes have been reported to be associated with the incidence of inherited thrombocytopenia, including

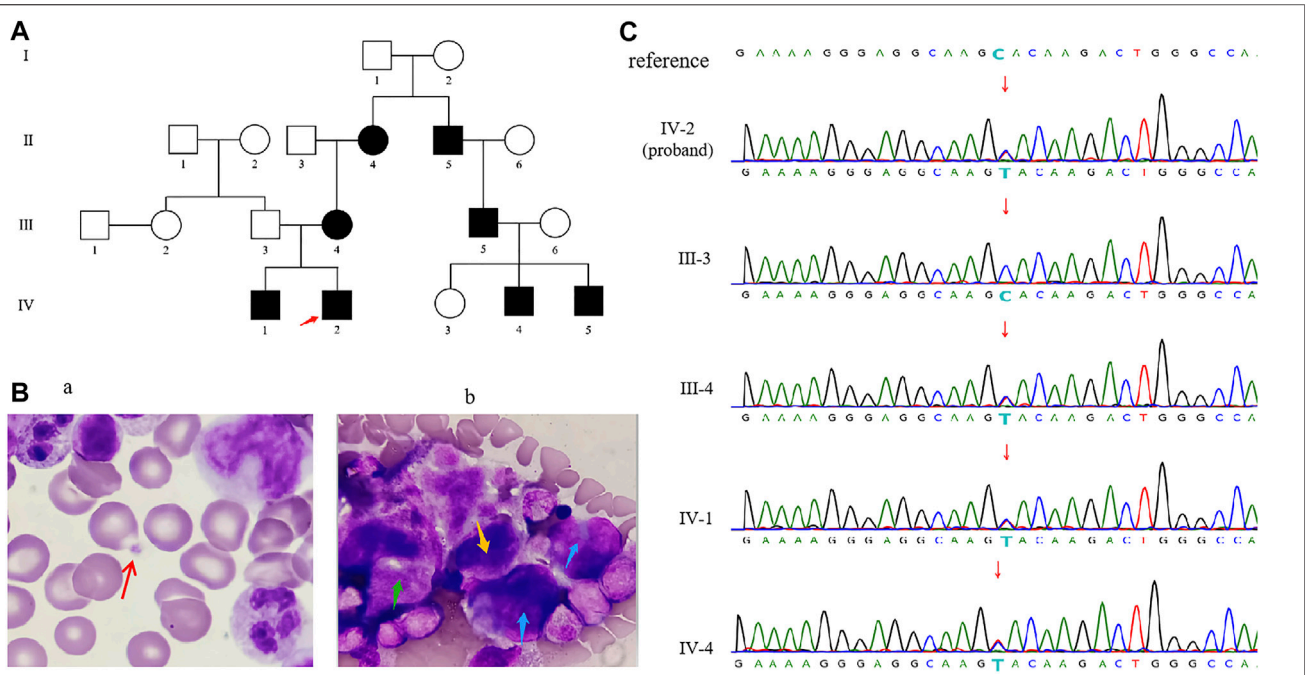



**FIGURE 1 | (A)** Familial pedigree of CYCS-mutated thrombocytopenia. The red arrow indicates the proband. **(B)** Optical microscopy of bone marrow cells from the proband showed the number of platelets were reduced. The red arrow on the left **(Ba)** showed a single platelet was sporadically visible, and the blue arrow points to promegakaryocytes **(Bb)**, the green arrow points to granular megakaryocytes **(Bb)**, and the yellow arrow points to thromocytogenic megakaryocytes **(Bb)**. **(C)** Sanger sequencing of *CYCS* gene c.79C > T (p.His27Tyr) variant in genomic DNA from the family.

syndromic, and non-syndromic (Nurden and Nurden, 2020). Hence, understanding the disease from a genetic perspective is a key of providing effective diagnosis and prognostic risk assessment.

The *CYCS* gene encodes the Cytochrome C (Cyt-c), which is a small and stable heme protein with heme C as its auxiliary group (Stevens, 2011). The Cyt-c protein is an essential component anchoring in the inner membrane of the mitochondrion for multiple bio-functions (Santucci et al., 2019). It is mainly responsible of transferring electrons from cytochrome b to the cytochrome oxidase complex (Kagan et al., 2005; Li et al., 2006; Ow et al., 2008), and also could initiate cell apoptosis as an antioxidant agent (Wang et al., 2003; Zhao et al., 2003; Lee and Xu, 2007). It is reported that the heterozygous mutation of *CYCS* gene on chromosome 7p15 can caused autosomal dominant non-syndromic thrombocytopenia-4 (THC4, OMIM: 612004) (Morison et al., 2008). The clinical features of THC4 were characterized with mild thrombocytopenia, normal platelet size and morphology, and no increased bleeding tendency. So far, there are only 4 pedigrees with nonsyndrome thrombocytopenia and 1 pedigree with hemophilia A, associated with variants in *CYCS* gene (Morison et al., 2008; De Rocco et al., 2014; Johnson et al., 2016; Uchiyama et al., 2018; Turro et al., 2020). None of the patients in these reports showed symptoms resulting from cell apoptosis and abnormal mitochondrial oxidative respiratory chain, whereas thrombocytopenia was present.

We studied a first pedigree with autosomal dominant non-syndrome thrombocytopenia 4 in a Chinese family. All patients in the family were found to carry a heterozygotic missense variant in *CYCS* gene by whole exome sequencing (WES), while the pathogenicity of the genetic variation was also evaluated. According to the clinical and genetic features of the patients in the pedigree, thrombocytopenia 4 was diagnosed. This work could be a scientific evidence to support on thrombocytopenia diagnosis and its prognostic management.

## 2 MATERIALS AND METHODS

### 2.1 Patient Clinical Information

The probands (IV-2) was the second son of a non-consanguineous Chinese couple (**Figure 1A**). He was born at gestational age of 39 + 5 weeks by spontaneous vaginal delivery with normal birth history (birth weight, 3500 g). Multiple hemorrhagic spots were found on the facial and back skin of the proband at birth. After 7 days, he was admitted for further treatment. Physical examination was normal. Brain magnetic resonance imaging (MRI) showed mild cerebral hemorrhage in the left ventricle and focal white matter injury near the posterior horn of the left ventricle. The full blood count test yielded a low platelet count at $54 \times 10^9$/L (reference range: 150–450 $\times 10^9$/L) and the increased leucocyte count at $30.68 \times 10^9$/L (reference range: 10.4–12.21 $\times 10^9$/L). After antibiotics and immunoglobulin treatment, the hemorrhagic spots disappeared with normal leucocyte counts and improved platelet counts (ranges from 56 to 96 $\times 10^9$/L).

**TABLE 1 |** Clinical features of the family members from the current study.

| Case clinical features | IV-2 (proband) | IV-1 | III-4 | II-4 | II-5 | III-5 | IV-4 | IV-5 | III-3 | III-6 | IV-3 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Brusied more than normal | – | + | + | + | + | – | – | – | – | – | – |
| Epistaxis | – | – | – | – | – | – | – | + | – | – | – |
| Bleeding spot | + | – | – | – | – | – | – | – | – | – | – |
| Mitchondial disease | – | – | – | – | – | – | – | – | – | – | – |
| Platelet counts 150–450 × $10^9$/L) | 54–108 | 89–110 | 41–45 | 50–70 | 60 | 38 | 98 | 76 | 200 | 205 | 217 |
| MPV | N | N | N | N | N | N | N | N | N | N | N |
| Variant | *CYCS*: NM_018947.6:c.79C > T (p.His27Tyr), heterozygous | | | | | | | | Wild-type | | |

*N, normal; MPV, mean platelet volumes.*

At 16 days of age, the patient showed no signs of hemorrhage with normal abdominal ultrasound and urine organic acids. Blood tandem mass spectrometry analysis showed camitine deficiency. Optical microscope of bone marrow showed megaloblasts and toxic granules below the stage of myelocyte and metamyelocyte, indicating as bone marrow infection. In addition, the whole bone marrow smear showed 550 megakaryocytes, among which promegakaryocytes accounted for 30%, granular megakaryocytes accounted for 62%, thromocytogenic megakaryocytes accounted for 8%. These abnormal proportions of megakaryocytes indicate that megakaryocytes maturation problem leaded to a decrease in platelet count (**Figure 1B**). Subsequently, he was instructed to continue antibiotics therapy and oral administration of levocarnitine. At 34 days of age, the patient showed no signs of hemorrhage. The results of the secondary MRI and blood tandem mass spectrometry analysis were normal. The platelet counts maintained at decreased levels in multiple regular examinations, one of which showed normal level at 108 × $10^9$/L (reference range: 150–450 × $10^9$/L).

Clinical features of his family members were summarized in supplemental **Table 1**, which details a history of thrombocytopenia in the proband's maternal family member. The patient's father is healthy. Of the 11 subjects, 8 presented low platelet counts (IV-2, IV-1, III-4, II-4, II-5, III-5, IV-4, IV-5), and the remaining subjects showed normal platelet level (III-3, III-6, IV-3). The platelet counts of IV-2, IV-1, III-4, and II-4 were examined multiple times, while the rest examined once. The studied family members, including the proband, had platelet counts ranging from 38 to 110 × $10^9$/L with a reference interval of 150–450 × $10^9$/L. The mean volumes and morphology of the sampled platelet were normal. Most studied patients had mild bleeding of skin mucosa, except IV-5 with epistaxis, and the proband showed the spontaneous skin bleeding (Supplemental **Table 1**). All the patients had no symptom of mitochondrial disease such as neurodegeneration, diabetes, myopathy, eye, and kidney diseases.

## 2.2 Sample Collection

This study was approved by the Ethics Committee of Xi'an Children's Hospital, and written informed consent was obtained from each participant or their guardian(s). The peripheral blood (3 ml) was individually sampled from the proband and his family members and collected in EDTA anticoagulant tube.

## 2.3 Whole Exome Sequencing

1 μg genomic DNA was extracted from 200 μL peripheral blood using a Qiagen DNA Blood Midi/Mini kit (Qiagen GmbH, Hilden, Germany) following the manufacturer's protocol. Library preparation was performed using NanoWES (Berry Genomics, China) according to the manufacturer's Protocol. Novaseq6000 platform (Illumina, San Diego, United States) was used for sequencing. The exome sequencing was performed with a minimum median coverage of 80X. The sequencing reads were aligned to the human reference genome (hg38/GRCh38) using Burrows–Wheeler Aligner tool. Verita Trekker® Variants Detection System by Berry Genomics and the third-party software GATK were employed for variant calling. Variants with lower quality (read depth < 10×, allele fraction < 30%) were eliminated. Variant annotation and interpretation were conducted by ANNOVAR (Wang et al., 2010) and the Enliven® Variants Annotation Interpretation System authorized by Berry Genomics. All variants were filtered through population databases including the 1,000 Genomes Project (1000G), Exome Aggregation Consortium (ExAC), and gnomAD, only those variants with population frequencies less than 1/1,000 in all databases were counted. Variant pathogenicity/deleteriousness prediction was evaluated using SIFT, Poly-Phen_2, Mutation Taster, REVEL, FATHMM, CADD. Prediction of variant impact on splicing was evaluated by dbscSNV, Human Splicing Finder (HSF), and SpliceAI. To maximize clinically diagnostic yield, the known pathogenic variants from Human Gene Mutation Database and ClinVar (Landrum et al., 2016; Stenson et al., 2020) were also retained for further evaluation. The variants were classified to five categories "pathogenic," "likely pathogenic," "uncertain significance," "likely benign," and "benign"--according to the American College of Medical Genetics and Genomics (ACMG) guidelines for interpretation of genetic variants (Richards et al., 2015). The suspected SNV/Indels were validated using Sanger sequencing.

The three-dimensional (3D) structure of CYCS protein of the missense variant was analyzed by Swiss-PDBviewer (PDB: 3ZOO).

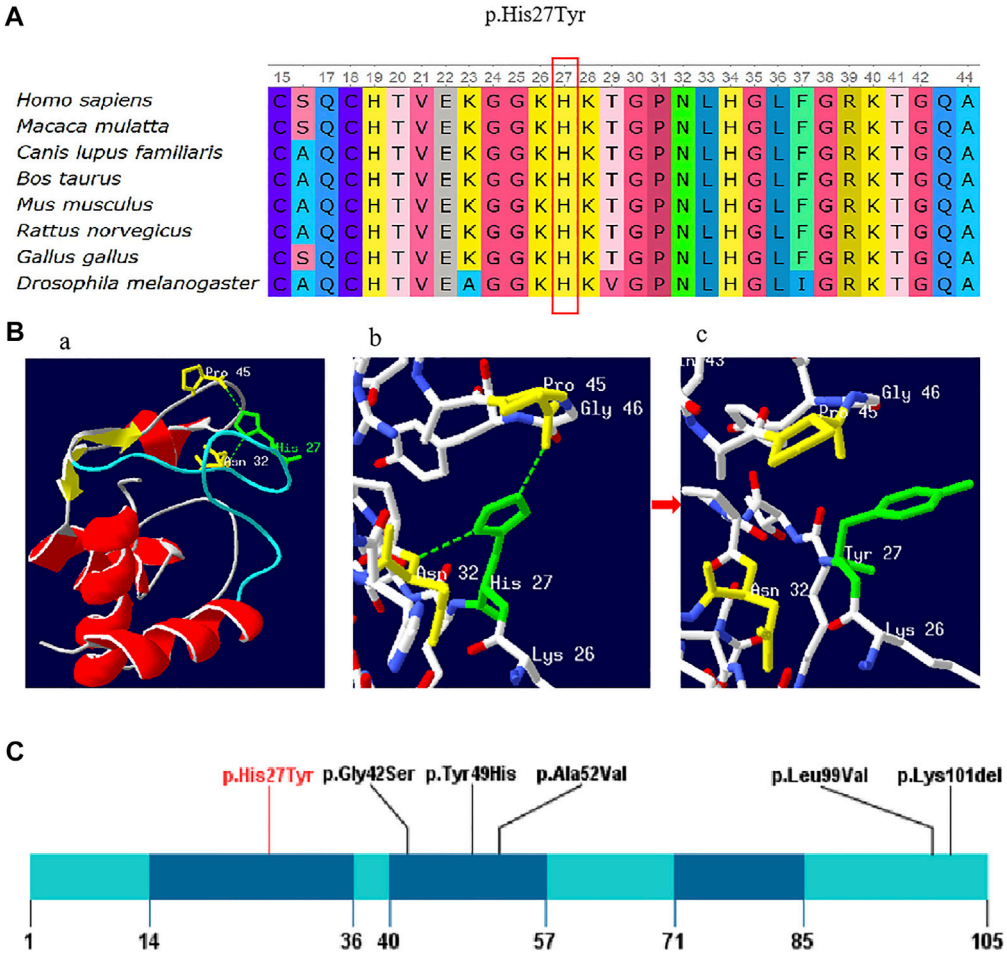



**FIGURE 2 | (A)** Conservation of the p. His27Tyr variant across various species. **(B)** Amino acid and conformation changes of the p.His27Tyr polypeptide wild-type and mutant type. His27 is located in the random coil domain of CYCS by colored green **(Ba)**; Stick models shows the amino acids around His27 and the selected side chains, wild-type His27 forms a hydrogen bond (green dotted line) with Asn32 and Pro45 (colored yellow) respectively **(Bb)**, and the hydrogen bond was lost in mutant type Tyr27 **(Bc)**. **(C)** Schematic presentation of linear CYCS protein (NM_018947.6) with all variants, red font indicate reported variant in the present study. The dark blue indicates three Ω-loops.

# 3 RESULTS

## 3.1 Genetic Analysis and Co-Segregation in the Family

A heterozygous variant in the *CYCS* (NM_018947.6: c.79C > T (p.His27Tyr) was identified for the proband by WES. Sanger sequencing further confirmed that all affected members (IV-2, IV-1, III-4, II-4, II-5, III-5, IV-4, IV-5) carried the same heterozygous variant, while other studied members (III-3, III-6, IV-3) did not carry it (**Figure 1C**). So, the variant segregated with the disorder in this family (PP1_Strong). This novel variant has not been reported in previous literature, and the c.79C > T substitution was not seen in gnomAD database (PM2_Supporting). Sequence alignment of CYCS protein among multiple species showed that p.His27 is highly conserved across evolution (**Figure 2A**), suggesting that His27 could play a vital role in maintaining the stability and function of proteins. The results of multiple statistical methods (REVEL) predicted that the variant could cause detrimental effect on gene function (PP3). According to the ACMG guidelines, c.79C > T (p.His27Tyr) is defined as "Likely Pathogenic". This novel variant has been submitted to ClinVar with the variation ID: 1210164.

## 3.2 Structure-Function Correlations of *CYCS* Variants

The impact of this missense variant was also evaluated by examining the 3D structural viewpoint of the variant protein. Our study mapped the mutation position onto the crystal structure of human cytochrome c (PDB:3ZOO). As shown in **Figure 2B**, p.His27Tyr variant is located in the random coil domain of CYCS (**Figure 2Ba**), which may affect the global conformation and activity of protein. When His27 is replaced by Tyr, the hydrogen bonds of His27 interacting with Asn32 and Pro45 were broken, and the amino acid side chain is changed from imidazole to benzene (**Figures 2Bb,Bc**).

## 4 DISCUSSION

In this study, we firstly reported a 3-generation family with autosomal dominant non-syndromic thrombocytopenia in China. The whole exome sequencing identified a novel missense heterozygous variant c.79C > T (p.His27Tyr) of *CYCS* gene in the pedigree with eight patients (IV-2, IV-1,III-4, II-4, II-5, III-5, IV-4, IV-5). The clinical features of the studied patients were consistent with that from other reports (Morison et al., 2008; De Rocco et al., 2014; Johnson et al., 2016). More, they had normal fertility and longevity, with no evidence of mitochondriopathies (Finsterer, 2004), while other family members (III-3, III-6, IV-3) of the study did not carry the variant. It was assumed that the variation p. His27Tyr of *CYCS* could be the genetic etiology of autosomal dominant non-syndromic thrombocytopenia in this family case. The finding enriches the mutation spectrum of *CYCS* gene and laid a foundation for future studies on the correlation between genotype and phenotype. This conclusion may help patients to prevent the disease risk associated with thrombocytopenia and help clinicians to do differential diagnoses including acquired and inherited forms.

In 2008, Morison *et al.* reported the first variant of the *CYCS* gene identified in a family with thrombocytopenia (Morison et al., 2008). Globally, total five variants have been reported to be associated with thrombocytopenia, including four missense variant (p.Gly42Ser, p.Tyr49His, p.Ala52Val, and p.Leu99Val) and one small deletion variant p.Lys301del (in-frame) (De Rocco et al., 2014; Johnson et al., 2016; Uchiyama et al., 2018; Turro et al., 2020). The p. His27Tyr variant found in this study was the fifth missense mutation. To date, multiple studies has investigated the missense mutation effect on physical structure and bio-function of CYCS protein. Liptak et al. demonstrated that the proapoptotic Gly42Ser mutation altered the heme electronic structure and increased the rate of electron self-exchange, resulting in the enhanced proapoptotic activity of Gly42Ser (Liptak et al., 2011). De Rocco *et al.* reported that the p. Gly42Ser and p.Tyr49His variants in yeast and mouse cellular models were responsible of the diminished respiratory level and increased apoptotic rate (De Rocco et al., 2014). Lei et al. showed that the Ala51Val variant enhanced peroxidase activity by destabilizing the native state of Cyt-c, and all three variants Gly42Ser, Tyr49His, and Ala51Val had reduced global and local stability than that of wild type Cyt-c (Lei and Bowler, 2019). Moreover, Uchiyama *et al.* provided that the mutation of p.Lys301del could significantly reduced cytochrome c protein expression and cause functional defects in the mitochondrial respiratory chain (Uchiyama et al., 2018). The existing findings are not enough to explain the molecular mechanism of thrombocytopenia, however, they could be used as scientific evidence for thrombocytopenia diagnosis.

Cyt-c, a highly evolved protein in different species (Stevens, 2011), contains three highly conserved Ω-loops including residues 14–36, 40–57, and 71–85 (Karsisiotis et al., 2016). Ω-loops play an important role in maintaining the protein function and stability (Fetrow, 1995). Of the five reported missense variants, three variants p.Gly42Ser, p.Tyr49His, and p.Ala52Val were located in the second Ω-loop domain (residues 40–57), two variants p.Leu99Val, and p.Lys301del were located in the C-terminal of the protein. The p.His27Tyr variant found in this study was in the first Ω-loop (residues 14–36) (**Figure 2C**), which may affect the stability of Cyt-c. Previous studies have reported that His27 in the wild type Cyt-c interacted with Pro45 through an hydrogen bond which is essential for sustaining the orientation of the heme conformation and the α-helices, inflicting the cardiolipin binding to cyt-c and subsequent apoptotic events (Balakrishnan et al., 2012). Additionally, the 3D structure of CYCS protein was predicted and it showed that the amino acid substitution (His27Tyr) could lead to the cleavage of hydrogen bonds between His27 and Pro45, causing the instability of the protein structure and destruction of its biological function. For mutant protein, the positive charged Histidine was replaced by the neutral Tyrosine, which negatively influence the electron transport in mitochondrial oxidative respiratory chain. Hence, we assumed that the variant p.His27Tyr could be the genetic etiology of the thrombocytopenia in the family case. Nevertheless, the genetic effect of this variant needs to be verified in future study.

In summary, with the rapid development of molecular biotechnology, genetic analysis has been widely used in rare disease diagnosis. Our findings demonstrated that a new missense variant of the *CYCS* gene was associated with non-syndromic thrombocytopenia identified by WES. Further research is required to understand the impact of *CYCS* variants on changing platelet production. The outcome could be beneficial for thrombocytopenia diagnosis and prognostic management.

## DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available due to ethical and privacy restrictions. Requests to access the datasets should be directed to the corresponding author.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of Xi'an Children's Hospital. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

## AUTHOR CONTRIBUTIONS

FC, JZ, and YY contributed to conception and design of the study. YZ organized the database. ZW and LZ performed the statistical analysis. FC wrote the first draft of the article. YY and JZ wrote sections of the article. All authors contributed to article revision, read, and approved the submitted version.

## FUNDING

This work was supported by Shaanxi Provincial Science and Technology Project (No. 2019KJXX-055) and Xi'an Children's Hospital research project (No. 2020A07).

## ACKNOWLEDGMENTS

The authors would like to thank all of the participation of the family members in this study. And to thank the Berry Gnomics Co., for the technical support.

## REFERENCES

Balakrishnan, G., Hu, Y., and Spiro, T. G. (2012). His26 Protonation in Cytochrome C Triggers Microsecond β-sheet Formation and Heme Exposure: Implications for Apoptosis. *J Am Chem Soc* 134 (46), 19061–19069. doi:10.1021/ja307100a

De Rocco, D., Cerqua, C., Goffrini, P., Russo, G., Pastore, A., Meloni, F., et al. (2014). Mutations of Cytochrome C Identified in Patients with Thrombocytopenia THC4 Affect Both Apoptosis and Cellular Bioenergetics. *Biochim Biophys Acta* 1842 (2), 269–274. doi:10.1016/j.bbadis.2013.12.002

Fetrow, J. S. (1995). Omega Loops: Nonregular Secondary Structures Significant in Protein Function and Stability. *FASEB J* 9 (9), 708–717. doi:10.1096/fasebj.9.9.7601335

Finsterer, J. (2004). Mitochondriopathies. *Eur J Neurol* 11 (3), 163–186. doi:10.1046/j.1351-5101.2003.00728.x

Franchini, M., Veneri, D., and Lippi, G. (2017). Thrombocytopenia and Infections. *Expert Rev Hematol* 10 (1), 99–106. doi:10.1080/17474086.2017.1271319

Greenberg, E. M., and Kaled, E. S. (2013). Thrombocytopenia. *Crit Care Nurs Clin North Am* 25 (4), 427–434. doi:10.1016/j.ccell.2013.08.003

Johnson, B., Lowe, G. C., Futterer, J., Lordkipanidzé, M., MacDonald, D., Simpson, M. A., et al. (2016). Whole Exome Sequencing Identifies Genetic Variants in Inherited Thrombocytopenia with Secondary Qualitative Function Defects. *Haematologica* 101 (10), 1170–1179. doi:10.3324/haematol.2016.146316

Kagan, V. E., Tyurin, V. A., Jiang, J., Tyurina, Y. Y., Ritov, V. B., Amoscato, A. A., et al. (2005). Cytochrome C Acts as a Cardiolipin Oxygenase Required for Release of Proapoptotic Factors. *Nat Chem Biol* 1 (4), 223–232. doi:10.1038/nchembio727

Karsisiotis, A. I., Deacon, O. M., Wilson, M. T., Macdonald, C., Blumenschein, T. M., Moore, G. R., et al. (2016). Increased Dynamics in the 40-57Ω-Loopof the G41S Variant of Human Cytochrome C Promote its Pro -apoptoticconformation. *Sci Rep* 6, 30447. doi:10.1038/srep30447

Landrum, M. J., Lee, J. M., Benson, M., Brown, G., Chao, C., Chitipiralla, S., et al. (2016). ClinVar: Public Archive of Interpretations of Clinically Relevant Variants. *Nucleic Acids Res* 44, D862–D868. doi:10.1093/nar/gkv1222

Lee, E. J., and Lee, A. I. (2016). Thrombocytopenia. *Prim Care* 43 (4), 543–557. doi:10.1016/j.pop.2016.07.008

Lei, H., and Bowler, B. E. (2019). Naturally Occurring A51V Variant of Human Cytochrome C Destabilizes the Native State and Enhances Peroxidase Activity. *J Phys Chem B* 123 (42), 8939–8953. doi:10.1021/acs.jpcb.9b05869

Li, Y., Park, J. S., Deng, J. H., and Bai, Y. (2006). Cytochrome C Oxidase Subunit IV Is Essential for Assembly and Respiratory Function of the Enzyme Complex. *J Bioenerg Biomembr* 38 (5-6), 283–291. doi:10.1007/s10863-006-9052-z

Liptak, M. D., Fagerlund, R. D., Ledgerwood, E. C., Wilbanks, S. M., and Bren, K. L. (2011). The Proapoptotic G41S Mutation to Human Cytochrome C Alters the Heme Electronic Structure and Increases the Electron Self-Exchange Rate. *J Am Chem Soc* 133 (5), 1153–1155. doi:10.1021/ja106328k

Lee, M., and Xu, J. X. (2007). Detoxifying Function of Cytochrome C against Oxygen Toxicity. *Mitochondrion* 7 (1-2), 13–16. doi:10.1016/j.mito.2006.11.011

Morison, I. M., Cramer Bordé, E. M., Cheesman, E. J., Cheong, P. L., Holyoake, A. J., Fichelson, S., et al. (2008). A Mutation of Human Cytochrome C Enhances the Intrinsic Apoptotic Pathway but Causes Only Thrombocytopenia. *Nat Genet* 40 (4), 387–389. doi:10.1038/ng.103

Nurden, A. T., and Nurden, P. (2020). Inherited Thrombocytopenias: History, Advances and Perspectives. *Haematologica* 105 (8), 2004–2019. doi:10.3324/haematol.2019.233197

Ow, Y. P., Green, D. R., Hao, Z., and Mak, T. W. (2008). Cytochrome C: Functions beyond Respiration. *Nat Rev Mol Cell Biol* 9 (7), 532–542. doi:10.1038/nrm2434

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

Santucci, R., Sinibaldi, F., Cozza, P., Polticelli, F., and Fiorucci, L. (2019). Cytochrome C: An Extreme Multifunctional Protein with a Key Role in Cell Fate. *Int J Biol Macromol* 136, 1237–1246. doi:10.1016/j.ijbiomac.2019.06.180

Stenson, P. D., Mort, M., Ball, E. V., Chapman, M., Evans, K., Azevedo, L., et al. (2020). The Human Gene Mutation Database (HGMD®): Optimizing its Use in a Clinical Diagnostic or Research Setting. *Hum Genet* 139 (10), 1197–1207. Epub 2020 Jun 28. doi:10.1007/s00439-020-02199-3

Stevens, J. M. (2011). Cytochrome C as an Experimental Model Protein. *Metallomics* 3 (4), 319–322. doi:10.1039/c0mt00089b

Turro, E., Astle, W. J., Megy, K., Gräf, S., Greene, D., Shamardina, O., et al. (2020). Whole-genome Sequencing of Patients with Rare Diseases in a National Health System. *Nature* 583 (7814), 96–102. doi:10.1038/s41586-020-2434-2

Uchiyama, Y., Yanagisawa, K., Kunishima, S., Shiina, M., Ogawa, Y., Nakashima, M., et al. (2018). A Novel CYCS Mutation in the α-helix of the CYCS C-Terminal Domain Causes Non-syndromic Thrombocytopenia. *Clin Genet* 94 (6), 548–553. doi:10.1111/cge.13423

Wang, K., Li, M., and Hakonarson, H. (2010). ANNOVAR: Functional Annotation of Genetic Variants from Next-Generation Sequencing Data. *Nucleic Acids Res* 38 (16), e164, 2010 . Epub 2010 Jul 3. doi:10.1093/nar/gkq603

Wang, Z. B., Li, M., Zhao, Y., and Xu, J. X. (2003). Cytochrome C Is a Hydrogen Peroxide Scavenger in Mitochondria. *Protein Pept Lett* 10 (3), 247–253. doi:10.2174/0929866033479013

Zhao, Y., Wang, Z. B., and Xu, J. X. (2003). Effect of Cytochrome C on the Generation and Elimination of O2*- and H2O2 in Mitochondria. *J Biol Chem* 278 (4), 2356–2360. doi:10.1074/jbc.M209681200

**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.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

*Copyright © 2022 Che, Zhao, Zhao, Wang, Zhang and Yang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*

frontiers in Genetics

ORIGINAL RESEARCH
published: 25 January 2022
doi: 10.3389/fgene.2022.768000

# Further Delineation of the Spectrum of XMEN Disease in Six Chinese Pediatric Patients

*Xiaomin Peng[1†], Yi Lu[2†], Huijun Wang[3], Bingbing Wu[3], Mingyu Gan[3], Suzhen Xu[3], Deyi Zhuang[4], Jianshe Wang[2], Jinqiao Sun[5], Xiaochuan Wang[5]* and Wenhao Zhou[1]**

[1]Center for Molecular Medicine of Children's Hospital of Fudan University and National Children's Medical Center, Institutes of Biomedical Sciences, Fudan University, Shanghai, China, [2]Center for Pediatric Liver Diseases, Children's Hospital of Fudan University and National Children's Medical Center, Shanghai, China, [3]Center for Molecular Medicine, Key Laboratory of Birth Defects, Pediatrics Research Institute, Children's Hospital of Fudan University and National Children's Medical Center, Shanghai, China, [4]Department of Pediatrics, Xiamen Children's Hospital, Xiamen, China, [5]Department of Clinical Immunology, Children's Hospital of Fudan University and National Children's Medical Center, Shanghai, China

OPEN ACCESS

**Edited by:**
*Natália Duarte Linhares, Genuity Science, Ireland*

**Reviewed by:**
*Fatima Dhalla, University of Oxford, United Kingdom*
*Juan Ravell, Hackensack University Medical Center, United States*

***Correspondence:**
*Xiaochuan Wang*
*xchwang@shmu.edu.cn*
*Wenhao Zhou*
*zhouwenhao@fudan.edu.cn*

*†These authors have contributed equally to this work and share first authorship.*

**Specialty section:**
*This article was submitted to Genetics of Common and Rare Diseases, a section of the journal Frontiers in Genetics*

**Received:** *31 August 2021*
**Accepted:** *05 January 2022*
**Published:** *25 January 2022*

**Citation:**
*Peng X, Lu Y, Wang H, Wu B, Gan M, Xu S, Zhuang D, Wang J, Sun J, Wang X and Zhou W (2022) Further Delineation of the Spectrum of XMEN Disease in Six Chinese Pediatric Patients. Front. Genet. 13:768000. doi: 10.3389/fgene.2022.768000*

X-linked MAGT1 deficiency with increased susceptibility to EBV-infection and N-linked glycosylation defect (XMEN) disease is a primary immunodeficiency caused by loss-of-function variants in the *MAGT1* gene. Only two patients from one family have been diagnosed with XMEN in China. In this study, we retrospectively analyzed the genetic, clinical, and immunological characteristics of six pediatric patients in a Chinese cohort. Medical records were retrieved, immunological phenotypes were assessed, and infectious microbes in patients were detected. Six male patients (mean age, 6.3 years) from five unrelated families were genetically diagnosed as XMEN. Five patients presented with a major complaint of elevated liver enzymes, while one patient was referred for recurrent fever, cough and skin rash. Five patients developed EBV viremia, and one patient developed non-Hodgkin's lymphoma. Histopathological findings from liver biopsy tissues showed variable hepatic steatosis, fibrosis, inflammatory infiltration, and glycogenosis. Immune phenotypes included CD4 T-cell lymphopenia, elevated B cells, inverted CD4/CD8 ratios, and elevated αβDNTs. No pathogenic microbes other than EBV were identified in these patients. This study reports the clinical and molecular features of Chinese patients with XMEN. For patients with transaminase elevation, chronic EBV infection and EBV-associated lymphoproliferative disease, the possibility of XMEN should be considered in addition to isolated liver diseases.

**Keywords: *MAGT1* gene, XMEN, elevated liver enzymes, immunodeficiency, genetic testing**

## INTRODUCTION

The magnesium transporter 1 (*MAGT1*) gene (OMIM: 300715) is located on chromosome Xq21.1 and comprises 10 exons. Its largest transcript variant, NM_032121.5, encodes a 367 amino acid protein conserved in eukaryotes (Zhou and Clapham 2009). MAGT1 protein consists of a signal peptide, a thioredoxin (TRX) domain that contains a bi-cysteine motif (CXXC), and four transmembrane (TM) regions (Zhou and Clapham 2009; Matsuda-Lennikov et al., 2019). As an $Mg^{2+}$-specific ion transporter, MAGT1 regulates free intracellular $Mg^{2+}$ and mediates a transient influx of $Mg^{2+}$ during T-cell activation. MAGT1 shares a 66% amino acid identity with TUSC3 and is a non-catalytic subunit of the oligosaccharyltransferase (OST) complex, facilitating glycosylation in a

STT3B-dependent manner (Zhou and Clapham 2009; Cherepanova et al., 2014). MAGT1 also promotes STT3B-dependent N-linked glycosylation (NLG) of a subset of glycoproteins crucial for proper immune function (Cherepanova et al., 2014; Matsuda-Lennikov et al., 2019; Ravell et al., 2020b).

Hemizygous loss-of-function variants in the *MAGT1* gene lead to a rare primary immunodeficiency known as X-linked MAGT1 deficiency with increased susceptibility to Epstein-Barr virus (EBV) infection and N-linked glycosylation defect (XMEN, previously known as X-linked immunodeficiency with magnesium defect, EBV infection, and neoplasia) disease (Ravell et al., 2020a). MAGT1 deficiency was initially recognized as a novel combined immunodeficiency because the affected patients showed defective development and function of T cells associated with chronic active EBV infections (Li et al., 2011). It has been demonstrated that loss of MAGT1 decreases $Mg^{2+}$ influx as well as the expression of NKG2D in natural killer (NK) cells and T lymphocytes, which are vital for antiviral and antitumor cytotoxicity (Chaigne-Delalande et al., 2013; Li et al., 2014). *MAGT1*-knockout mice also showed impaired $Mg^{2+}$ homeostasis and increased numbers and hyper-activation of B cells, although there were no defects in T-cell activation (Gotru et al., 2018). Further observations expanded the phenotypic spectrum of MAGT1 deficiency to intellectual and developmental disability caused by abnormal glycosylation (Blommaert et al., 2019; Ravell et al., 2020b). Clinical manifestations and severity of XMEN are highly variable, including recurrent ear and sinus infections, chronic EBV infections, lymphadenopathy, molluscum contagiosum, autoimmune cytopenia, EBV-positive lymphoma, and central nervous system abnormalities (Ravell et al., 2020b). Elevated liver enzymes and autoimmune hepatitis were first reported in an adolescent boy who presented with severe autoimmune disorders mimicking autoimmune lymphoproliferative disease (ALPS) (Patiroglu et al., 2015). Recently, Ravell *et al.* found noninfectious liver abnormalities with evidence of defective glycosylation instead of autoimmune hepatitis in patients with XMEN and further distinguished XMEN and ALPS using leukocyte surface marker clusters (Ravell et al., 2020b). Therefore, XMEN has been demonstrated to be a selective congenital disorder of glycosylation that predominantly manifests as immunodeficiency (Ravell et al., 2020a).

Forty-five unique male patients have been reported since XMEN was first reported in 2011 (Li et al., 2011; Chaigne-Delalande et al., 2013; Li et al., 2014; Dhalla et al., 2015; Patiroglu et al., 2015; Brigida et al., 2017; He et al., 2018; Blommaert et al., 2019; Dimitrova et al., 2019; Ravell et al., 2020b; Hoyos-Bachiloglu et al., 2020; Klinken et al., 2020; Brault et al., 2021). However, only two cases from one family have been reported in China to date. These two cases suffered from recurrent upper respiratory tract infections and sinusitis (He et al., 2018). In this study, we reported six patients in five unrelated Chinese families with pathogenic/likely pathogenic (P/LP) variants in *MAGT1*, which lead to a diagnosis of XMEN. Interestingly, most patients presented with symptoms of elevated transaminase levels and chronic EBV infections.

# MATERIALS AND METHODS

## Patients

We retrospectively analyzed six pediatric patients who were diagnosed with XMEN in the genetic laboratory of Children's Hospital of Fudan University from January 2019 to January 2021. Electronic medical records were retrieved and reviewed. Archived diagnostic slides of liver biopsies were re-assessed by an experienced clinical pathologist. This study was approved by the Ethics Committee of Children's Hospital of Fudan University (2015–130). Written informed consent was obtained from the patients' guardians.

## Genetic Testing and Data Analysis

Peripheral blood samples were obtained from the probands and their parents. Genomic DNA was isolated from peripheral blood mononuclear cells using Qiagen genomic DNA extraction kit (Qiagen, Hilden, Germany). Extracted genomic DNA was enriched using the Agilent ClearSeq Inherited disease Kit (Agilent Technologies, Santa Clara, CA, United States) including 2,742 genes for clinical exome sequencing (CES), or the Agilent SureSelect All Exon Human V5 Kit for exome sequencing (ES) as previously described (Dong et al., 2020; Yang et al., 2020). After amplification using polymerase chain reaction (PCR), DNA libraries were sequenced on an Illumina HiSeq2500 platform (Illumina, San Diego, CA, United States) to obtain 150 bp paired-end sequencing reads.

Sequences were aligned with the human genome assembly GRCh37/hg19 using Burrows-Wheeler Aligner (BWA) version 0.7.15-r1140 (Li and Durbin 2009). Subsequent processing of sorting, merging and removing duplicated BAM files was performed using SAMtools (v.1.8) (Li et al., 2009) and Picard tools (v.2.20.1). Variant callings, including single-nucleotide variations (SNV) and small indels, were obtained using the GATK tool and best practice pipelines with default parameters. Home-modified CANOES was applied to detect CNVs from exon-based read counts (Dong et al., 2020). Candidate pathogenic variants in *MAGT1* were validated using Sanger sequencing. The nonsense-mediated mRNA decay (NMD) efficacy of novel *MAGT1* variants was predicted using an online NMDective resource (Lindeboom et al., 2019). Pathogenicity was defined based on the standards and guidelines of the American College of Medical Genetics and Genomics (ACMG) (Richards et al., 2015) and the ClinGen Sequence Variant Interpretation (SVI) Working Group.

## Identification of Pathogenic Microbes

For metagenomic sequencing, DNA was extracted from liver biopsy tissues of patient 4 and 5 (P4 and P5, respectively) using a QIAamp UCP Pathogen Mini Kit (Qiagen, Hilden, Germany). The sequencing library was constructed using KAPA Hyper Prep Kit (Kapa Biosystems, Wilmington, MA, United States). Libraries were then pooled and sequenced on the Illumina HiSeq 2000

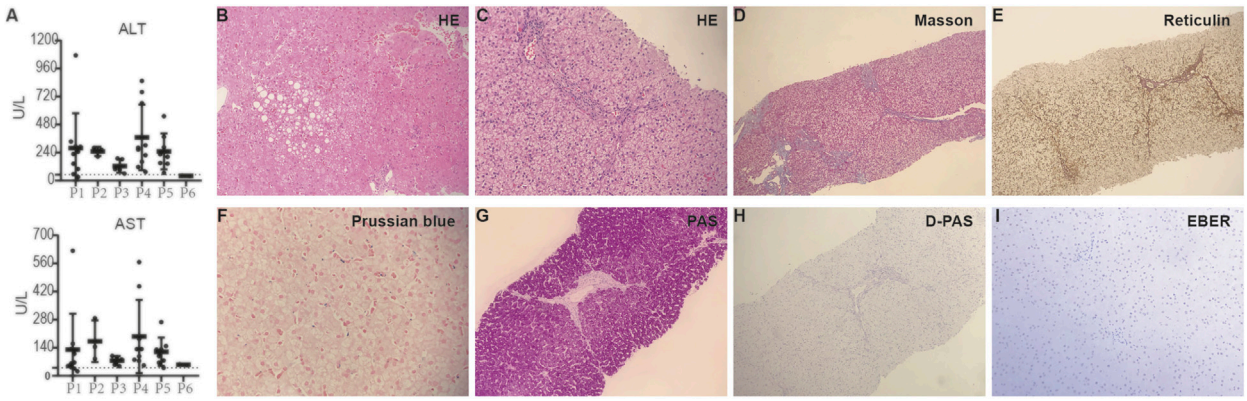



**FIGURE 1 |** Abnormal liver function tests and pathological findings of liver tissues in patients with XMEN disease in this study. **(A)** ALT (upper panel) and AST (lower panel) levels in all patients with a normal range (the dotted line, except for P6) and patients' identities (shown on the *X*-axis). **(B–H)** Representative images depicting histopathologic findings from liver biopsy tissues of patients. **(B & C)** Hematoxylin and eosin staining results showing mild hepatocyte swelling, hepatosteatosis, periportal fibrosis, inflammatory infiltrations and sinusoidal iron deposits, ×200. **(D)** Masson's Trichrome staining results showing periportal fibrosis and bridging fibrosis, ×100. **(E)** Reticulin staining results showing preserved despite focal disrupted reticulin framework, ×100. **(F)** Prussian blue staining results showing mild iron deposition in hepatocytes, ×100. PAS **(G)** and D-PAS **(H)** staining depicting diffuse glycogenosis, ×100. **(I)** Liver tissues from patient 1 were negative for EBER staining, ×200. *P*, patient.

platform (Illumina, San Diego, CA, United States) using a 75-cycle single-end index sequencing kit. At least 20 million single-end reads were obtained from each sample. Reads were quality trimmed with a 10-base sliding window. Trimming was performed when the average base quality dropped below 15 using Trimmomatic v0.39 (Bolger et al., 2014). Reads shorter than 40 bases and human reads were subsequently removed. The remaining reads were aligned to the microorganism reference genome database using Centrifuge v1.0.3. Reads with an alignment rate >70% were retained. In addition, reads that aligned to multiple different species genomes were excluded.

To detect infectious pathogens including EBV in the blood, DNA was extracted from whole blood of these six patients using QIAamp UCP Pathogen Mini Kit. The diluted pre-amplification products were mixed with TaqMan OpenArray Real-Time PCR Master Mix and automatically dispensed on the OpenArray plate. Real-time PCR of the OpenArray plate was performed using the QuantStudio 12K Flex instrument (Thermo Fisher Scientific, Waltham, MA, United States) following the manufacturer's instructions. Non-template control (NTC) was added to each batch of real-time PCR. Cq confidence, amplification score, and Ct value was calculated using the Thermo Fisher Scientific's proprietary algorithm. Amplifications with an amplification score >1.2 and Cq confidence >0.7 were considered of sufficient quality and retained for the following analyses. Positive detection was identified only when all the replicated assays were qualified amplified and no amplification was detected in the NTC.

## Immunological Assays

Routine blood counts and immunological phenotypes were evaluated as previously reported (Wang et al., 2018). The levels of serum immunoglobulins, including IgG, IgA, IgM, and IgE, were detected using nephelometry, while lymphocyte subsets were analyzed using a FACSCalibur flow cytometer (Becton Dickinson, Franklin Lakes, NJ, United States). The following validated antibodies were used for flow cytometry: anti-CD3 (UCHT1), anti-CD4 (RPA-T4), anti-CD8 (RPA-T8), anti-CD27 (M-T271), anti-TCRαβ (T10B9.1A-31), anti-TCRγδ (B1), anti-CD45RA (HI100), anti-CD19 (HIB19), anti-CD24 (ML5), anti-CD38 (HIT2), and anti-IgD (IA6–2) (all from BD Biosciences, San Jose, CA, United States).

# RESULTS

## Clinical Manifestations

Six male patients from five unrelated families were clinically diagnosed with XMEN based on genetic testing results, laboratory studies, and clinical manifestations. The mean age at diagnosis was 6.3 years (range, 19 months to 10-years-7-months). All patients were born to healthy, non-consanguineous parents. P4 and P5 were from the same family, and the maternal uncle died of lymphoma at 10 years of age.

Interestingly, five of six patients (except for P6) were referred to the center for liver diseases in our hospital because of abnormal liver function tests (LFTs) (**Figure 1A**). These five patients were negative for conventional autoantibodies associated with autoimmune hepatitis, including anti-ANA, anti-LKM, anti-AMA, and anti-SMA. Liver diseases such as intrahepatic cholestasis, biliary atresia and Wilson's disease were excluded because the five patients had normal levels of serum γ-glutamyltransferase (GGT), muscle creatine phosphokinase (CPK), bilirubin, bile acid, and ceruloplasmin. However, P6 presented with recurrent fever, cough, oral ulcers, and urticaria; these symptoms lead to clinical suspicion of an immunodeficiency disease. He developed neutropenia ($1.51 \times 10^9$ cells/L at the last evaluation, range = $1.18–1.74 \times 10^9$ cells/L).

**TABLE 1 |** Clinical and genetic characteristics of six XMEN patients in this study.

| | Patient 1 | Patient 2 | Patient 3 | Patient 4 | Patient 5 | Patient 6 |
|---|---|---|---|---|---|---|
| **Sex** | **male** | **male** | **male** | **male** | **male** | **male** |
| Age at onset | 4 y6 m | 3 y2 m | 4 y3 m | 5 y | 7 y2 m | 13 m |
| Age at genetic testing | 10 y7 m | 4 y8 m | 6 y3 m | 6 y2 m | 8 y5 m | 19 m |
| Major complaint | abnormal LFTs for 5 years and intrahepatic nodules for 1 y | abnormal LFTs for more than 1 y | abnormal LFTs for 6 m | abnormal LFTs for 1 y | abnormal LFTs for 1 y | recurrent upper respiratory tract infection |
| *MAGT1* variants | | | | | | |
| chromosome location (hg19) | chrX:77111082 | chrX:77111053 | chrX:77150909 | chrX:77096749 | chrX:77096749 | chrX:77112895 |
| cDNA (NM_032121.5) | exon6:c.774dupT | exon6:c.803G > A | exon1:c.95dupA | exon8:c.991C > T | exon8:c.991C > T | exon4:c.586G > T |
| protein | p.(Val259CysfsTer21) | p.(Trp268Ter) | p.(Asn32LysfsTer40) | p.(Arg331Ter) | p.(Arg331Ter) | p.(Glu196Ter) |
| inheritance | *De novo* | *De novo* | Maternal | Maternal | Maternal | Maternal |
| classification | pathogenic (PVS1+PM2_Supporting + PM6) | pathogenic (PVS1+PM2_Supporting + PM6) | likely pathogenic (PVS1+PM2_Supporting) | likely pathogenic (PVS1+PM2_Supporting) | likely pathogenic (PVS1+PM2_Supporting) | likely pathogenic (PVS1+PM2_Supporting) |
| ClinVar accession number | SCV001827201.1 | SCV001827202.1 | SCV001827203.1 | — | — | SCV001827204.1 |
| Liver function tests | | | | | | |
| ALT (U/L) | 30–1072 (9–50) | 205–280 (9–50) | 59–189.8 (9–50) | 74.5–853.1 (9–50) | 65–551 (9–50) | 38–42 (0–40) |
| AST (U/L) | 23–623 (15–40) | 85–287.4 (15–40) | 48–99.6 (15–40) | 52.20–566.8 (15–40) | 40–268 (15–40) | 53–58 (0–40) |
| GGT (U/L) | 21.3–37.9 (8–57) | 16–39.2 (8–57) | 25–41.4 (8–57) | 34–202 (8–57) | 23.6–44.6 (8–57) | 12–15 (7–50) |
| CPK (U/L) | 80–117 (0–164) | 78–97 (0–164) | 147–176 (0–164) | 125–221 (0–164) | 134–148 (0–164) | NA |
| Liver histopathology | mild steatosis, sinusoidal iron deposits | NA | moderate steatosis, glycogenosis | bridging fibrosis, glycogenosis | bridging fibrosis, glycogenosis | NA |
| Neutropenia | + | - | - | - | - | + |
| Sinopulmonary infections | + | + | + | + | + | + |
| Lymphadenopathy | + | + | - | + | + | - |
| Hepatosplenomegaly | + | - | - | + | + | - |
| EBV-DNA in PBMC (copies/ml) † | $8.02 \times 10^4$ | $8.12 \times 10^5$ | $5.83 \times 10^3$ | $1.19 \times 10^5$ | $1.52 \times 10^5$ | - |
| Pan-pathogen detection | EBV | EBV | EBV | EBV | EBV | negative |
| Lymphoma (age) | non-Hodgkin's lymphoma (5y2m) | no | no | no | no | no |
| NKG2D expression | NA | NA | NA | NA | NA | NA |
| Immunologic data | | | | | | |
| IgG (g/L) | 7.50 (6.09–12.85) | 6.13 (6.94–16.2) | 10.50 (4.95–12.74) | 5.60 (4.95–12.74) | 5.40 (6.09–12.85) | 7.6 (6.94–16.2) |
| IgA (g/L) | 0.64 (0.52–2.16) | 0.12 (0.68–3.78) | 1.19 (0.33–1.89) | 0.35 (0.33–1.89) | 0.21 (0.52–2.16) | 0.165 (0.82–4.53) |
| IgM (g/L) | 2.55 (0.67–2.01) | 0.83 (0.6–2.63) | 1.10 (0.65–2.01) | 1.24 (0.65–2.01) | 0.91 (0.67–2.01) | 1.12 (1–3) |
| total T cells (%) | 61.20 (64–73) | 72.68 (59–84) | 68.51 (64–73) | 44.72 (64–73) | 59.28 (64–73) | 47.12↑* |
| total T cells (cells/μL) | 1755.3 (1325–2276) | NA | 2042.3 (1424–2664) | 1155.3 (1424–2664) | 1866.2 (1325–2276) | 4208 (1563–3929) |
| CD4 T cells (%) | 20.15 (29–36) | 26.17 (31–60) | 34.20 (29–36) | 18.23 (29–36) | 27.84 (29–36) | 11.20* |
| CD4 T cells (cells/μL) | 578.09 (531–1110) | NA | 1019.65 (686–1358) | 470.99 (686–1358) | 876.31 (531–1110) | 916 (738–2001) |
| CD8 T cells (%) | 31.65 (24–34) | 40.20 (13–38) | 25.59 (24–34) | 14.82 (24–34) | 21.66 (24–34) | 25.16↑* |
| CD8 T cells (cells/μL) | 907.87 (480–1112) | NA | 762.93 (518–1125) | 382.90 (518–1125) | 681.76 (480–1112) | 2058 (532–1549) |
| CD4/CD8 | 0.59 (0.81–1.66) | 0.65 (0.9–3.6) | 1.34 (0.87–1.94) | 1.23 (0.87–1.94) | 1.29 (0.81–1.66) | 0.45 (0.90–2.13) |
| total B cells (%) | 24.95 (14–21) | 17.22 (7–22) | 14.27 (14–21) | 42.89 (14–21) | 28.94 (14–21) | 37.08↑* |

(Continued on following page)

**TABLE 1 |** (*Continued*) Clinical and genetic characteristics of six XMEN patients in this study.

| | Patient 1 | Patient 2 | Patient 3 | Patient 4 | Patient 5 | Patient 6 |
|---|---|---|---|---|---|---|
| **Sex** | **male** | **male** | **male** | **male** | **male** | **male** |
| total B cells (cells/μL) | 715.54 (216–536) | NA | 425.46 (280–623) | 1107.97 (280–623) | 911.03 (216–536) | 3311 (261–960) |
| NK cells (%) | 13.37 (11–23) | 4.18 (6–27) | 16.07 (11–23) | 11.29 (11–23) | 10.15 (11–23) | 11.21†* |
| NK cells (cells/μL) | 383.54 (246–792) | NA | 479.09 (258–727) | 291.76 (258–727) | 319.66 (246–792) | 1001 (197–786) |
| αβDNT (%) | 3.95 (0.61–2.31) | NA | NA | NA | NA | NA |
| Follow-up | | | | | | |
| age at last follow-up | 13 y | 6 y6 m | 7 y5 m | 7 y | 9y3m | 26 m |
| outcome | remission of lymphoma, AS | AS | AS | AS | AS | AS |

*The number in the round bracket presents the age-specific reference values according to reference values for liver enzyme levels or peripheral blood lymphocyte subsets of healthy children in China; † prior to treatment; *the reference values are not available; NA, data not available; AS, alive symptomatic.*

Two patients (P4 and P5) underwent brain MRI, and showed no specific neurological findings, although half of the patients had cavum septum pellucidum in a recent study (Ravell et al., 2020b). Intellectual disability or facial dysmorphism was not observed in any of these patients.

Recurrent upper respiratory tract infections were observed in P1, P2, P3 and P4. Aphthous ulcers and sinusitis were observed in P4 and P5, rhinitis in P4, and otitis media in P5. P2 experienced an episode of erythema multiforme during pneumonia. Neck and submandibular lymphadenopathy was observed in P1 and P2, respectively. Generalized lymphadenopathy and lymphadenitis occurred in both P4 and P5. Moreover, P1 developed a painless mass on the left lower lip at 4-years-6-months and was diagnosed with mature B-cell non-Hodgkin's lymphoma. He also had a history of suppurative osteomyelitis of the right knee and intrahepatic nodules and presented with mild hepatosplenomegaly. Abdominal ultrasonography revealed increased spleen thickness and splenomegaly in P4 and P5.

Four patients (P1, P3, P4, and P5) underwent percutaneous core needle liver biopsy. Histopathological analysis revealed varying degrees of liver damage (**Figures 1B–H**). P1 had mild steatosis, mild sinusoidal iron deposits, and lymphocyte infiltration. P3 showed moderate hepatic steatosis, small vacuoles and eosinophilic bodies in the hepatocyte cytoplasm, and glycogenated hepatocyte nuclei. P4 and P5 showed a similar pattern of injury, including hepatocyte swelling, periportal fibrous proliferation, bridging fibrosis, inflammatory cell infiltration, and glycogenosis. The liver specimen from P1 was EBV-encoded small RNA (EBER) negative (**Figure 1I**). According to metagenomic sequencing, the liver specimens from P4 and P5 were negative for EBV sequences. All patients, except P6, had EBV infection, confirmed through the detection of EBV-specific antibodies or EBV DNA in blood samples. Probe-based qPCR of whole blood samples revealed that no patient had cytomegalovirus (CMV), adenovirus, or herpes simplex virus type 1 (HSV-1). Infection studies showed negative results for hepatitis A, hepatitis B, hepatitis C, CMV, rubella virus, parvovirus B19, and human immunodeficiency virus (HIV) in all patients, though P1 had a history of transient HSV-1 infection. The clinical and genetic characteristics of these six patients are summarized in **Table 1** and **Supplementary Material**.

## Genetic Testing Results

Three of the six patients had novel variants in *MAGT1* (NM_032121.5, P1: c.774dupT; P2: c.803G > A and P3: c.95dupA) (**Table 1**; **Figures 2**, **3**). The variants found in P1 and P2 were not detected in their mothers, indicating that they were *de novo* variants. P1 underwent proband CES testing, and his mother underwent Sanger sequencing. Genetic studies revealed a hemizygous variant c.774dupT in *MAGT1* in P1; this variant is predicted to cause a frameshift with premature termination of the protein (p.(Val259CysfsTer21)). Furthermore, this variant was not detected in his mother through Sanger sequencing. P2 was found to harbor a *de novo* nonsense variant in *MAGT1* (c.803G > A); this variant is predicted to result in a stop codon with premature termination of the protein (p.(Trp268Ter)). Trio-ES was

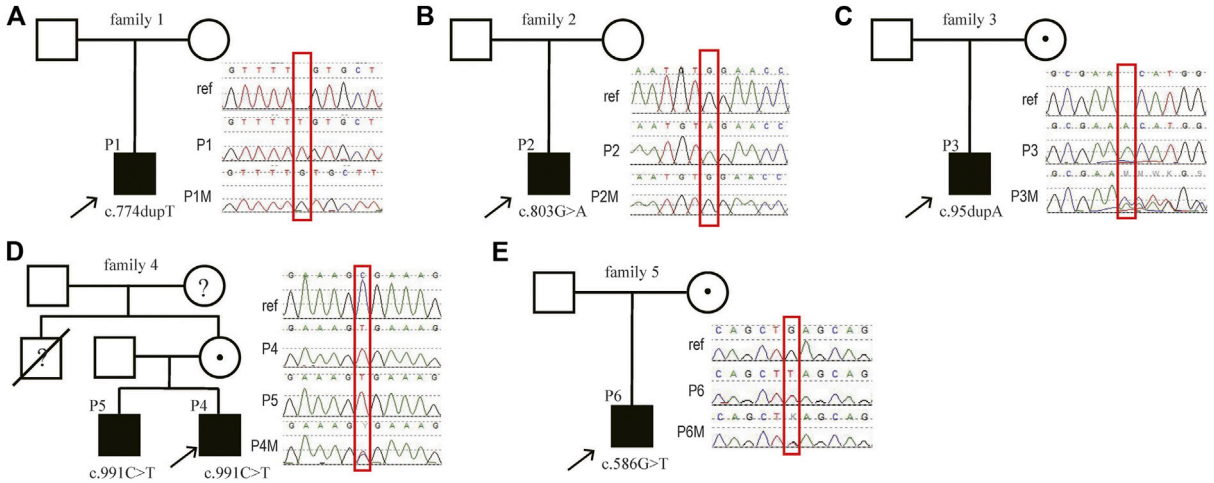



**FIGURE 2 |** Identification and validation of the *MAGT1* variants in six patients with XMEN disease in this study. Pedigrees of patient 1 **(A)**, patient 2 **(B)**, patient 3 **(C)**, patient 4 and 5 **(D)** and patient 6 **(E)** carrying disease-causing variants in the *MAGT1* gene. *P*, patient; *M*, mother; *ref*, reference sequence.

conducted in family 3 and revealed a frameshift variant (c.95dupA, p.(Asn32LysfsTer40)) in *MAGT1* in the proband inherited from his mother. Trio-ES showed a hemizygous nonsense variant (c.586G > T, p.(Glu196Ter)) in *MAGT1* in P6 inherited from his mother. P4 and P5 were found to inherit a previously reported disease-causing variant of c.991C > T (Blommaert et al., 2019; Ravell et al., 2020b; Brault et al., 2021) from their mother, introducing an amino acid exchange from arginine to a stop codon at position 331 (p.(Arg331Ter)). The maternal uncle died at 10 years of age without diagnosis (**Figure 2**).

None of these novel variants in *MAGT1* has been previously reported in known public databases (gnomAD and 1000 Genomes) or our in-house database (more than 40,000 individuals). The loss-of-function intolerance (pLI) score for *MAGT1* is 0.96, and the observed/expected (o/e) ratio is 0.07; suggesting that *MAGT1* is highly intolerant to loss-of-function variants. Most reported disease-causing mutations in *MAGT1* are loss-of-function variants (Li et al., 2011; Chaigne-Delalande et al., 2013; Li et al., 2014; Dhalla et al., 2015; Patiroglu et al., 2015; Brigida et al., 2017; He et al., 2018; Blommaert et al., 2019; Dimitrova et al., 2019; Ravell et al., 2020b; Hoyos-Bachiloglu et al., 2020; Klinken et al., 2020; Brault et al., 2021). All these novel variants occur upstream to the last *MAGT1* exon. In addition, all variants except c.95dupA are predicted to efficiently trigger NMD according to the results of the NMDective algorithm. However, the adjacent nonsense variant c.110G > A has been reported in patients diagnosed with XMEN who presented with EBV viremia, inverted CD4/CD8 ratios, and decreased NKG2D expression (Li et al., 2014; Ravell et al., 2020b). According to the ACMG guidelines and the ClinGen SVI recommendations, these novel variants are classified as P/LP variants.

## Autoimmune Characteristics and Immune Phenotype Evaluation

Routine blood analysis revealed transient mild neutropenia in P1 and P6, and lymphocytosis in P2 and P6. No other blood system abnormalities were observed in these patients (data not shown). IgA level was decreased in P2, P5, and P6, whereas IgG level was slightly decreased in P2 and P5. CD4 T-cell lymphopenia was observd in four patients (P1, P2, P4, and P5), although opportunistic infections were rarely observed. Three patients (P1, P2, and P6) had inverted CD4/CD8 T-cell ratios, and four patients (P1, P4, P5, and P6) had increased numbers of B cells. NK-cell counts were slightly decreased in P2 and P5. P1 underwent T-cell and B-lymphocyte subpopulation analysis, which revealed increased numbers of $CD4^{-}CD8^{-}TCR\alpha\beta^{+}$ T cells (αβDNTs) and naïve cytotoxic T cells ($CD8^{+}CD27^{+}CD45RA^{+}$) (**Table 1** and **Supplementary Table S1**). Unfortunately, NKG2D expression on NK and CD8 cells was not evaluated. No other significant abnormalities were found in the routine assessment of the immune phenotypes.

## Treatment and Outcome

P1 received five courses of chemotherapy according to a modified B-NHL-BFM 95 protocol for non-Hodgkin's lymphoma. Rituximab was added in his last chemotherapy session considering bone marrow involvement, and the patient achieved complete remission of lymphoma for 7 years. All six patients underwent prophylactic antibiotic and symptomatic treatment. None of them received intravenous immunoglobulin (IVIG) or hematopoietic stem cell transplantation (HSCT). All patients were in stable condition during follow-ups. Specifically, all patients had sinopulmonary infections for 5–10 times per year; all patients except P6 had persistent EBV viremia accompanied by fluctuated liver transaminase levels.

# DISCUSSION

In this study, we reported five MAGT1 deficient patients with a major complaint of abnormal LFTs. P1 had persistently elevated liver enzymes prior to and throughout chemotherapy for non-

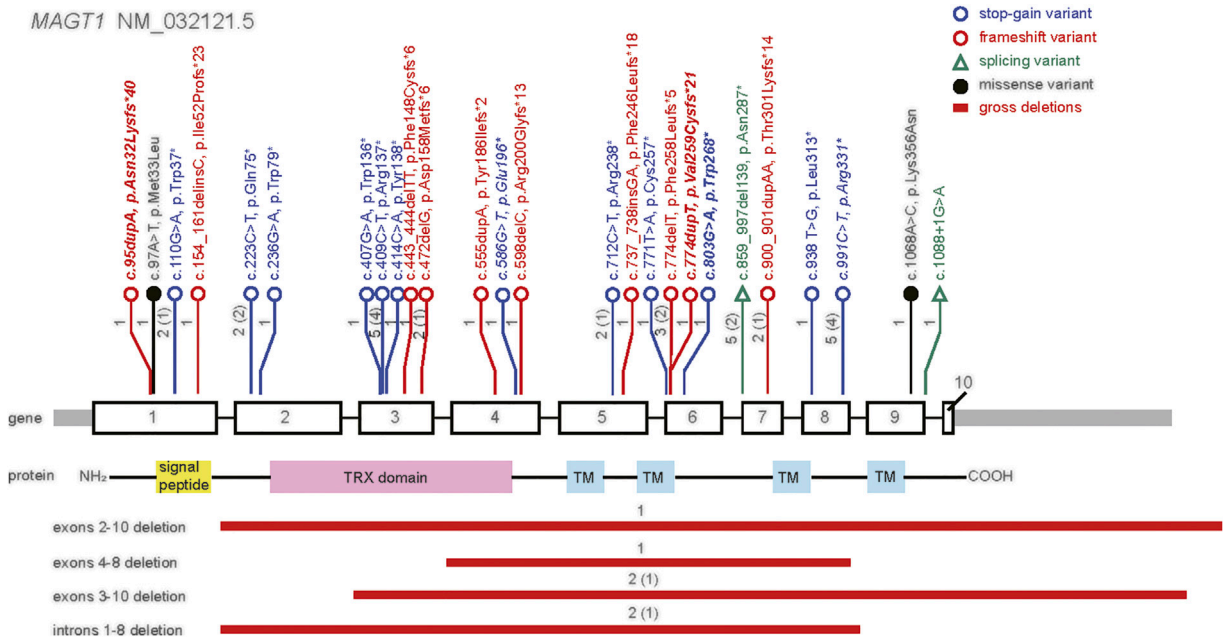



**FIGURE 3 |** Distribution of disease-causing variants in *MAGT1* gene in previous literatures and the present study. MAGT1 protein consists of a signal peptide, a thioredoxin (TRX) domain and four transmembrane (TM) regions. Variants identified in this study are in italics, and novel variants are in bold. The number of patients carrying each variant is labeled with family numbers in round brackets.

Hodgkin's lymphoma, and the abnormal LFTs had been detected for the past 5 years. P2, P3, P4 and P5 presented with mild to moderate elevated serum transaminases (ALT and AST) levels and mostly normal CPK and GGT levels. The symptom of elevated LFTs was first described in a 15-year-old Turkish boy in 2015, who presented with frequent sinopulmonary infections and severe autoimmune disorders, and was finally diagnosed with XMEN (Patiroglu et al., 2015). However, autoimmune hepatitis was excluded from our patients because detection results of conventional autoantibodies were negative. Infection studies showed negative results for hepatitis A, hepatitis B, and hepatitis C. EBER staining of the liver tissue from P1 was negative. The EBV sequence could not be detected in the liver tissues from P4 and P5 using metagenomic sequencing. Liver biopsies from four patients with EBV infection showed no evidence of EBV-associated hepatitis. Instead, histopathology revealed a non-specific pattern of injury with variable degrees of hepatosteatosis, inflammatory infiltrates, fibrosis, and glycogenosis; this finding is consistent with that of a previous study (Ravell et al., 2020b). In the largest XMEN cohort study, transient and asymptomatic elevations in ALT and AST levels were observed in all 23 patients, whereas elevation in serum CPK was detected in 12 cases (Ravell et al., 2020b). Both EBV-infected and EBV-negative patients exhibited similar histopathological patterns and liver enzyme levels (Ravell et al., 2020b). Blommaert *et al.* (Blommaert et al., 2019) also demonstrated that MAGT1-deficient patients have a defect in glycosylation with enhanced expression of the MAGT1 homolog protein TUSC3. Taken together, these findings suggest that liver diseases in XMEN are related to selective NLG defects but not to EBV infections.

All patients in our study, except P6, presented with EBV infection. Chronic EBV infections were common in XMEN patients because of the defective expression of NKG2D as well as decreased glycosylation of NKG2D and CD70. Decreased NKG2D expression on CD8 T cells and NK cells is the most consistent diagnostic finding of XMEN (Ravell et al., 2020a). Unfortunately, our patients were not available for NKG2D protein evaluation. *In vitro* studies also proved the impaired EBV-specific cytotoxic function on both CD8 T-cells and NK cells of XMEN patients (Chaigne-Delalande et al., 2013; Li et al., 2014; Ravell et al., 2020b). The earliest patients were identified based on persistently high EBV levels and inverted CD4/CD8 ratios (Chaigne-Delalande et al., 2013; Li et al., 2014). Additionally, young children without persistent EBV viremia with an immune phenotype similar to that of EBV-infected patients have been identified (Ravell et al., 2020b). In a review article involving 36 XMEN patients, 1/3 of them developed EBV-driven lymphoproliferation disease or lymphoma, with the age at diagnosis ranging from 5 to 57 years. And five were reported to experience more than one malignancy in their life (Ravell et al., 2020a). A young adult subsequently developed an EBV-negative liposarcoma after the successful treatment of Hodgkin's lymphoma (Dimitrova et al., 2019). P1 in this study developed non-Hodgkin's lymphoma at 4 years, and he is now of complete remission after chemotherapy.

Patients with XMEN had a higher predisposition to infectious diseases. Besides persistent EBV viremia, other infections such as CMV, HSV, herpes zoster, and streptococcal infections have also been recorded in XMEN (Li et al., 2014; Patiroglu et al., 2015; Hoyos-Bachiloglu et al., 2020). In this study, all patients were

negative for CMV, rubella virus, parvovirus B19, and HIV. To comprehensively evaluate the potential infectious pathogens, we used a probe-based qPCR method to identify the presence of pathogenic microbes in stored whole blood samples. No pathogens other than EBV were detected, although P1 had a history of transient HSV infection. P6 was referred to clinicians for recurrent upper respiratory tract infection, oral ulcers and neutropenia. He was negative for EBV serology and EBV-DNA, and the pathogen detection result was also negative. However, two EBV-negative patients were reported to develop EBV infection during follow-up (Ravell et al., 2020b); therefore, routine checks for EBV infection would be needed for P6 in later life. We also applied metagenomic sequencing for sequence-based identification of pathogenic microbes, including viruses, bacteria, fungi, and parasites, in liver biopsy tissues from P4 and P5 and obtained negative results.

We evaluated the immune phenotypes in our patients and observed low IgA and IgG levels, inverted CD4/CD8 ratios and increased numbers of B cells. P1 underwent lymphocyte subpopulation examination, and elevated αβDNTs was observed, which was reported in 95% of patients with XMEN (Ravell et al., 2020b). All patients in our study had characteristic peripheral blood abnormalities, and they also presented with typical immunologic symptoms. Two patients with *MAGT1* variants were reported to present with different phenotypes of intellectual and developmental disability. They were subsequently diagnosed as congenital glycosylation defects (Blommaert et al., 2019). The finding of selective NLG defect in XMEN disease expanded the clinical presentations to multisystem abnormalities, including hepatic and neurological abnormalities (Ravell et al., 2020b). Taken together, these findings highlight the clinical diversity observed in XMEN patients.

Since its first report in 2011, XMEN disease has been reported in a total of 45 patients aged between 18 months and 58 years. All reported patients were detected P/LP variants in the *MAGT1* gene (Li et al., 2011; Chaigne-Delalande et al., 2013; Li et al., 2014; Dhalla et al., 2015; Patiroglu et al., 2015; Brigida et al., 2017; He et al., 2018; Blommaert et al., 2019; Dimitrova et al., 2019; Hoyos-Bachiloglu et al., 2020; Klinken et al., 2020; Ravell et al., 2020b; Brault et al., 2021). In total, 30 P/LP variants have been reported, including 3 novel variants identified in this study. These variants are distributed over all exons of the gene, except for exon 10 (**Figure 3**). There are 12 stop-gain, 10 frameshift, 2 splicing, 2 missense variants, and 4 gross deletions in *MAGT1*; this finding indicates a loss-of-function mechanism of this gene. Of note, the missense variant c.1068A > C was detected in a patient characterized by developmental disability (Blommaert et al., 2019). Another missense variant c.97A > T was observed in a patient suffered from EBV-driven cutaneous lymphoproliferation, three times of distinct lymphomas, and multiple infection histories (Klinken et al., 2020). The most common variant, c.409C > T, appeared in five patients from four unrelated families (Ravell et al., 2020b). Another common variant, c.991C > T, was detected in P4 and P5 from family 4, in two reported patients with classical XMEN (Ravell et al., 2020b; Brault et al., 2021), and in another case with a variable phenotype of developmental disability (Blommaert et al., 2019). Therefore, XMEN appears to exhibit variable severity of clinical phenotypes even with identical genotypes. There do not appear to be specific genotype-phenotype correlations among individuals with the disease.

Currently, there lack of effective treatments for XMEN. IVIG combined with antibiotic therapy may reduce opportunistic infections. *In vitro* studies have revealed that MAGT1 deficiency abrogates magnesium influx into T cells. However, oral magnesium supplementation has not proven successful improvement in patients (Chaigne-Delalande et al., 2013; Li et al., 2014; Klinken et al., 2020; Chauvin et al., 2021). Allogeneic HSCT has also been applied to treat XMEN; the disease phenotype reversed in two out of five patients, and three died of transplant-related complications (Li et al., 2014; Dimitrova et al., 2019). Of note, all patients who received HSCT presented with thrombocytopenia and significant hemorrhage after transplantation; this result may suggest a qualitative platelet defect in XMEN patients. Considering the unpredictable complications of HSCT and the stable clinical conditions of our six patients, none of them underwent HSCT. Recently, a pre-clinical study demonstrated that *ex vivo* correcting autologous T and NK cells from XMEN patients through *MAGT1* mRNA electroporation can restore NKG2D receptor expression and rescue cytotoxic activity to normal levels (Brault et al., 2021). This highly efficient method avoids genomic integration and the risks of alloimmunity, supporting cell therapy as an attractive immunologic cure for this monogenetic disease.

In this study, we diagnosed six pediatric XMEN cases predominantly based on genetic testing results. Five patients were referred to the center for pediatric liver diseases, and immunological phenotypes were not evaluated until the molecular diagnosis was made. The combined data from clinical phenotypes and genetic testing suggested a diagnosis of XMEN disease. XMEN is a rare, underdiagnosed disease with manifestations ranging from immunodeficiency and immune dysregulation, to extra-immune symptoms. Some patients were not diagnosed until adulthood because of the insidious onset and indolent course of the disease (Dhalla et al., 2015; Ravell et al., 2020a; Ravell et al., 2020b). To date, 48.9% (22/45) of the patients were diagnosed in their childhood (<18 years of age). The clinical and immunological features of 18 patients with detailed records (Li et al., 2011; Chaigne-Delalande et al., 2013; Li et al., 2014; Dhalla et al., 2015; Patiroglu et al., 2015; Brigida et al., 2017; He et al., 2018; Blommaert et al., 2019; Dimitrova et al., 2019; Hoyos-Bachiloglu et al., 2020; Klinken et al., 2020; Ravell et al., 2020b; Brault et al., 2021) are summarized in **Supplementary Table S2**. Long-term follow-up is necessary in pediatric patients as the clinical and immunological phenotypes of XMEN may evolve with age. Most cases developed EBV-associated malignancies in their second and third decades of life (Ravell et al., 2020a). P1 was the first diagnosed case and was followed up for 2 years. The patient achieved complete remission of lymphoma for 7 years. Patients may have brain atrophy that is more significant than expected for their age or white matter abnormalities consistent with leukoencephalopathy (Ravell et al., 2020b). Cavum septum pellucidum was found in half patients with XMEN, with a significantly higher rate than that in the general population (Ravell et al., 2020b). Further management and treatment should be modified according to the symptoms of each patient.

In conclusion, this study expands our understanding of the clinical and immunological features as well as the mutational

spectrum of XMEN disease. The combined immunodeficiency should be suspected in male patients with abnormal LFTs, chronic EBV infection and EBV-associated lymphoproliferative disease. Genetic diagnosis with *MAGT1* will be crucial for the management and follow-up of the patients, even for effective clinical trials of potential therapies.

## DATA AVAILABILITY STATEMENT

The datasets for this article are not publicly available due to concerns regarding participant/patient anonymity. Requests to access the datasets should be directed to the corresponding authors.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the ethics committees of Children's Hospital of Fudan University. 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 minor(s)' legal guardian/next of kin for the publication of any potentially identifiable images or data included in this article.

## AUTHOR CONTRIBUTIONS

WZ, XW, and HW contributed to the conception and design. XP, HW, BW, and SX contributed to data acquisition and analysis, and drafting of the manuscript. YL, DZ, JS, JW, and XW confirmed the diagnosis, performed communication with patients and the follow-up. MG and XP performed the identification of pathogenic microbes. All authors gave final approval of the submitted version.

## FUNDING

This work was supported by Shanghai Science and Technology Innovation Action Plan (20Z11900600, Dr. Zhou) and Clinical Skill and Innovation program of Shanghai Shenkang Hospital Development Center (SHDC2020CR6028-002, Dr. Zhou).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.768000/full#supplementary-material

## REFERENCES

Blommaert, E., Péanne, R., Cherepanova, N. A., Rymen, D., Staels, F., Jaeken, J., et al. (2019). Mutations in MAGT1 lead to a Glycosylation Disorder with a Variable Phenotype. *Proc. Natl. Acad. Sci. USA* 116 (20), 9865–9870. doi:10.1073/pnas.1817815116

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a Flexible Trimmer for Illumina Sequence Data. *Bioinformatics* 30 (15), 2114–2120. doi:10.1093/bioinformatics/btu170

Brault, J., Meis, R. J., Li, L., Bello, E., Liu, T., Sweeney, C. L., et al. (2021). MAGT1 Messenger RNA-Corrected Autologous T and Natural Killer Cells for Potential Cell Therapy in X-Linked Immunodeficiency with Magnesium Defect, Epstein-Barr Virus Infection and Neoplasia Disease. *Cytotherapy* 23 (3), 203–210. doi:10.1016/j.jcyt.2020.08.013

Brigida, I., Chiriaco, M., Di Cesare, S., Cittaro, D., Di Matteo, G., Giannelli, S., et al. (2017). Large Deletion of MAGT1 Gene in a Patient with Classic Kaposi Sarcoma, CD4 Lymphopenia, and EBV Infection. *J. Clin. Immunol.* 37 (1), 32–35. doi:10.1007/s10875-016-0341-y

Chaigne-Delalande, B., Li, F.-Y., O'Connor, G. M., Lukacs, M. J., Jiang, P., Zheng, L., et al. (2013). Mg 2+ Regulates Cytotoxic Functions of NK and CD8 T Cells in Chronic EBV Infection through NKG2D. *Science* 341 (6142), 186–191. doi:10.1126/science.1240094

Chauvin, S. D., Price, S., Zou, J., Hunsberger, S., Brofferio, A., Matthews, H., et al. (2021). A Double-Blind, Placebo-Controlled, Crossover Study of Magnesium Supplementation in Patients with XMEN Disease. *J. Clin. Immunol.*. doi:10.1007/s10875-021-01137-w

Cherepanova, N. A., Shrimal, S., and Gilmore, R. (2014). Oxidoreductase Activity Is Necessary for N-Glycosylation of Cysteine-Proximal Acceptor Sites in Glycoproteins. *J. Cell Biol* 206 (4), 525–539. doi:10.1083/jcb.201404083

Dhalla, F., Murray, S., Sadler, R., Chaigne-Delalande, B., Sadaoka, T., Soilleux, E., et al. (2015). Identification of a Novel Mutation in MAGT1 and Progressive Multifocal Leucoencephalopathy in a 58-Year-Old Man with XMEN Disease. *J. Clin. Immunol.* 35 (2), 112–118. doi:10.1007/s10875-014-0116-2

Dimitrova, D., Rose, J. J., Uzel, G., Cohen, J. I., Rao, K. V., Bleesing, J. H., et al. (2019). Successful Bone Marrow Transplantation for XMEN: Hemorrhagic Risk Uncovered. *J. Clin. Immunol.* 39 (1), 1–3. doi:10.1007/s10875-018-0573-0

Dong, X., Liu, B., Yang, L., Wang, H., Wu, B., Liu, R., et al. (2020). Clinical Exome Sequencing as the First-Tier Test for Diagnosing Developmental Disorders Covering Both CNV and SNV: a Chinese Cohort. *J. Med. Genet.* 57 (8), 558–566. doi:10.1136/jmedgenet-2019-106377

Gotru, S. K., Gil-Pulido, J., Beyersdorf, N., Diefenbach, A., Becker, I. C., Vögtle, T., et al. (2018). Cutting Edge: Imbalanced Cation Homeostasis in MAGT1-Deficient B Cells Dysregulates B Cell Development and Signaling in Mice. *J.I.* 200 (8), 2529–2534. doi:10.4049/jimmunol.1701467

He, T. Y., Xia, Y., Li, C. G., Li, C. R., Qi, Z. X., and Yang, J. (2018). X-linked Immunodeficiency with Magnesium Defect, Epstein-Barr Virus Infection, and Neoplasia: Report of a Family and Literature Review. *Zhonghua Er Ke Za Zhi* 56 (1), 48–52. doi:10.3760/cma.j.issn.0578-1310.2018.01.013

Hoyos-Bachiloglu, R., Concha, S., Sepúlveda, P., Campos, R., Perez-Mateluna, G., King, A., et al. (2020). The Many Faces of XMEN Disease, Report of Two Patients with Novel Mutations. *J. Clin. Immunol.* 40 (2), 415–417. doi:10.1007/s10875-020-00746-1

Klinken, E. M., Gray, P. E., Pillay, B., Worley, L., Edwards, E. S. J., Payne, K., et al. (2020). Diversity of XMEN Disease: Description of 2 Novel Variants and Analysis of the Lymphocyte Phenotype. *J. Clin. Immunol.* 40 (2), 299–309. doi:10.1007/s10875-019-00732-2

Li, F.-Y., Chaigne-Delalande, B., Kanellopoulou, C., Davis, J. C., Matthews, H. F., Douek, D. C., et al. (2011). Second Messenger Role for Mg2+ Revealed by Human T-Cell Immunodeficiency. *Nature* 475 (7357), 471–476. doi:10.1038/nature10246

Li, F.-Y., Chaigne-Delalande, B., Su, H., Uzel, G., Matthews, H., and Lenardo, M. J. (2014). XMEN Disease: a New Primary Immunodeficiency Affecting Mg2+ Regulation of Immunity against Epstein-Barr Virus. *Blood* 123 (14), 2148–2152. doi:10.1182/blood-2013-11-538686

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

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The Sequence Alignment/Map Format and SAMtools. *Bioinformatics* 25 (16), 2078–2079. doi:10.1093/bioinformatics/btp352

Lindeboom, R. G. H., Vermeulen, M., Lehner, B., and Supek, F. (2019). The Impact of Nonsense-Mediated mRNA Decay on Genetic Disease, Gene Editing and Cancer Immunotherapy. *Nat. Genet.* 51 (11), 1645–1651. doi:10.1038/s41588-019-0517-5

Matsuda-Lennikov, M., Biancalana, M., Zou, J., Ravell, J. C., Zheng, L., Kanellopoulou, C., et al. (2019). Magnesium Transporter 1 (MAGT1) Deficiency Causes Selective Defects in N-Linked Glycosylation and Expression of Immune-Response Genes. *J. Biol. Chem.* 294 (37), 13638–13656. doi:10.1074/jbc.RA119.008903

Patiroglu, T., Haluk Akar, H., Gilmour, K., Unal, E., Akif Ozdemir, M., Bibi, S., et al. (2015). A Case of XMEN Syndrome Presented with Severe Auto-Immune Disorders Mimicking Autoimmune Lymphoproliferative Disease. *Clin. Immunol.* 159 (1), 58–62. doi:10.1016/j.clim.2015.04.015

Ravell, J. C., Chauvin, S. D., He, T., and Lenardo, M. (2020a). An Update on XMEN Disease. *J. Clin. Immunol.* 40 (5), 671–681. doi:10.1007/s10875-020-00790-x

Ravell, J. C., Matsuda-Lennikov, M., Chauvin, S. D., Zou, J., Biancalana, M., Deeb, S. J., et al. (2019b). Defective Glycosylation and Multisystem Abnormalities Characterize the Primary Immunodeficiency XMEN Disease. *J. Clin. Invest.* 130 (1), 507–522. doi:10.1172/JCI131116

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

Wang, Y., Wang, W., Liu, L., Hou, J., Ying, W., Hui, X., et al. (2018). Report of a Chinese Cohort with Activated Phosphoinositide 3-Kinase δ Syndrome. *J. Clin. Immunol.* 38 (8), 854–863. doi:10.1007/s10875-018-0568-x

Yang, L., Liu, X., Li, Z., Zhang, P., Wu, B., Wang, H., et al. (2020). Genetic Aetiology of Early Infant Deaths in a Neonatal Intensive Care Unit. *J. Med. Genet.* 57 (3), 169–177. doi:10.1136/jmedgenet-2019-106221

Zhou, H., and Clapham, D. E. (2009). Mammalian MagT1 and TUSC3 Are Required for Cellular Magnesium Uptake and Vertebrate Embryonic Development. *Pnas* 106 (37), 15750–15755. doi:10.1073/pnas.0908332106

**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.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

*Copyright © 2022 Peng, Lu, Wang, Wu, Gan, Xu, Zhuang, Wang, Sun, Wang and Zhou. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*

frontiers in Genetics

REVIEW
published: 26 January 2022
doi: 10.3389/fgene.2021.776807

# Primary Coenzyme Q10 Deficiency-7 and Pathogenic *COQ4* Variants: Clinical Presentation, Biochemical Analyses, and Treatment

*Jieqiong Xie*[1], *Jiayang Jiang*[1,2] *and Qiwei Guo*[1]*

[1]United Diagnostic and Research Center for Clinical Genetics, Women and Children's Hospital, School of Medicine and School of Public Health, Xiamen University, Xiamen, China, [2]School of Medicine, Huaqiao University, Quanzhou, China

Primary Coenzyme Q10 Deficiency-7 (COQ10D7) is a rare mitochondrial disorder caused by pathogenic *COQ4* variants. In this review, we discuss the correlation of *COQ4* genotypes, particularly the East Asian-specific c.370G > A variant, with the clinical presentations and therapeutic effectiveness of coenzyme Q10 supplementation from an exon-dependent perspective. Pathogenic *COQ4* variants in exons 1–4 are associated with less life-threating presentations, late onset, responsiveness to CoQ10 therapy, and a relatively long lifespan. In contrast, pathogenic *COQ4* variants in exons 5–7 are associated with early onset, unresponsiveness to CoQ10 therapy, and early death and are more fatal. Patients with the East Asian-specific c.370G > A variant displays intermediate disease severity with multi-systemic dysfunction, which is between that of the patients with variants in exons 1–4 and 5–7. The mechanism underlying this exon-dependent genotype-phenotype correlation may be associated with the structure and function of COQ4. Sex is shown unlikely to be associated with disease severity. While point-of-care high-throughput sequencing would be useful for the rapid diagnosis of pathogenic *COQ4* variants, whereas biochemical analyses of the characteristic impairments in CoQ10 biosynthesis and mitochondrial respiratory chain activity, as well as the phenotypic rescue of the CoQ10 treatment, are necessary to confirm the pathogenicity of suspicious variants. In addition to CoQ10 derivatives, targeted drugs and gene therapy could be useful treatments for COQ10D7 depending on the in-depth functional investigations and the development of gene editing technologies. This review provides a fundamental reference for the sub-classification of COQ10D7 and aim to advance our knowledge of the pathogenesis, clinical diagnosis, and prognosis of this disease and possible interventions.

**Keywords: COQ4, c.370G>A variant, phenotype-genotype correlation, primary coenzyme Q10 deficiency, exon**

OPEN ACCESS

***Edited by:***
*Kathleen M. Gorman, Temple Street Children's University Hospital, Ireland*

***Reviewed by:***
*Ann Saada, Hebrew University of Jerusalem, Israel*
*Ying Wang, McGill University, Canada*

***Correspondence:**
*Qiwei Guo*
*guoqiwei@xmu.edu.cn*

***Specialty section:***
*This article was submitted to Genetics of Common and Rare Diseases, a section of the journal Frontiers in Genetics*

***Received:*** *14 September 2021*
***Accepted:*** *14 December 2021*
***Published:*** *26 January 2022*

***Citation:***
*Xie J, Jiang J and Guo Q (2022) Primary Coenzyme Q10 Deficiency-7 and Pathogenic COQ4 Variants: Clinical Presentation, Biochemical Analyses, and Treatment. Front. Genet. 12:776807. doi: 10.3389/fgene.2021.776807*

## INTRODUCTION

Coenzyme Q (CoQ) is a lipophilic molecule composed of 1, 4-benzoquinone and a tail of isoprenoid units. The length of the isoprenoid tail is species specific, and in humans, it is 10 units long (CoQ10). CoQ10 is ubiquitously distributed in all cells and is mostly located in the mitochondrial inner membrane, where it plays an important role in the mitochondrial respiratory chain (MRC). The

oxidized form of CoQ10 (ubiquinone) can be reduced to ubiquinol by the addition of two electrons, which functions as an electron carrier that shuttles electrons from MRC complexes I (NADH: ubiquinone oxidoreductase) and II (succinate dehydrogenase) to complex III (decylubiquinol cytochrome c oxidoreductase) (Turunen, et al., 2004; Alcazar-Fabra, et al., 2016; Wang and Hekimi 2016). In addition, CoQ10 has been suggested to be involved in several other biological processes, such as pyrimidine biosynthesis (Evans and Guy 2004), $\beta$-oxidation of fatty acids (Watmough and Frerman 2010), structural stabilization of other MRC complexes (Cramer, et al., 2011; Letts and Sazanov 2017), production of reactive oxygen species (Guaras, et al., 2016), and inhibition of ferroptosis (Bersuker, et al., 2019; Doll, et al., 2019).

Although some CoQ10 can be obtained through dietary sources, due to its poor distribution and bioavailability, endogenous synthesis is the major source of CoQ10 (Stefely and Pagliarini 2017). Primary CoQ10 deficiency, which is defined as reduced levels of CoQ10 in tissues due to impairment of CoQ10 biosynthesis, results in a group of rare, clinically heterogeneous disorders presenting as multisystem manifestations (Desbats, et al., 2015; Quinzii, et al., 2014). To date, more than 10 genes (*ADCK3*, *PDSS1*, *PDSS2*, *COQ2*, *COQ3*, *COQ4*, *COQ5*, *COQ6*, *COQ7*, *COQ8A*, *COQ8B*, *COQ9*, *COQ10A*, and *COQ10B*) have been suggested to be involved in human CoQ10 biosynthesis (Shalata, et al., 2019; Brea-Calvo et al., 2021). Theoretically, pathogenic variants in any of these genes would result in primary CoQ10 deficiency. Based on the predicted prevalence of CoQ10 deficiency derived from the allelic frequencies of pathogenic variants, there are >120,000 individuals with CoQ10 deficiency, and most of these are undiagnosed (Hughes, et al., 2017). The clinical profile of primary CoQ10 deficiency is very complex, as there are not only pathogenic variants in different causative genes but also different pathogenic variants in the same causative gene, which result in highly heterogeneous manifestations with different ages of onset and outcomes (Alcazar-Fabra, et al., 2021). An exhaustive profile of primary CoQ10 deficiency has been recently reviewed elsewhere (Alcazar-Fabra, et al., 2021).

Among the causative genes of primary CoQ10 deficiency, *COQ4* (MIM 612898) is relatively newly confirmed, and its exact function remains largely unknown. *COQ4* is located on chromosome 9q34.11 and encodes a ubiquitously expressed 265-amino acid COQ4 protein, which is mainly localized to the matrix side of the mitochondrial inner membrane (Casarin, et al., 2008). Based on the finding using a yeast model, it has been speculated that COQ4 is essential for the stabilization of a multi-heteromeric complex containing several CoQ10 biosynthetic enzymes rather than functioning as a catalytic enzyme (Marbois, et al., 2009). According to the literature, the first patient with a *COQ4* variant and primary CoQ10 deficiency was reported by Salviati et al.(2012), in 2012; they reported a boy with a severe encephalomyopathic disorder carrying a *de novo* heterozygous 3.9 Mb deletion affecting at least 80 genes, including *COQ4*. Thus, haploinsufficiency of *COQ4* was considered to be the cause of CoQ10 deficiency in this patient (Salviati, et al., 2012). In 2015, Brea-Calvo et al. (2015) and Chung et al. (2015) further validated the association between *COQ4* variants and primary CoQ10 deficiency. Their studies demonstrated that the pathogenicity of *COQ4* variants has an autosomal recessive mode of inheritance (Chung, et al., 2015; Brea-Calvo, et al., 2015). Subsequently, the primary CoQ10 deficiency caused by *COQ4* variants was termed primary CoQ10 deficiency-7 (COQ10D7). With the increasing use of clinical whole exome sequencing (WES), more COQ10D7 patients with various phenotypes and causative variants have been found worldwide (Finsterer and Zarrouk-Mahjoub 2017; Romero-Moya et al., 2017a; Romero-Moya, et al., 2017b; Sondheimer, et al., 2017; Bosch, et al., 2018; Caglayan, et al., 2019; Ge, et al., 2019; Ling, et al., 2019; Lu, et al., 2019; Yu, et al., 2019; Chen, et al., 2020; Tsang, et al., 2020; Hashemi, et al., 2021; Mero, et al., 2021). In 2019, we reported a Chinese family with a c.370G > A *COQ4* variant along with an associated Leigh syndrome phenotype, the results of biomedical analyses, and intervention (Lu, et al., 2019). The c.370G > A variant has been shown to be East Asian population-specific, with an allele frequency of 0.001504 (gnomAD database), and was validated as a founder variant in the southern Chinese population, which was inherited from a common ancestor approximately 27 generations ago (Yu, et al., 2019). At the time this review was prepared (1 June, 2021), 28 *COQ4* variants had been reported in 38 patients, and 15 of these patients (~40%) carried the c.370G > A variant, highlighting the high prevalence of this founder variant. Although the current number of reported patients is limited, these data will enable us to take a small step towards creating a comprehensive profile for this monogenic disease. In this review, we focus on the correlation of *COQ4* genotypes, particularly the East Asian-specific c.370G > A variant, with the clinical presentations and therapeutic effectiveness of treatments from an exon-dependent perspective, and discuss the associated biochemical analyses with an aim to advance our knowledge of the pathogenesis, clinical diagnosis, prognosis, and intervention of this disease.

## CLINICAL PRESENTATION OF COQ10D7

Similar to that of primary CoQ10 deficiency, the phenotypic spectrum of COQ10D7 is wide. We categorized the reported patients into three cohorts based on the location of their pathogenic variants: cohort 1, which includes nine patients carrying variants in exons 1–4 in both alleles except for those carrying the c.370G > A variant (**Figure 1A** and **Table 1**); cohort 2, which includes 14 patients carrying variants in exons 5–7 (**Figure 1B** and **Table 2**); and cohort 3, which consists of 15 patients carrying the c.370G > A variant (**Figure 1C** and **Table 3**). It should be noted that a compound heterozygous patient (c.370G > A/c.533G > A) was analyzed in both cohorts 2 and 3, and a patient who carried a *de novo* heterozygous complete deletion of *COQ4* was not included in any cohort (Salviati, et al., 2012).

### Presentations in Patients With *COQ4* Variants in Exons 1–4 in Both Alleles

Three of the nine (33.3%) patients in cohort 1 (**Table 1**) had neonatal onset (onset within 1 month of age), and two (66.7%) of these had infantile death (death between 1 month and 1 year of age). One patient with neonatal onset and infantile death carried a frameshift variant

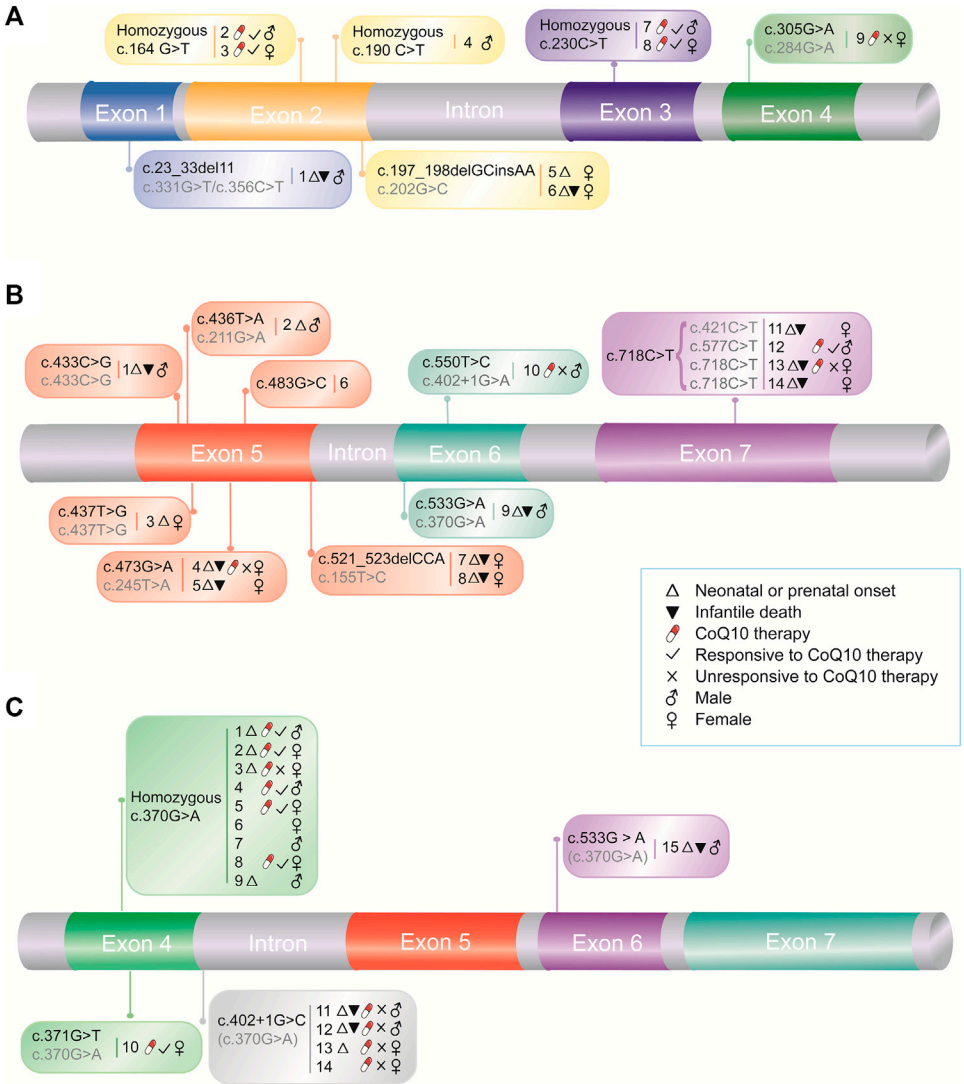



**FIGURE 1 |** Pathogenic *COQ4* variants and associated medical information. Medical information for patients with *COQ4* variants in exons 1–4 **(A)** and exons 5–7 **(B)**, and with East Asian-specific c.370G > A variant **(C)**, respectively. The numbers indicate case numbers in **Tables 1**, **2**, and **3**. The second alleles were indicated in grey.

(c.23_33del11) with a loss of function. In contrast, five other patients (5/9, 55.6%) had late onset (at age 4–8 years) and were alive at 14–28 years, suggesting a relatively mild manifestation of the disease. The most frequent presentations in this cohort were less life-threating and included central nervous system (CNS) symptoms, such as seizures (8/9, 88.9%), ataxia (6/9, 66.7%), and cognitive defects (5/9, 55.6%). In contrast, lactic acidosis (2/9, 22.2%), cardiomyopathy (2/9, 22.2%), respiratory distress or failure (3/9, 33.3%), and hypotonia (3/9, 33.3%) were the presentations of the patients who had early disease onset and death. Imaging indicated that cerebral or cerebellar abnormalities were common (8/9, 88.9%) in this cohort (**Table 1**). There was no sex bias in this cohort, as the female-to-male ratio was 5 to 4 (**Table 1**).

## Presentations in Patients With *COQ4* Variants in Exons 5–7

In contrast to cohort 1, 11 of the 14 (78.6%) patients in cohort 2 (**Table 2**) had neonatal onset, and 9/11 (81.8%) died in the first year of life, suggesting greater disease severity. The most frequent presentations in this cohort were multi-systemic, including severe lactic acidosis (10/14, 71.4%), respiratory failure or distress (11/14, 78.6%), cardiomyopathy (9/14, 64.3%), and hypotonia (7/14, 50.0%). Similar to the findings in cohort 1, these presentations were associated with early onset and death, suggesting that they are more life-threatening. CNS symptoms, such as seizures (5/14, 35.7%), ataxia (2/14, 14.3%), and cognitive defects (2/14, 14.3%) were less frequent, probably because these presentations are relatively late-onset and thus undetectable due to early death. The main imaging findings were cerebral and cerebellar abnormalities (7/

**TABLE 1 |** Clinical presentation, treatment, outcomes, and brain findings for patients with *COQ4* variants in exons 1-4.

| Case number | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| Reference | Sondheimer et al. (2017) | Caglayan et al. (2019) | Caglayan et al. (2019) | Brea-Calvo et al. (2015) | Chung et al. (2015) | Chung et al. (2015) | Bosch et al. (2018) | Bosch et al. (2018) | Mero et al. (2021) |
| *COQ4* variants | c.23_33del11/ c.331G>T/ c.356C>T | Homozygous c.164 G>T | Homozygous c.164 G>T | Homozygous c.190C>T | c.197_198delGCinsAA/ c.202G>C | c.197_198delGCinsAA/ c.202G>C | Homozygous c.230C>T | Homozygous c.230C>T | c.284G>A/ c.305G>A |
| Affected exons | 1,4 | 2 | 2 | 2 | 2 | 2 | 3 | 3 | 3, 4 |
| Demography | Unclear | Turkey | Turkey | Italian | Caucasian | Caucasian | Unclear | Unclear | Unclear |
| Sex | Male | Male | Female | Male | Female | Female | Male | Female | Female |
| **Clinical presentations** | | | | | | | | | |
| Onset age | At birth | 8 years | 8 years | 10 months | At birth | Before birth | 4 years | 9 years | 5 years |
| Seizures | √ | √ | √ | √ | √ | | √ | √ | √ |
| Ataxia | | √ | √ | √ | | | √ | √ | √ |
| Cognitive defect | | √ | √ | √ | | | √ | | √ |
| Spasticity | | √ | √ | √ | | | | | |
| Oromotor dysfunction | | | | √ | √ | √ | | | |
| Respiratory distress | √ | | | | √ | √ | | | |
| Hypotonia | √ | | | | √ | √ | | | |
| Developmental delay | | | | | √ | | √ | | |
| Lactic acidosis (lactate level) | √(4-6 mM) | | | | | √(6.2 mM) | | | |
| Cardiomyopathy | √ | | | | √ | | | | |
| Other symptoms | Hearing loss | | | Scoliosis | | | | | |
| Brain findings | First week: focal regions of cortical increased T1 signal and MRS identified enlarged lactate peaks. Tenth week: microcephaly with volume loss and increasing prominence of lactate peaks. | Age 26 years: brain MRI showed cortical and subcortical T2 hyperintensity, not limited to a specific vascular territory. | Age 27 years: brain MRI showed cerebral and cerebellar atrophy. | Age 12 years: MRI showed bilateral increased signal intensity in FLAIR and T2W sequencing in both occipital cortical and juxtacortical areas; Electrophysiological examination showed a sensory motor polyneuropathy with slowed conduction velocities. Age 17 years: MRI showed cerebellar atrophy, widened ventricular space, scars from cortical necrotic lesions in both occipital areas. | First day: cerebellar hypoplasia, prominent extra axial space in posterior fossa and mild lateral ventricle enlargement; EEG showed multiple areas of amplitude suppression, suggesting a generalised encephalopathy and multifocal-onset seizures. | Fetal MRI: normal intracranial anatomy, transverse cerebellar diameter 10–15th percentile. Age 2 days: MRI showed decreased cerebellar hemisphere volume. | Age 5 years: brain MRI revealed a suspected tectal glioma (treated with radiotherapy) | Age 10 years: cavernoma in the left parietal lobe. Age 13 y: lesion at left occipital lobe with clear diffusion restriction. | Age 15 years: brain and spine MRI showed discrete enlargement of the pericerebellar sulci without focal abnormalities in supratentorial structures. |

(Continued on following page)

**TABLE 1 |** (*Continued*) Clinical presentation, treatment, outcomes, and brain findings for patients with *COQ4* variants in exons 1-4.

| Case number | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| **Treatment** | | | | | | | | | |
| Symptomatic treatment | Phenobarbital, topiramate, clobazam | Carbamazepine, clonazepam | Levatiracetam | | | | | Carbamazepine | Valproate |
| CoQ10 supplement | | 2000 mg/day | 2000 mg/day | | | | 1000 mg/day | 1000 mg/day | ubiquinol, 100 mg/kg/day |
| Age at initiating CoQ10 supplement | | 26 years | 27 years | | | | 13 years | 11 years | 19 years |
| Responsiveness to CoQ10 therapy | | Marked improvement in the SARA score. | An improvement in the SARA score. | | | | Walk test stable over the period of a year. | Walk test stable over the period of a year. | No improvement in the SARA score. |
| **Outcomes** | Died of acidosis and respiratory failure at the age of 4 months. | Alived at the age of 27 years. | Alived at the age of 28 years. | Alived at the age of 18 years. | Died of cardiorespiratory failure at the age of 19 months. | Died at the age of 10 weeks. | Alived at the age of 15 years. | Alived at the age of 14 years. | Alived at the age of 19 years. |

*SARA, the scale for the assessment and rating of ataxia; MRI, magnetic resonance imaging; MRS, magnetic resonance spectroscopy; EEG, electroencephalograph.*

77.8%). In rare cases, abnormalities in the basal ganglia (1 case) and brainstem (1 case) were noted (**Table 2**). The female-to-male ratio in this cohort was 9 to 5, indicating a sex bias (**Table 2**), which suggests that male patients have greater disease severity and may have a higher prenatal lethality rate. However, seven of the nine patients with infantile death were females, suggesting that females were more sensitive to the disease (**Table 2**). Therefore, whether sex contributes to disease severity remains to be further investigated.

## Presentations in Patients With the East Asian-specific *COQ4* c.370G > A Variant

Nine patients in cohort 3 (**Table 3**) were homozygous for c.370G > A, and six patients were compound heterozygotes. Four of the nine (44.4%) homozygous patients had neonatal onset, but only 1 (25%) had infantile death. Among the five patients who were alive at the last follow-up, the oldest patient was ~8 years old. The six compound heterozygous patients carried the c.370G > A variant with c.371G > T, c.402+1G > C, and c.533G > A variants. The c.371G > T variant resulted in a different amino acid change (p.Gly124Val) at the same codon as the c.370G > A variant (p.Gly124Ser). The patient with this genotype had neonatal onset and was alive at 3 years of age, which was the last follow-up. Similar to the patient with a loss of function frameshift variant (c.23_33del11) in cohort 1, three of four (75%) patients with the splicing variant c.402+1G > C had neonatal onset, and two (66.7%) of them had infantile death, suggesting that this splicing variant could be a loss of function variant, which is associated with greater disease severity. As was observed in cohort 2, the patient with the compound heterozygous c.370G > A/c.533G > A variants had neonatal onset and infantile death. The most prevalent presentations among the c.370G > A homozygotes and heterozygotes were similar and included multi-systemic symptoms, such as lactic acidosis (14/15, 93.3%), developmental delay (13/15, 86.9%), seizures (12/15, 80.0%), respiratory distress (10/15, 66.7%), cardiomyopathy (9/15, 60.0%), hypotonia (9/15, 60.0%), and dystonia (5/15, 33.3%). Similar to the variation in clinical presentations, imaging showed more diverse abnormalities in this cohort than in cohorts 1 and 2 (**Table 3**). The frequently affected tissues included the cerebrum (7/12, 58.3%), cerebellum (6/12, 50.0%), and corpus callosum (5/12, 41.7%). Noticeably, abnormalities in the basal ganglia and midbrain were also noted in five (5/12, 41.7%) patients, including two c.370G > A homozygotes and three c.370G > A heterozygotes, suggesting that a specific Leigh syndrome-like phenotype is correlated with this genotype (**Table 3**). Therefore, it can be concluded that the East Asian-specific c.370G > A variant conferred intermediate disease severity that was between the severity of the patients with variants in exons 1–4 and 5–7. Lastly, similar to cohort 1, the female-to-male ratios for c.370G > A homozygous and heterozygous patients were 5 to 4 and 3 to 3, respectively, indicating no sex bias.

# BIOCHEMICAL ANALYSES OF PATIENTS WITH COQ10D7

Elevated lactate levels are frequently detected in the plasma or serum of patients with primary CoQ10 deficiency, including

**TABLE 2 |** Clinical presentation, treatment, outcomes, and brain findings for patients with *COQ4* variants in exons 5-7.

| Case number | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| References | Brea-Calvo et al. (2015) | Ge et al. (2019) | Hashemi et al. (2021) | Chung et al. (2015) | Chung et al. (2015) | Romero-Moya et al. (2017a) | Brea-Calvo et al. (2015) | Brea-Calvo et al. (2015) |
| *COQ4* variants | Homozygous c.433C>G | c.211G>A/c.436T>A | Homozygous c.437T>G | c.245T>A/c.473G>A | c.245T>A/c.473G>A | c.483G>C | c.155T>C/ c.521_523delCCA | c.155T>C/ c.521_523delCCA |
| Affected exons | 5 | 3, 5 | 5 | 3, 5 | 3, 5 | 5 | 2, 5 | 2, 5 |
| Demography | Italian | Chinese | Iranian | Caucasian–Hispanic ancestry | Caucasian–Hispanic ancestry | Unclear | Austrian | Austrian |
| Sex | Male | Male | Female | Female | Female | Female | Female | Female |
| **Clinical presentations** | | | | | | | | |
| Onset age | At birth | At birth | At birth | At birth | At birth | Before 4 years | Before birth | Before birth |
| Lactic acidosis (lactate level) | √(20.1 mM) | | √ | √(19.5mM) | √(22 mM) | | √(6.4-14 mM) | √(3.5-9 mM) |
| Respiratory distress | √ | √ | √ | √ | √ | | √ | √ |
| Cardiomyopathy | √ | √ | | √ | √ | | | |
| Hypotonia | √ | | √ | √ | | | | |
| Seizures | | √ | √ | √ | | | √ | √ |
| Oromotor dysfunction | | √ | √ | | | | | |
| IUGR | | | | | | | | |
| Developmental delay | | √ | | | | | | |
| Ataxia | | | √ | | | | | |
| Cognitive defect | | | √ | | | | | |
| Other symptoms | Areflexia, acrocyanosis, bradycardia | | Visual impairment, hearing loss, sensorimotor polyneuropathy | | | Severe metabolic, mitochondrial defects | Distal arthrogryposis | |
| Brain findings | | Age 3 months: head MRI showed bilateral brain atrophy, bilateral occipital parietal lobes, bilateral basal ganglia, bilateral cerebro-foot small patches DWI hyperintensity (cytotoxic edema); white matter myelin development lags behind the same age. | Age 9 years: brain MRI revealed cerebellar atrophy, and nonspecific white matter abnormal signal intensities on T2 weighted and FLAIR sequences. | At birth: brain MRI showed small cerebellar size and diffuse T2 white matter hyperintensity, and MR spectroscopy demonstrated decreased NAA and a lactate peak. | | | At birth: USG brain showed cerebellar hypoplasia. Autopsy: severe olivopontocerebellar and thalamic hypoplasia and scattered cavitations in the white matter. | At birth: a cranial ultrasound confirmed cerebellar hypoplasia. |

(Continued on following page)

**TABLE 2 |** (*Continued*) Clinical presentation, treatment, outcomes, and brain findings for patients with *COQ4* variants in exons 5-7.

| **Case number** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** |
|---|---|---|---|---|---|---|---|---|
| **Treatment** | | | | | | | | |
| Symptomatic treatment | Dobutamine infusion | | | Carnitine, thiamine, riboflavin, biotin, hydroxocobalamin, dopamine, milrinone, phenobarbital, fosphenytoi, levetiracetam, clobazam. | Antiepileptics | | | |
| CoQ10 supplement | | | | Intravenous, 20 mg/kg/day. | | | | |
| Age at initiating CoQ10 supplement | | | | 2 days | | | | |
| Responsiveness to CoQ10 therapy | | | | No significant improvement. | | | | |
| **Outcomes** | Died 4 hours after birth. | Unknown | Alived at the age of 9 years. | Died of acidosis and respiratory failure at the age of 2 months. | Died of acidosis and respiratory failure at the age of 2 days. | Died of lethal rhabdomyolysis at the age of 4 years. | Died of acidosis and multiorgan failure at the age of 3 days. | Died of lactic acidosis at the age of 2 days. |

| **Case number** | **9** | **10** | **11** | **12** | **13** | **14** |
|---|---|---|---|---|---|---|
| Reference | Ling et al. (2019) | Yu et al. (2019) | Brea-Calvo et al. (2015) | Mero et al. (2021) | Chung et al. (2015) | Chung et al. (2015) |
| *COQ4* variants | c.370G>A/ c.533G > A | c.550T>C/c.402+1G>A | c.421C>T/c.718C>T | c.577C>T/c.718C>T | Homozygous c.718C>T | Homozygous c.718C>T |
| Affected exons | 4, 6 | 6, unknown | 5, 7 | 6, 7 | 7 | 7 |
| Demography | Chinese | Chinese | Japanese | Unclear | Ashkenazi Jewish | Ashkenazi Jewish |
| Sex | Male | Male | Female | Male | Female | Female |
| **Clinical presentations** | | | | | | |
| Onset age | Before birth | Before 8 months | Before birth | Before 30 months | At birth | Before birth |
| Lactic acidosis (lactate level) | √ | √(2.5-5.9 mM) | √(11.2-18.8 mM) | | √(2.4 mM) | |
| Respiratory distress | √ | | √ | | √ | √ |
| Cardiomyopathy | √ | | √ | √ | √ | √ |
| Hypotonia | | √ | | √ | √ | √ |
| Seizures | | | | | | |
| Oromotor dysfunction | | √ | | | √ | |
| IUGR | √ | | √ | | | √ |
| Developmental delay | | √ | | √ | | |
| Ataxia | | | | √ | | |
| Cognitive defect | | | | √ | | |
| Other symptoms | | Microcephaly, dystonia, cortical visual impairment | | | Apnoea, bradycardia | |
| Brain findings | | Age 21 days, 40 days, and 1 year 4 months: mild cerebral atrophy with bilateral frontal predominance. | | Age 22 months and 4 years: brain MRI scans were normal. | MRS showed elevated lactate and depressed NAA peaks; EEG showed burst suppression. | An EEG was abnormal with a burst suppression pattern. Autopsy: cerebellar and brainstem hypoplasia and microdysgenesis. |

(Continued on following page)

**TABLE 2 |** *(Continued)* Clinical presentation, treatment, outcomes, and brain findings for patients with *COQ4* variants in exons 5-7.

| Case number | 9 | 10 | 11 | 12 | 13 | 14 |
|---|---|---|---|---|---|---|
| **Treatment** | | | | | | |
| Symptomatic treatment | Caffeine | | | | Pyridoxal phosphate, folinic acid, and riboflavin | Ampicillin, gentamicin, bicarbonate, dopamine, and dobutamine |
| CoQ10 supplement | | √ | | Oral, ubiquinol, 30 mg/kg/day | 15 mg/kg/day divided into two doses | |
| Age at initiating CoQ10 supplement | | | | 3 years | 1 month | |
| Responsiveness to CoQ10 therapy | | No significant improvement. | | Improvement in the neuromuscular symptoms and cardiac health. | No significant improvement. | |
| **Outcomes** | Died at the age of 28 days. | Alived at the age of 3 years 7 months. | Died of lactic acidosis and heart failure at the age of 2 days. | Alived at the age of 4 years. | Died at the age of 7 weeks. | Died of cardiopulmonary arrest at the age of 4 days. |

*IUGR, intrauterine growth retardation; DWI, diffusion-weighted imaging; FLAIR, fluid attenuated inversion recovery; MRI, magnetic resonance imaging; MRS, magnetic resonance spectroscopy; MR, magnetic resonance; EEG, electroencephalograph; NAA, N-acetylaspartic acid; USG, ultrasonography.*

COQ10D7, indicating metabolic impairment. As shown in **Tables 1**, **2**, and **3**, elevated levels of lactate were observed in the three cohorts. However, greater elevation of lactate was observed in cohorts 2 and 3, in accordance with the frequent manifestation of lactic acidosis, early onset, and worse outcomes among patients in cohorts 2 and 3. In contrast, data on plasma or serum CoQ10 levels are limited. However, data from three patients indicated relatively low CoQ10 levels in serum samples (Caglayan, et al., 2019; Yu, et al., 2019).

In addition to plasma and serum biomarkers, MRC activity was evaluated in several patients and was considered to be a diagnostic indicator of primary CoQ10 deficiency. Generally, biochemical analysis of MRC activity is performed using a muscle biopsy or cultured skin fibroblasts. Currently, the correlations between this biochemical analysis and most *COQ4* genotypes are unknown because data on the MRC activity of specific genotypes are not only limited but also derived from different sample types. However, the MRC activity profiles of patients with COQ10D7 are traceable according to sample type. Generally, reduced CoQ10 levels were observed in both muscle biopsies and fibroblasts from patients with COQ10D7. MRC complex activity testing results derived from fibroblasts were highly consistent among patients, and complex II + III activity was characteristically decreased (**Table 4**). These results were reasonable because the shuttling of electrons from complexes I and II to complex III was most likely affected by CoQ10 deficiency. Complex I + III activity was not evaluated in most cultured cells, as the results would be unreliable (Spinazzi, et al., 2012). In contrast, MRC complex activity levels in muscle biopsies varied among patients, although reduced complex II + III activity was also common (**Table 5**). Interestingly, elevated levels of complexes I, III, and IV were reported in frozen post-mortem muscle samples. These results indicate that reduced levels of CoQ10 and complex II + III activity in muscle or fibroblast samples are hallmarks of COQ10D7, and these biochemical changes were more specific in fibroblasts than in muscle samples, which is consistent with the conclusion of a previous study (Montero, et al., 2008).

## TREATMENT OF COQ10D7

Early oral supplementation with high-dose CoQ10 has been reported to improve a wide spectrum of phenotypes of primary CoQ10 deficiency (Montini, et al., 2008; Hargreaves 2014). Therefore, some patients with COQ10D7 were treated with CoQ10 therapy, in the form of ubiquinone or ubiquinol, and various responses were reported.

### Treatment of Patients With *COQ4* Variants in Exons 1–4 in Both Alleles

In this cohort (**Figure 1A**), five patients initiated CoQ10 therapy at different ages and at different doses (1,000–~5,000 mg/day), and four (80%) showed positive effects, including improvements

**TABLE 3 |** Clinical presentation, treatment, outcomes, and brain findings for patients with East Asian-specific *COQ4* c.370G > A variant.

| Case number | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| Reference | Ling et al. (2019) | Yu et al. (2019) | Yu et al. (2019) | Yu et al. (2019) | Yu et al. (2019) | Yu et al. (2019) | Lu et al. (2019) | Lu et al. (2019) |
| *COQ4* variants | Homozygous c.370G>A | Homozygous c.370G>A | Homozygous c.370G>A | Homozygous c.370G>A | Homozygous c.370G>A | Homozygous c.370G>A | Homozygous c.370G>A | Homozygous c.370G>A |
| Affected exons | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 |
| Demography | Chinese | Chinese | Chinese | Chinese | Chinese | Chinese | Chinese | Chinese |
| Sex | Male | Female | Female | Male | Female | Female | Male | Female |
| **Clinical presentations** | | | | | | | | |
| Onset age | 1 month | At birth | At birth | 2 months | 2 months | 4 months | 2 months | 2 months |
| Lactic acidosis (lactate level) | √(4.1 mM) | √(24 mM) | √(2.4–3.2 mM) | √ | √ | | √ (15.1 mM) | √ (17.6 mM) |
| Developmental delay | √ | √ | √ | √ | √ | √ | √ | √ |
| Seizures | √ | √ | | √ | √ | √ | √ | √ |
| Respiratory distress | √ | √ | | | √ | | √ | √ |
| Cardiomyopathy | | √ | | | √ | √ | √ | |
| Hypotonia | | √ | | | √ | √ | | |
| Dystonia | | | √ | | | | √ | √ |
| Oromotor dysfunction | | | √ | | | √ | √ | √ |
| Hearing impairment | √ | | | | | | √ | |
| Ophthalmic impairment | | | √ | | √ | | | √ |
| IUGR | | | | | | | | |
| Weak responsiveness | | | | | | | √ | √ |
| Spasticity | | | √ | | | √ | | |
| Ataxia | | | | | | | √ | √ |
| Brain findings | Age 4 months: MRI showed cerebral atrophy. The auditory brainstem response test showed features of conduction delay in the central auditory pathway. The nerve conduction study showed features of early demyelinating motor neuropathy. Age 2 years: MRI showed rapid progression of bifrontal cerebral atrophy, with the involvement of bifrontal cortical grey matter and white matter. | Age 7 weeks: mild cerebellar hypoplasia, mild thinning of corpus callosum. | Age 6 months: severe cerebral atrophy. | Age 32 months: moderate cerebellar atrophy without isolated vermian hypoplasia, cerebral atrophy, symmetrical loss of cerebral white matter particularly in bilateral frontal and anterior temporal regions. Corpus callosum was thinned, basal ganglia and pons unremarkable. | Age 14 months: mild thinning of corpus callosum. | Age 1 year 2 months: mild cerebellar atrophy and cerebral atrophy, white matter cystic changes with bilateral frontal and anterior temporal predominance. Corpus callosum thinning, preserved basal ganglia and brainstem. | Age 2 months: slightly widened frontal and temporal lobe. Age 5 months 18 days: bilateral, symmetrical lesions in the midbrain. | Age 1 month: normal. Age 4 months: lesions in midbrain and basal ganglia. Age 3 year 8 months: symmetrical, patchy, low density shadow in bilateral basal ganglia and diffuse brain atrophy. |

(Continued on following page)

**TABLE 3 |** (*Continued*) Clinical presentation, treatment, outcomes, and brain findings for patients with East Asian-specific *COQ4* c.370G > A variant.

| Case number | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| **Treatment** | | | | | | | | |
| Symptomatic treatment | Phenobarbital, multiple antiepileptics | Multiple anticonvulsants | | | Multiple anticonvulsants | | Phenobarbital, levetiracetam | Thiamine, levetiracetam, phenobarbitone, carnitine |
| CoQ10 supplement | 200 mg, 3 times a day | Intravenous | √ | √ | 30 mg/kg/day | | | Oral, 50 mg/kg/day |
| Age at initiating CoQ10 supplement | 6 years | 22 days | 2 years | 7 years | 11 months | | | 12 months |
| Responsiveness to CoQ10 therapy | A subjective improvement of the patient's awareness to the surroundings was the major observation. | Cardiac function stable. | No significant improvement. | Stable condition. | Some improvement in seizure control and development. | | | Seizures, screaming, and respiratory distress improved. No evident improvements in nystagmus, dystonia, psychomotor development, or ambulation. |
| **Outcome** | Died of recurrent central apnea, aspiration pneumonia, and respiratory failure at the age of 6 years. | Alived at the age of 9 months. | Died at the age of 3 years 6 months. | Alived at the age of 8 years. | Alived at the age of 1 year 6 months. | Die of sepsis at the age of 1 year 8 months. | Died at the age of 5 months. | Alived at the age of 3 years 8 months. |

| Case number | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|
| Reference | Chen et al. (2020) | Yu et al. (2019); Ling et al. (2019) | Yu et al. (2019) | Yu et al. (2019) | Yu et al. (2019) | Yu et al. (2019) | Ling et al. (2019) |
| *COQ4* variants | Homozygous c.370G>A | c.370G>A/c.371G>T | c.370G>A/ c.402+1G>C | c.370G>A/ c.402+1G>C | c.370G>A/ c.402+1G>C | c.370G>A/ c.402+1G>C | c.370G>A/ c.533G>A |
| Affected exons | 4 | 4 | 4, unknown | 4, unknown | 4, unknown | 4, unknown | 4, 6 |
| Demography | Chinese | Chinese | Chinese | Chinese | Chinese | Chinese | Chinese |
| Sex | Male | Female | Male | Male | Female | Female | Male |
| **Clinical presentations** | | | | | | | |
| Onset age | 1 month | 6 months | Before birth | At birth | Before birth | Before birth | Before birth |
| Lactic acidosis (lactate level) | √ | √ (2.3 mM) | √(28.36 mM) | √ (2.6 mM) | √ (10 mM) | √ | √(3.6 mM) |
| Developmental delay | √ | √ | √ | | √ | √ | |
| Seizures | √ | √ | √ | | √ | √ | |
| Respiratory distress | | | √ | √ | √ | √ | √ |
| Cardiomyopathy | | | √ | √ | √ | √ | √ |
| Hypotonia | √ | √ | √ | | √ | √ | √ |
| Dystonia | √ | √ | | | | | |
| Oromotor dysfunction | √ | | | | √ | | |
| Hearing impairment | √ | | √ | | | | |
| Ophthalmic impairment | √ | √ | √ | | | | |
| IUGR | | | √ | | √ | √ | √ |
| Weak responsiveness | √ | | | | | | |
| Spasticity | | √ | | | | | |
| Ataxia | | | | | | | |

(Continued on following page)

**TABLE 3 |** (*Continued*) Clinical presentation, treatment, outcomes, and brain findings for patients with East Asian-specific *COQ4* c.370G > A variant.

| Case number | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|
| **Brain findings** | Age 3 months: MRI showed increased T1 and T2 signal in the white matter of the right parietal lobe, and displasia of white matter myelin sheath; EEG showed multifocal-onset seizures. | (Yu, et al. 2019): mild cerebral and cerebellar hypoplasia. Age 35 months:small focus of T2 and FLAIR hyperintensity at the left lentiform nucleus. Age 6 months: MRS showed raised lactate peaks at bilateral basal ganglia and frontal white matter, which normalized at the age of 7 months (Ling, et al. 2019): EEG showed potential epileptogenic discharge in the background, which was normalized after one month. MRI of the brain showed prominent subarachnoid and sulcal spaces of the cerebral hemispheres, which are consistent with microcephaly. Brain MRS revealed a lactate peak signal of uncertain significance. | Age 3 weeks: MRI showed symmetrical T1 and T2B hyperintensity with restricted diffusion at bilateral lentiform nuclei. Foci of restricted diffusion were also detected at bilateral frontal white matter. MRS showed raised lactate peaks at bilateral basal ganglia and cerebral white matter. Subsequent follow-up MRI showed established infarcts with cystic changes at bilateral lentiform nuclei. Mild cerebellar hypoplasia was also noted. | | Neonatal stage: symmetrical T1 hyperintensity at bilateral basal ganglia, mild cerebellar hypoplasia, later with generalized progressive cerebellar and cerebral atrophy with diffuse white matter loss, thinning of corpus callosum. Cystic changes within cerebral white matter, bilateral basal ganglia, and thalami. MRS showed raised lactate peaks at bilateral basal ganglia. Age 9 months: MRI showed generalized progressive cerebellar and cerebral atrophy, with diffuse white matter loss including thinning of the corpus callosum. Cystic changes were seen within the cerebral white matter, bilateral basal ganglia, and thalami. | | |
| **Treatment** | | | | | | | |
| Symptomatic treatment | Phenobarbital, levetiracetam | Steroids, levetiracetam | Phenobarbitone, levetiracetam | Carnitine | | | Caffeine |
| CoQ10 supplement | Oral, 30 mg/kg/day. | Ubiquinol, 250 mg/day; then increased to 400 mg/day. After a year, the regimen was switched to liquid liposomal ubiquinol at a dose of 100 mg/day. | 40 mg/kg/day | 15 mg/kg/day | √ | √ | |
| Age at initiating CoQ10 supplement | 5 months | 9 months | 5 months | 2 days | 4 years 5 months | 1 years | |
| Responsiveness to CoQ10 therapy | Unclear | Improvement in responsiveness andseizures. | No significant improvement. | No significant improvement. | No significant improvement. | No significant Improvement. | |
| **Outcome** | Alived at the age of 5 months. | Alived at the age of 3 years. | Died of respiratory distress at the age of 8 months. | Died at the age of 2.5 days. | Alived at the age of 4 years 6 months. | Died of respiratory failure at the age of 1 year 1 month. | Died at the age of 28 days. |

*IUGR, intrauterine growth retardation; MRI, magnetic resonance imaging; MRS, magnetic resonance spectroscopy; CT, computed tomography; EEG, electroencephalograph; IUGR, intrauterine growth retardation.*

in neurological condition and walking ability (**Table 1**). Compared to the responsive patients, the unresponsive patient carried variants in exon 4 and had a relatively early onset. As was observed in patients with the c.370G > A variant, the disease manifestations associated with changes in this exon may be more severe than those associated with variants in the upstream exons. Moreover, CoQ10 therapy was initiated relatively late (at 19 years of age). These could be the reasons for this patient's unresponsiveness to CoQ10 therapy. Nevertheless, CoQ10 therapy was effective in most patients in this cohort.

## Treatment of Patients With *COQ4* Variants in Exons 5–7

In contrast to cohort 1, four patients in cohort 2 (**Figure 1B**) received CoQ10 therapy, and three (75%) were unresponsive, which is in accordance with the life-threatening presentations of patients in this cohort (**Table 2**). The responsive male patient had both a relatively late onset and early initiation of CoQ10 therapy, suggesting that disease severity and initiation of therapy are the essential determinants of CoQ10 therapy effectiveness.

## Treatment of Patients With the East Asian-specific c.370G > A *COQ4* Variant

Six homozygous patients in this cohort were treated with CoQ10 supplementation, and five (83.3%) (three females and two males) were responsive to the therapy (**Figure 1C**), suggesting that CoQ10 therapy effectively improved the health of these homozygous patients.

For the compound heterozygous patients with the c.370G > A variant, responsiveness to CoQ10 therapy depended on the other variant they carried. For example, the patient with the c.371G > T variant responded to CoQ10 therapy, which is in accordance with the similar clinical presentations of this patient and those of c.370G > A homozygotes. In contrast, none of the four patients carrying the c.402+1G > C variant benefited from CoQ10 therapy, which is also in accordance with the severe manifestations caused by this loss-of-function variant. These results further support the hypothesis that the effectiveness of CoQ10 therapy is associated with disease severity and the timing of therapy initiation.

# DISCUSSION

The highly heterogeneous clinical profiles of primary CoQ10 deficiencies like COQ10D7, including variations in clinical manifestations, age of onset, treatment effectiveness, and outcomes, could be attributed to many speculative factors (Alcazar-Fabra, et al., 2021). For example, genetic factors, such as differences in pathogenic variations within the diverse genetic backgrounds of populations or ethnicities, epigenetic modifications, and defects in biological processes other than MRC activity as well as environmental factors, including the nutritional or metabolic status of specific tissues have been suggested. Moreover, the combined effects of these factors may manifest very early, even in the first stages of embryonic development and may eventually determine the degree of tissue damage and disease severity (Alcazar-Fabra, et al., 2021; Fabozzi, et al., 2021). Due to the limited number of cases and the evident bias in many studies, it is impossible to fully decipher the correlations between the clinical profiles and putative causes, which hampers the diagnosis and management of this set of diseases (Alcazar-Fabra, et al., 2021). However, the fragmented pieces of information derived from existing COQ10D7 cases enable us to take a small step towards creating a comprehensive profile for this monogenic disease.

## Genotype-Phenotype Correlations

Generally, the major clinical presentations caused by pathogenic *COQ4* variants in exons 1–4 (i.e., amino acid changes in the N-terminus of COQ4) are CNS symptoms which are less life-threating, and are associated with late onset, responsiveness to CoQ10 therapy, and a relatively long lifespan. In contrast, the clinical presentations caused by pathogenic *COQ4* variants in exons 5–7 (i.e., amino acid changes in the C-terminus of COQ4) included multi-systemic symptoms which are more fatal and are associated with early onset, unresponsiveness to CoQ10 therapy, and early death. Patients with the East Asian-specific c.370G > A variant displayed intermediate disease severity with multi-systemic dysfunction, which was between that of the patients with variants in exons 1–4 and 5–7. Specifically, the c.370G > A homozygous patients had early disease onset, were responsive to CoQ10 therapy, and a relatively long lifespan. For the c.370G > A heterozygotes, case with other allelic variant located in exon 4 had late onset, was responsive to CoQ10 therapy, and relatively long lifespan, whereas case with allelic variant located in exon 6 had early onset, was unresponsive to CoQ10 therapy, and early death. Moreover, loss-of-function variants, such as frameshift and splicing variants, were associated with early onset, unresponsiveness to CoQ10 therapy, and poor outcomes. Lastly, based on the current data, sex is unlikely to be associated with disease severity.

The mechanism underlying the exon-dependent genotype-phenotype correlation may be associated with the structure and function of COQ4. Yeast Coq4, which is a functional ortholog of human COQ4, is located at the inner mitochondrial membrane on the matrix side and interacts with other Coq proteins, such as Coq3, Coq5, Coq6, Coq7, and Coq9 (Belogrudov, et al., 2001; Marbois, et al., 2005; Awad, et al., 2018). Two functional motifs have been identified in the C-terminus of Coq4; one is a geranylgeranyl monophosphate lipid-bound long hydrophobic α-helix (PDB: 3KB4, Northeastern structural genomics program), and the other is a highly conserved putative zinc ligand motif [HDxxH-(x)11-E] (Marbois et al., 2009). Coq4 is thought to function as a scaffold or organizer that simultaneously anchors the inner mitochondrial membrane, CoQ synthetic complex, and long polyisoprenyl tail of CoQ intermediates and/or CoQ, to organize the ring modifications in the CoQ biosynthetic pathways, and the C-terminus of Coq4 is essential for maintaining these functions (Marbois et al., 2009; Rea et al., 2010; He et al., 2014). In the predicted structure of human COQ4 with Phyre2 (Kelley, et al., 2015), there are

**TABLE 4 |** Biochemical analysis of skin fibroblasts.

| Case number | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| Reference | Sondheimer, et al. (2017) | Bosch, et al. (2018) | Chung, et al. (2015) | Chung, et al. (2015) | Lu, et al. (2019) | Yu, et al. (2019); Ling, et al., (2019) | Chung, et al. (2015) | Chung, et al. (2015) |
| *COQ4* variants | c.23_33del11/ c.331G > T/c.356C > T | Homozygous c.230C > T | Homozygous c.370G > A | Homozygous c.370G > A | Homozygous c.370G > A | c.370G > A/ c.371G > T | c.370G > A/ c.402+1G > C | c.550 T > C/ c.402+1G > A |
| Affected exons | 1, 4 | 3 | 4 | 4 | 4 | 4 | 4, unknown | 6, unknown |
| CoQ10 level | ↓ 19.1 μg/mg protein (45.4–65.7) | ↓ 0.26 nmol/U (0.77–1.61) | ↓ 0.4 pmol/U COX (1.64–3.32) | ↓ 16.4 ng/mg protein (46.1 ± 3) | ↓ | ↓ 0.29 nmol/U COX (1.64–3.32) | ↓ 0.4 pmol/U COX (1.64–3.32) | ↓ 0.63 pmol/U CS (1.04–2.92) |
| CI | ↓ 0.63 (0.86–1.70) (nmol/min/mg protein) | | | | N | N | N | |
| CI + III | ↓ 0.11 (0.55–1.30) (nmol/min/mg protein) | | | ↓ 64% of CS | | | | |
| CII | N | | N | ↓ 90% of CS | N | N | N | N |
| CII + III | ↓ 0.08 (0.20–0.79) (nmol/min/mg protein) | | ↓ 183 mU/U COX (269–781) | ↓ 55% of CS | ↓ | ↓ 135 mU/U COX (269–781) | ↓ 130 mU/U COX (269–781) | ↓ 183 mU/U COX (269–781) |
| CIII | | | N | | | N | N | N |
| CIV | N | | | ↓ 67% of CS | N | N | N | |
| CV | N | | | | | | | |

*CI, complex I; CII, complex II; CIII, complex III; CIV, complex IV; CS, citrate synthase; COX, cyclooxygenase; N, normal; ↓ decreased level.*

transmembrane helices in the C-terminus (**Figure 2**). Therefore, it is reasonable to speculate that the C-terminus of human COQ4 is likely also important for protein function; thus, *COQ4* variants in exons 5–7, which are more likely to disrupt the function of the C-terminus, could cause more serious manifestations than variants in exons 1–4. Similar associations between pathogenic variants in specific domains and certain phenotypes have been also reported in other *COQ* genes. For example, missense variants within the COQ8 protein-specific KxGQ domain were shown to be associated with cortical and pyramidal tract dysfunction (Traschutz, et al., 2020). Nonetheless, more data, including a high-resolution protein structure and in-depth mechanistic investigations, are required to confirm this speculation.

Notably, a dominant mode of inheritance was suggested for two patients: a 3-year-old boy who carried a *de novo* heterozygous 3.9-Mb deletion affecting at least 80 genes, including *COQ4*, who displayed intellectual disability, encephalomyopathy, and dysmorphic features (Salviati, et al., 2012), and a 4-year-old girl with a heterozygous *COQ4* c.483 G > C variant who displayed severe metabolic/mitochondrial deficits, extensive muscle damage, and lethal rhabdomyolysis (Romero-Moya, et al., 2017b). However, some speculative reasons may explain these exceptions. For the first patient, the presence of a second allelic variant, for example, a variant in a regulatory region, would have been missed by only sequencing the exon regions. Moreover, the clinical presentations could be attributed to, either fully or in part, other causative genes other than *COQ4*, since at least 80 genes were affected by the deletion. Similarly, the presence of a second allelic variant may have been missed in the second patient, which could explain why genetic rescue of mitochondrial and skeletal muscle impairment was observed with patient-origin induced pluripotent stem cells (iPSCs) when the c.483 G > C variant was corrected. However, as mentioned above, the phenotypic heterogeneity is high among COQ10D7 patients, even for siblings who share similar genetic backgrounds, suggesting that other genetic or environmental contributors are involved in the pathogenesis of COQ10D7. Moreover, CoQ10 is involved in several biological processes, but it remains unclear whether any of the clinical presentations are attributed to the impairment of other biological processes instead of MRC deficiency in COQ10D7. Therefore, it is possible that, in rare cases, clinical manifestations could be associated with hemizygous or heterozygous *COQ4* variants.

## Biochemical Analyses

In accordance with the role of CoQ10 in the MRC, reduced levels of CoQ10 and complex II + III activity in muscle or fibroblast samples are hallmarks of COQ10D7. *In vivo* assessments of the CoQ biosynthetic rate using cultured fibroblasts were available, which enabled us to discriminate primary deficiency from secondary deficiency (Rodriguez-Aguilera, et al., 2017). Although the detection of CoQ10 biosynthesis and MRC activity have some diagnostic value, their clinical utility for COQ10D7 diagnosis is limited because of the invasive sampling, relatively long turnaround time for cell culture, limited data regarding diagnostic accuracy, and the lack of a consensus or comprehensive evaluation of detection methodologies. Therefore, genetic testing, particularly high-throughput sequencing, either whole exome and genome-based, should be used as the first-line tool to diagnose

**TABLE 5 |** Biochemical analysis of muscle biopsies.

| Case number | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| Reference | Sondheimer, et al. (2017) | Brea-Calvo, et al. (2015) | Brea-Calvo, et al. (2015) | Brea-Calvo, et al. (2015) | Chung, et al. (2015) | Yu, et al. (2019); Ling, et al. (2019) | Brea-Calvo, et al. (2015) | Brea-Calvo, et al. (2015) | Chung, et al. (2015) |
| *COQ4* variants | c.23_33del11/ c.331G > T/ c.356C > T | c.155 T > C/ c.521_523delCCA | c.155 T > C/ c.521_523delCCA | Homozygous c.190C > T | c.245 T > A/ c.473G > A | c.370G > A/c.371G > T | c.421C > T/ c.718C > T | Homozygous c.433C > G | Homozygous c.718C > T |
| Affected exons | 1, 4 | 2, 5 | 2, 5 | 2 | 3, 5 | 4 | 5, 7 | 5 | 7 |
| CoQ10 level | ↓ 6.9 μg/g tissue (19.6–46.8) | | | | ↓ 5.2 μg/g (16% of N) | N | | | |
| CI | | N | ↑ 145% of CS | ↓ <5% of CS | | N | ↓ 6% of CS | ↓ 36% of CS | N |
| CI + III | | N | N | | | | | ↓ 24% of CS | |
| CII | | N | N | N | | N | ↓ 42% of CS | N | N |
| CII + III | | ↓ 55% of CS | N | ↓ 30% of CS | ↓ 14% of N | | ↓ 43% of CS | ↓ 34% of CS | |
| CIII | | N | ↑ 222% of CS | ↓ 50% of CS | | N | ↓ 10% of CS | N | N |
| CIV | | ↓ 50% of CS | ↑ 189% of CS | N | | N | ↓ 30% of CS | N | N |

*CI, complex I; CII, complex II; CIII, complex III; CIV, complex IV; CS, citrate synthase; N, normal; ↓, decreased level; ↑, increased level.*

COQ10D7 in a rapid and definitive manner, especially when point-of-care high-throughput sequencing is available (Sun, et al., 2020; Sanford, et al., 2019). In fact, all reported *COQ4* variants were diagnosed using WES. However, it should be noted that, for new potentially pathogenic variants identified by genetic testing, evaluations of CoQ10 biosynthesis and MRC activity are necessary to confirm their pathogenicity (Brea-Calvo et al., 2021). Phenotypic rescue by CoQ10 treatment would also provide convincing additional evidence for the pathogenicity of a *COQ* gene variant (Romero-Moya et al., 2017a).

In addition to muscle and skin biopsies, measurements of biomarkers in plasma, serum, white blood cells, or urine show promise for clinical applications due to their less invasive sampling. Although an elevated lactate level is not specific for the diagnosis of COQ10D7, it is valuable for evaluating disease severity and prognosis. Another biomarker, serum CoQ10, was detected at relatively low levels in three COQ10D7 patients (Yu, et al., 2019; Caglayan, et al., 2019), suggesting that it could be used for the preliminary diagnosis of COQ10D7 when accompanied with clinical presentations. Moreover, assessment of plasma or serum CoQ10 is simple and rapid, and could provide the first clue and promote timely CoQ10 supplementation. However, more plasma and serum CoQ10 data are needed to establish reference intervals and to validate this concept. Although decreased plasma or serum CoQ10 is not COQ10D7 specific and is associated with all types of primary CoQ deficiency, patients with all types of primary CoQ deficiency could likely benefit from CoQ10 supplementation, and no detrimental side effects have been associated with this biomolecule (Hargreaves 2014). Nevertheless, if skin fibroblasts could be obtained less invasively, they would be an ideal option due to the analytical accuracy of their biochemical analysis and the ability of functional studies (Romero-Moya, et al., 2017b; Alcazar-Fabra, et al., 2021).

## Treatment of COQ10D7

Clear evidence supporting the use of any specific drug for the treatment of mitochondrial diseases (MD) is lacking. However, based on numerous cases with positive responses and the lack of noticeable toxicity, exogenous CoQ10 is widely used for the treatment of MD, particularly primary CoQ10 deficiency (Hargreaves 2014; Orsucci, et al., 2019). Recently, ubiquinol was approved as an orphan drug for the treatment of primary CoQ10 deficiency (Hernandez-Camacho, et al., 2018). Although the delivery and utilization of exogenous CoQ10 is inefficient because of its high molecular weight and low aqueous solubility, when CoQ10 is administered at high doses or in specific formulations, CoQ10 levels are increased in all tissues, including the heart and brain (Zaki 2016). However, approximately 43% (9/21) of patients with COQ10D7 did not benefit from CoQ10 therapy. Due to inconsistent treatment regimens, including variations in formulation and dose, and the lack of data regarding *in vivo* CoQ10 levels, it is currently impossible to determine the conclusive reasons for unresponsiveness in these patients. However, according to the limited data available, the responsiveness to CoQ10 therapy was highly associated with the location of the pathogenic *COQ4* variants and the severity of clinical presentation. A rational speculation was that CoQ10 is more likely to prevent further harm rather than reverse the present impairments. Therefore, unresponsiveness to CoQ10 therapy was more likely to be observed in patients with early onset and life-threatening manifestations. In general, CoQ10 therapy, when administered in time, would be effective, to various extents, for those patients with *COQ4* variants in exons 1–4

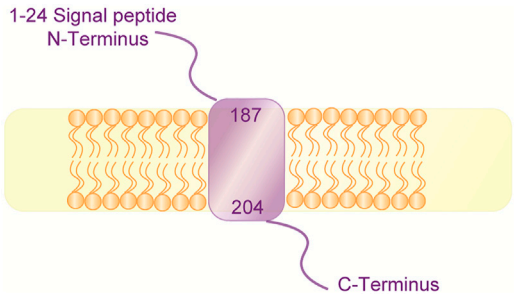



**FIGURE 2 |** Predictive structure of COQ4.

in both alleles, including the c.370G > A homozygotes, thus highlighting the importance of early genetic diagnosis and intervention.

For responsive patients, the symptoms of multisystem dysfunction, including the CNS, were improved, although it is difficult for CoQ10 to penetrate the blood-brain barrier (Musumeci, et al., 2001; Lamperti, et al., 2003; Artuch, et al., 2006; Mancuso, et al., 2012; Mortensen, et al., 2014). It has been shown that ubiquinol has greater therapeutic effectiveness than ubiquinone due to the improved mitochondrial bioenergetics in the brain and reduced oxidative stress and astrogliosis (Garcia-Corzo, et al., 2014). However, the differences in the therapeutic efficacy of ubiquinol and ubiquinone for COQ10D7 treatment are unclear due to limited data. Since CoQ10 is a lipophilic molecule, its gastrointestinal uptake and assimilation may be inefficient (Masotta, et al., 2019). Therefore, CoQ10 derivatives of a more absorbable nature could be promising alternatives to enhance CoQ10 biosynthesis or bypass deficient steps in the CoQ10 biosynthetic pathway (Awad, et al., 2018). Recently, $\beta$-resorcylic acid, a structural analog of the CoQ precursor 4-hydroxybenzoic acid (4-HB), was shown to be a powerful therapy for patients with COQ9 or COQ7 variants, thus shedding light on the application of 4-HB analogs for other forms of primary CoQ10 deficiency, including COQ10D7 (Hidalgo-Gutierrez, et al., 2019). In addition to CoQ therapy, symptomatic treatments were also helpful to some extent for improving the health condition of patients with MD. In general, there is no substantial difference in symptomatic treatments for the same conditions between patients with MD and non-MD patients (Orsucci, et al., 2019). However, in terms of COQ10D7 treatment, the effectiveness of drugs, such as creatine, thiamine, riboflavin, and L-carnitine, requires further evaluation, although some of these drugs may be associated with positive outcomes in rare cases. Moreover, other drugs, such as valproate, linezolid, and aminoglycosides, should be used carefully because of their potential for mitochondrial toxicity (Mancuso, et al., 2012; Orsucci, et al., 2019). Notably, two c.370G > A homozygous patients died due to infections and associated complications (Ling, et al., 2019; Yu, et al., 2019), suggesting that infection management is crucial for COQ10D7 patients since they could lead to metabolic crises in patients with MD (Orsucci, et al., 2019).

Based on findings regarding genotype-phenotype correlations, in-depth functional investigations of COQ4 domains would be useful for the development of targeted drugs for COQ10D7 patients with specific variants. Similarly, targeted drugs could be developed based on an understanding of the presentations associated with impairment of other biological processes besides MRC. Revolutionary advances in multifaceted omics technologies, including genomics, metabolomics, proteomics, epigenomics, transcriptomics, and interactomics, will continue to expand our knowledge of mitochondrial pathology (Rahman and Rahman 2018) and provide us with the possibility of developing targeted drugs based on newly identified molecular phenotypes.

In addition to conventional therapies, targeted gene therapy is a novel and attractive prospect for COQ10D7. With the development of induced pluripotent stem cells (iPSCs) and gene editing technologies, *in vitro* and *in vivo* disease models can be established more readily to advance the development of gene therapy (Karpe, et al., 2021). For example, phenotypic rescue was achieved in patient-derived iPSCs and mouse models via rectification of pathogenic variants in amyotrophic lateral sclerosis and disruption of splicing regulatory elements in spinal muscular atrophy (SMA) using gene editing (Wang, et al., 2017; Son, et al., 2019; Lin, et al., 2020). A similar investigation was initiated for COQ10D7 in which genetic rescue of mitochondrial and skeletal muscle impairment was attained in iPSCs with a *COQ4* c.483G > C variant (Romero-Moya et al., 2017a). Currently, three targeted drugs (nusinersen, onasemnogene abeparvovec-xioi, and risdiplam) based on gene replacement or splicing modification have been approved by the US Food and Drug Administration for use in the clinical treatment of SMA (Menduti, et al., 2020), confirming the promising clinical utility of targeted gene therapy. However, more comprehensive *in vivo* evaluations of the efficacy and safety of gene therapy for COQ10D7 are required because the genetic abnormalities in COQ10D7 are more heterozygous than those in SMA.

Whether conventional or targeted gene therapies are administered, a unified, sensitive, and valid design for measuring outcomes is a prerequisite for accurate evaluation of therapeutic efficacy (Mancuso, et al., 2017; Steele, et al., 2017; Koene, et al., 2018). Such comprehensive evaluations were lacking in previous COQ10D7 studies, partially due to the rapid deterioration of patients who have the disease. However, this should be emphasized in future studies, if possible. Randomized, double-blind, placebo-controlled, crossover trials are also essential for accurate evaluation of therapeutic efficacy in a subset of COQ10D7 patients, for example, patients who have *COQ4* variants in exons 1–4 in both alleles and present with mild symptoms.

## CONCLUSION AND PERSPECTIVES

We showed that COQ10D7 displays distinct clinical presentations, and responsiveness to CoQ10 therapy depends on the location of *COQ4* variants, providing a

fundamental reference for the sub-classification of and prognosis prediction for this disease. Moreover, sex is unlikely to be associated with disease severity. Point-of-care high-throughput sequencing would be useful for the rapid diagnosis of pathogenic *COQ4* variants, and biochemical analyses to detect the characteristic impairments in CoQ10 biosynthesis and MRC activity and, phenotypic rescue by CoQ10 treatment, are necessary to confirm the pathogenicity of suspicious variants. In addition to CoQ10 derivatives, targeted drugs and gene therapy could be useful for COQ10D7 treatment relying on in-depth functional investigations and the development of gene editing technologies.

## AUTHOR CONTRIBUTIONS

JX: Investigation, Data curation, Visualization, Writing—Original draft preparation. JJ: Investigation, Data curation, Writing—Original draft preparation. QG: Conceptualization, Data curation, Writing—Original draft—Review and Editing.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.776807/full#supplementary-material

## REFERENCES


Alcázar-Fabra, M., Navas, P., and Brea-Calvo, G. (2016). Coenzyme Q Biosynthesis and its Role in the Respiratory Chain Structure. *Biochim. Biophys. Acta (Bba) - Bioenerg.* 1857, 1073–1078. doi:10.1016/j.bbabio.2016.03.010

Alcázar-Fabra, M., Rodríguez-Sánchez, F., Trevisson, E., and Brea-Calvo, G. (2021). Primary Coenzyme Q Deficiencies: A Literature Review and Online Platform of Clinical Features to Uncover Genotype-Phenotype Correlations. *Free Radic. Biol. Med.* 167, 141–180. doi:10.1016/j.freeradbiomed.2021.02.046

Artuch, R., Brea-Calvo, G., Briones, P., Aracil, A., Galván, M., Espinós, C., et al. (2006). Cerebellar Ataxia with Coenzyme Q10 Deficiency: Diagnosis and Follow-Up after Coenzyme Q10 Supplementation. *J. Neurol. Sci.* 246, 153–158. doi:10.1016/j.jns.2006.01.021

Awad, A. M., Bradley, M. C., Fernández-del-Río, L., Nag, A., Tsui, H. S., and Clarke, C. F. (2018). Coenzyme Q10 Deficiencies: Pathways in Yeast and Humans. *Essays Biochem.* 62, 361–376. doi:10.1042/ebc20170106

Belogrudov, G. I., Lee, P. T., Jonassen, T., Hsu, A. Y., Gin, P., and Clarke, C. F. (2001). Yeast *COQ4* Encodes a Mitochondrial Protein Required for Coenzyme Q Synthesis. *Arch. Biochem. Biophys.* 392, 48–58. doi:10.1006/abbi.2001.2448

Bersuker, K., Hendricks, J. M., Li, Z., Magtanong, L., Ford, B., Tang, P. H., et al. (2019). The Coq Oxidoreductase Fsp1 Acts Parallel to Gpx4 to Inhibit Ferroptosis. *Nature* 575, 688–692. doi:10.1038/s41586-019-1705-2

Bosch, A. M., Kamsteeg, E.-J., Rodenburg, R. J., van Deutekom, A. W., Buis, D. R., Engelen, M., et al. (2018). Coenzyme Q10 Deficiency Due to a *COQ4* Gene Defect Causes Childhood-Onset Spinocerebellar Ataxia and Stroke-like Episodes. *Mol. Genet. Metab. Rep.* 17, 19–21. doi:10.1016/j.ymgmr.2018.09.002

Brea-Calvo, G., Alcázar-Fabra, M., Trevisson, E., and Navas, P. (2021). "Coenzyme Q Biosynthesis Disorders," in *Mitochondrial Diseases*. Sevilla, Spain: Springer Nature. doi:10.1007/978-3-030-70147-5_6

Brea-Calvo, G., Haack, T. B., Karall, D., Ohtake, A., Invernizzi, F., Carrozzo, R., et al. (2015). *COQ4* Mutations Cause a Broad Spectrum of Mitochondrial Disorders Associated with Coq10 Deficiency. *Am. J. Hum. Genet.* 96, 309–317. doi:10.1016/j.ajhg.2014.12.023

Caglayan, A. O., Gumus, H., Sandford, E., Kubisiak, T. L., Ma, Q., Ozel, A. B., et al. (2019). *COQ4* Mutation Leads to Childhood-Onset Ataxia Improved by Coq10 Administration. *Cerebellum* 18, 665–669. doi:10.1007/s12311-019-01011-x

Casarin, A., Jimenez-Ortega, J. C., Trevisson, E., Pertegato, V., Doimo, M., Ferrero-Gomez, M. L., et al. (2008). Functional Characterization of Human *COQ4*, a Gene Required for Coenzyme Q10 Biosynthesis. *Biochem. Biophysical Res. Commun.* 372, 35–39. doi:10.1016/j.bbrc.2008.04.172

Chen, X. R., Xu, J. P., and Yao, Y. H. (2020). Primary Coenzyme Q10 Deficiency-7: A Case Report and Literature Review. *Zhonghua Er Ke Za Zhi* 58, 928–932. doi:10.3760/cma.j.cn112140-20200610-00601

Chung, W. K., Martin, K., Jalas, C., Braddock, S. R., Juusola, J., Monaghan, K. G., et al. (2015). Mutations inCOQ4, an Essential Component of Coenzyme Q Biosynthesis, Cause Lethal Neonatal Mitochondrial Encephalomyopathy. *J. Med. Genet.* 52, 627–635. doi:10.1136/jmedgenet-2015-103140

Cramer, W. A., Hasan, S. S., and Yamashita, E. (2011). The Q Cycle of Cytochrome Bc Complexes: A Structure Perspective. *Biochim. Biophys. Acta (Bba) - Bioenerg.* 1807, 788–802. doi:10.1016/j.bbabio.2011.02.006

Desbats, M. A., Lunardi, G., Doimo, M., Trevisson, E., and Salviati, L. (2015). Genetic Bases and Clinical Manifestations of Coenzyme Q10 (CoQ10) Deficiency. *J. Inherit. Metab. Dis.* 38, 145–156. doi:10.1007/s10545-014-9749-9

Doll, S., Freitas, F. P., Shah, R., Aldrovandi, M., da Silva, M. C., Ingold, I., et al. (2019). Fsp1 Is a Glutathione-independent Ferroptosis Suppressor. *Nature* 575, 693–698. doi:10.1038/s41586-019-1707-0

Evans, D. R., and Guy, H. I. (2004). Mammalian Pyrimidine Biosynthesis: Fresh Insights into an Ancient Pathway. *J. Biol. Chem.* 279, 33035–33038. doi:10.1074/jbc.r400007200

Fabozzi, G., Iussig, B., Cimadomo, D., Vaiarelli, A., Maggiulli, R., Ubaldi, N., et al. (2021). The Impact of Unbalanced Maternal Nutritional Intakes on Oocyte Mitochondrial Activity: Implications for Reproductive Function. *Antioxidants* 10, 91. doi:10.3390/antiox10010091

Finsterer, J., and Zarrouk-Mahjoub, S. (2017). Mitochondrial Cardioencephalopathy Due to a *COQ4* Mutation. *Mol. Genet. Metab. Rep.* 13, 7–8. doi:10.1016/j.ymgmr.2017.07.003

García-Corzo, L., Luna-Sánchez, M., Doerrier, C., Ortiz, F., Escames, G., Acuña-Castroviejo, D., et al. (2014). Ubiquinol-10 Ameliorates Mitochondrial Encephalopathy Associated with Coq Deficiency. *Biochim. Biophys. Acta (Bba) - Mol. Basis Dis.* 1842, 893–901. doi:10.1016/j.bbadis.2014.02.008

Ge, L., Chen, Z., Liu, L., Zheng, X., Kong, J., and Yinsen, S. (2019). Clinical Characteristics and Genetic Analysis of Primary Coenzyme Q10 Deficiency Caused by *COQ4* Gene Mutation. *Chin. J. Endocrinol. Metab.* 35, 1014–1018. doi:10.3760/cma.j.issn.1000-6699.2019.12.004

Guarás, A., Perales-Clemente, E., Calvo, E., Acín-Pérez, R., Loureiro-Lopez, M., Pujol, C., et al. (2016). The Coqh2/coq Ratio Serves as a Sensor of Respiratory Chain Efficiency. *Cel Rep.* 15, 197–209. doi:10.1016/j.celrep.2016.03.009

Hargreaves, I. P. (2014). Coenzyme Q10 as a Therapy for Mitochondrial Disease. *Int. J. Biochem. Cel Biol.* 49, 105–111. doi:10.1016/j.biocel.2014.01.020

Hashemi, S. S., Zare-Abdollahi, D., Bakhshandeh, M. K., Vafaee, A., Abolhasani, S., Inanloo Rahatloo, K., et al. (2021). Clinical Spectrum in Multiple Families with Primary COQ 10 Deficiency. *Am. J. Med. Genet.* 185, 440–452. doi:10.1002/ajmg.a.61983

He, C. H., Xie, L. X., Allan, C. M., Tran, U. C., and Clarke, C. F. (2014). Coenzyme Q Supplementation or Over-expression of the Yeast Coq8 Putative Kinase Stabilizes Multi-Subunit Coq Polypeptide Complexes in Yeast Coq Null Mutants. *Biochim. Biophys. Acta (Bba) - Mol. Cel Biol. Lipids* 1841, 630–644. doi:10.1016/j.bbalip.2013.12.017

Hernández-Camacho, J. D., Bernier, M., López-Lluch, G., and Navas, P. (2018). Coenzyme Q10 Supplementation in Aging and Disease. *Front. Physiol.* 9, 44. doi:10.3389/fphys.2018.00044

Hidalgo-Gutiérrez, A., Barriocanal-Casado, E., Bakkali, M., Díaz-Casado, M. E., Sánchez-Maldonado, L., Romero, M., et al. (2019). β-RA Reduces DMQ/CoQ Ratio and Rescues the Encephalopathic Phenotype in Coq9 R239X Mice. *EMBO Mol. Med.* 11, e9466. doi:10.15252/emmm.201809466

Hughes, B. G., Harrison, P. M., and Hekimi, S. (2017). Estimating the Occurrence of Primary Ubiquinone Deficiency by Analysis of Large-Scale Sequencing Data. *Sci. Rep.* 7, 17744. doi:10.1038/s41598-017-17564-y

Karpe, Y., Chen, Z., and Li, X.-J. (2021). Stem Cell Models and Gene Targeting for Human Motor Neuron Diseases. *Pharmaceuticals* 14, 565. doi:10.3390/ph14060565

Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., and Sternberg, M. J. E. (2015). The Phyre2 Web portal for Protein Modeling, Prediction and Analysis. *Nat. Protoc.* 10, 845–858. doi:10.1038/nprot.2015.053

Koene, S., van Bon, L., Bertini, E., Jimenez-Moreno, C., van der Giessen, L., de Groot, I., et al. (2018). Outcome Measures for Children with Mitochondrial Disease: Consensus Recommendations for Future Studies from a delphi-based International Workshop. *J. Inherit. Metab. Dis.* 41, 1267–1273. doi:10.1007/s10545-018-0229-5

Lamperti, C., Naini, A., Hirano, M., De Vivo, D. C., Bertini, E., Servidei, S., et al. (2003). Cerebellar Ataxia and Coenzyme Q10 Deficiency. *Neurology* 60, 1206–1208. doi:10.1212/01.wnl.0000055089.39373.fc

Letts, J. A., and Sazanov, L. A. (2017). Clarifying the Supercomplex: The Higher-Order Organization of the Mitochondrial Electron Transport Chain. *Nat. Struct. Mol. Biol.* 24, 800–808. doi:10.1038/nsmb.3460

Lin, X., Chen, H., Lu, Y.-Q., Hong, S., Hu, X., Gao, Y., et al. (2020). Base Editing-Mediated Splicing Correction Therapy for Spinal Muscular Atrophy. *Cell Res* 30, 548–550. doi:10.1038/s41422-020-0304-y

Ling, T.-k., Law, C.-y., Yan, K.-w., Fong, N.-c., Wong, K.-c., Lee, K.-l., et al. (2019). Clinical Whole-Exome Sequencing Reveals a Common Pathogenic Variant in Patients with Coq10 Deficiency: An Underdiagnosed Cause of Mitochondriopathy. *Clinica Chim. Acta* 497, 88–94. doi:10.1016/j.cca.2019.07.016

Lu, M., Zhou, Y., Wang, Z., Xia, Z., Ren, J., and Guo, Q. (2019). Clinical Phenotype, In Silico and Biomedical Analyses, and Intervention for an East Asian Population-specific c.370G>A (p.G124S) COQ4 Mutation in a Chinese Family with CoQ10 Deficiency-Associated Leigh Syndrome. *J. Hum. Genet.* 64, 297–304. doi:10.1038/s10038-019-0563-y

Mancuso, M., McFarland, R., Klopstock, T., Hirano, M., Artuch, R., Bertini, E., et al. (2017). International Workshop:. *Neuromuscul. Disord.* 27, 1126–1137. doi:10.1016/j.nmd.2017.08.006

Mancuso, M., Orsucci, D., Filosto, M., Simoncini, C., and Siciliano, G. (2012). Drugs and Mitochondrial Diseases: 40 Queries and Answers. *Expert Opin. Pharmacother.* 13, 527–543. doi:10.1517/14656566.2012.657177

Marbois, B., Gin, P., Faull, K. F., Poon, W. W., Lee, P. T., Strahan, J., et al. (2005). Coq3 and *COQ4* Define a Polypeptide Complex in Yeast Mitochondria for the Biosynthesis of Coenzyme Q. *J. Biol. Chem.* 280, 20231–20238. doi:10.1074/jbc.m501315200

Marbois, B., Gin, P., Gulmezian, M., and Clarke, C. F. (2009). The Yeast *COQ4* Polypeptide Organizes a Mitochondrial Protein Complex Essential for Coenzyme Q Biosynthesis. *Biochim. Biophys. Acta (Bba) - Mol. Cel Biol. Lipids* 1791, 69–75. doi:10.1016/j.bbalip.2008.10.006

Masotta, N. E., Martinefski, M. R., Lucangioli, S., Rojas, A. M., and Tripodi, V. P. (2019). High-dose Coenzyme Q10-Loaded Oleogels for Oral Therapeutic Supplementation. *Int. J. Pharmaceutics* 556, 9–20. doi:10.1016/j.ijpharm.2018.12.003

Menduti, G., Rasà, D. M., Stanga, S., and Boido, M. (2020). Drug Screening and Drug Repositioning as Promising Therapeutic Approaches for Spinal Muscular Atrophy Treatment. *Front. Pharmacol.* 11, 592234. doi:10.3389/fphar.2020.592234

Mero, S., Salviati, L., Leuzzi, V., Rubegni, A., Calderan, C., Nardecchia, F., et al. (2021). New Pathogenic Variants in *COQ4* Cause Ataxia and Neurodevelopmental Disorder without Detectable Coq10 Deficiency in Muscle or Skin Fibroblasts. *J. Neurol.* 268, 3381–3389. doi:10.1007/s00415-021-10509-6

Montero, R., Sánchez-Alcázar, J. A., Briones, P., Hernández, Á. R., Cordero, M. D., Trevisson, E., et al. (2008). Analysis of Coenzyme Q10 in Muscle and Fibroblasts for the Diagnosis of Coq10 Deficiency Syndromes. *Clin. Biochem.* 41, 697–700. doi:10.1016/j.clinbiochem.2008.03.007

Montini, G., Malaventura, C., and Salviati, L. (2008). Early Coenzyme Q10 Supplementation in Primary Coenzyme Q10 Deficiency. *N. Engl. J. Med.* 358, 2849–2850. doi:10.1056/nejmc0800582

Mortensen, S. A., Rosenfeldt, F., Kumar, A., Dolliner, P., Filipiak, K. J., Pella, D., et al. (2014). The Effect of Coenzyme Q 10 on Morbidity and Mortality in Chronic Heart Failure. *JACC: Heart Fail.* 2, 641–649. doi:10.1016/j.jchf.2014.06.008

Musumeci, O., Naini, A., Slonim, A. E., Skavin, N., Hadjigeorgiou, G. L., Krawiecki, N., et al. (2001). Familial Cerebellar Ataxia with Muscle Coenzyme Q10 Deficiency. *Neurology* 56, 849–855. doi:10.1212/wnl.56.7.849

Orsucci, D., Ienco, E. C., Siciliano, G., and Mancuso, M. (2019). Mitochondrial Disorders and Drugs: What Every Physician Should Know. *Drugs Context* 8, 212588. doi:10.7573/dic.212588

Quinzii, C. M., Emmanuele, V., and Hirano, M. (2014). Clinical Presentations of Coenzyme Q10 Deficiency Syndrome. *Mol. Syndromol* 5, 141–146. doi:10.1159/000360490

Rahman, J., and Rahman, S. (2018). Mitochondrial Medicine in the Omics Era. *The Lancet* 391, 2560–2574. doi:10.1016/s0140-6736(18)30727-x

Rea, S. L., Graham, B. H., Nakamaru-Ogiso, E., Kar, A., and Falk, M. J. (2010). Bacteria, Yeast, Worms, and Flies: Exploiting Simple Model Organisms to Investigate Human Mitochondrial Diseases. *Dev. Disabil. Res. Revs* 16, 200–218. doi:10.1002/ddrr.114

Rodríguez-Aguilera, J., Cortés, A., Fernández-Ayala, D., and Navas, P. (2017). Biochemical Assessment of Coenzyme Q10 Deficiency. *Jcm* 6, 27. doi:10.3390/jcm6030027

Romero-Moya, D., Castaño, J., Santos-Ocaña, C., Navas, P., and Menendez, P. (2017a). Generation, Genome Edition and Characterization of iPSC Lines from a Patient with Coenzyme Q 10 Deficiency Harboring a Heterozygous Mutation in COQ4 Gene. *Stem Cel Res.* 24, 144–147. doi:10.1016/j.scr.2016.09.007

Romero-Moya, D., Santos-Ocaña, C., Castaño, J., Garrabou, G., Rodríguez-Gómez, J. A., Ruiz-Bonilla, V., et al. (2017b). Genetic rescue of Mitochondrial and Skeletal Muscle Impairment in an Induced Pluripotent Stem Cells Model of Coenzyme Q10 Deficiency. *Stem Cells* 35, 1687–1703. doi:10.1002/stem.2634

Salviati, L., Trevisson, E., Rodriguez Hernandez, M. A., Casarin, A., Pertegato, V., Doimo, M., et al. (2012). Haploinsufficiency ofCOQ4causes Coenzyme Q10deficiency. *J. Med. Genet.* 49, 187–191. doi:10.1136/jmedgenet-2011-100394

Sanford, E. F., Clark, M. M., Farnaes, L., Williams, M. R., Perry, J. C., Ingulli, E. G., et al. (2019). Rapid Whole Genome Sequencing Has Clinical Utility in Children in the PICU*. *Pediatr. Crit. Care Med.* 20, 1007–1020. doi:10.1097/pcc.0000000000002056

Shalata, A., Edery, M., Habib, C., Genizi, J., Mahroum, M., Khalaily, L., et al. (2019). Primary Coenzyme Q Deficiency Due to Novel Adck3 Variants, Studies in Fibroblasts and Review of Literature. *Neurochem. Res.* 44, 2372–2384. doi:10.1007/s11064-019-02786-5

Son, Y. S., Choi, K., Lee, H., Kwon, O., Jung, K. B., Cho, S., et al. (2019). ASMN2Splicing Modifier Rescues the Disease Phenotypes in an *In Vitro* Human Spinal Muscular Atrophy Model. *Stem Cell Dev.* 28, 438–453. doi:10.1089/scd.2018.0181

Sondheimer, N., Hewson, S., Cameron, J. M., Somers, G. R., Broadbent, J. D., Ziosi, M., et al. (2017). Novel Recessive Mutations in COQ4 Cause Severe Infantile Cardiomyopathy and Encephalopathy Associated with CoQ 10 Deficiency. *Mol. Genet. Metab. Rep.* 12, 23–27. doi:10.1016/j.ymgmr.2017.05.001

Spinazzi, M., Casarin, A., Pertegato, V., Salviati, L., and Angelini, C. (2012). Assessment of Mitochondrial Respiratory Chain Enzymatic Activities on Tissues and Cultured Cells. *Nat. Protoc.* 7, 1235–1246. doi:10.1038/nprot.2012.058

Steele, H. E., Horvath, R., Lyon, J. J., and Chinnery, P. F. (2017). Monitoring Clinical Progression with Mitochondrial Disease Biomarkers. *Brain* 140, 2530–2540. doi:10.1093/brain/awx168

Stefely, J. A., and Pagliarini, D. J. (2017). Biochemistry of Mitochondrial Coenzyme Q Biosynthesis. *Trends Biochem. Sci.* 42, 824–843. doi:10.1016/j.tibs.2017.06.008

Sun, X., Song, L., Yang, W., Zhang, L., Liu, M., Li, X., et al. (2020). Nanopore Sequencing and its Clinical Applications. *Methods Mol. Biol.* 2204, 13–32. doi:10.1007/978-1-0716-0904-0_2

Traschütz, A., Schirinzi, T., Laugwitz, L., Murray, N. H., Bingman, C. A., Reich, S., et al. (2020). Clinico-Genetic, Imaging and Molecular Delineation of COQ8A -Ataxia: A Multicenter Study of 59 Patients. *Ann. Neurol.* 88, 251–263. doi:10.1002/ana.25751

Tsang, M. H. Y., Kwong, A. K. Y., Chan, K. L. S., Fung, J. L. F., Yu, M. H. C., Mak, C. C. Y., et al. (2020). Delineation of Molecular Findings by Whole-Exome

Sequencing for Suspected Cases of Paediatric-Onset Mitochondrial Diseases in the Southern Chinese Population. *Hum. Genomics* 14, 28. doi:10.1186/s40246-020-00278-0

Turunen, M., Olsson, J., and Dallner, G. (2004). Metabolism and Function of Coenzyme Q. *Biochim. Biophys. Acta (Bba) - Biomembranes* 1660, 171–199. doi:10.1016/j.bbamem.2003.11.012

Wang, L., Yi, F., Fu, L., Yang, J., Wang, S., Wang, Z., et al. (2017). Crispr/cas9-mediated Targeted Gene Correction in Amyotrophic Lateral Sclerosis Patient Ipscs. *Protein Cell* 8, 365–378. doi:10.1007/s13238-017-0397-3

Wang, Y., and Hekimi, S. (2016). Understanding Ubiquinone. *Trends Cel Biol.* 26, 367–378. doi:10.1016/j.tcb.2015.12.007

Watmough, N. J., and Frerman, F. E. (2010). The Electron Transfer Flavoprotein: Ubiquinone Oxidoreductases. *Biochim. Biophys. Acta (Bba) - Bioenerg.* 1797, 1910–1916. doi:10.1016/j.bbabio.2010.10.007

Yu, M. H.-C., Tsang, M. H.-Y., Lai, S., Ho, M. S.-P., Tse, D. M. L., Willis, B., et al. (2019). Primary Coenzyme Q10 Deficiency-7: Expanded Phenotypic Spectrum and a Founder Mutation in Southern Chinese. *Npj Genom. Med.* 4, 18. doi:10.1038/s41525-019-0091-x

Zaki, N. M. (2016). Strategies for Oral Delivery and Mitochondrial Targeting of Coq10. *Drug Deliv.* 23, 1868–1881. doi:10.3109/10717544.2014.993747

**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.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

*Copyright © 2022 Xie, Jiang and Guo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*

frontiers in Pediatrics

CASE REPORT
published: 31 January 2022
doi: 10.3389/fped.2021.790075

# Andersen–Tawil Syndrome With Novel Mutation in *KCNJ2*: Case Report

*Jisook Yim*[1,2], *Kyoung Bo Kim*[3], *Minsun Kim*[1], *Gun Dong Lee*[1] *and Myungshin Kim*[1,2]*

[1] Department of Laboratory Medicine, College of Medicine, The Catholic University of Korea, Seoul, South Korea, [2] Catholic Genetic Laboratory Center, Seoul St. Mary's Hospital, College of Medicine, The Catholic University of Korea, Seoul, South Korea, [3] Department of Laboratory Medicine, Keimyung University School of Medicine, Daegu, South Korea

Andersen–Tawil syndrome (ATS) is a rare autosomal dominant disorder characterized by a classic symptom triad: periodic paralysis, ventricular arrhythmias associated with prolonged QT interval, and dysmorphic skeletal and facial features. Pathogenic variants of the inwardly rectifying potassium channel subfamily J member 2 (KCNJ2) gene have been linked to the ATS. Herein, we report a novel *KCNJ2* causative variant in a proband and her father showing different ATS-associated symptoms. A 15-year-old girl was referred because of episodic weakness and periodic paralysis in both legs for 2–3 months. The symptoms occurred either when she was tired or after strenuous exercise. These attacks made walking or climbing stairs difficult and lasted from one to several days. She had a short stature (142 cm, <3rd percentile) and weighed 40 kg. The proband also showed orbital hypertelorism, dental crowding, mandibular hypoplasia, fifth-digit clinodactyly, and small hands. Scoliosis in the thoracolumbar region was detected by chest X-ray. Since she was 7 years old, she had been treated for arrhythmia-associated long QT interval and underwent periodic echocardiography. Brain MRI revealed cerebrovascular abnormalities indicating absence or hypoplasia of bilateral internal carotid arteries, and compensation of other collateral vessels was observed. There were no specific findings related to intellectual development. The proband's father also had a history of periodic paralysis similar to the proband. He did not show any cardiac symptoms. Interestingly, he was diagnosed with hyperthyroidism during an evaluation for paralytic symptoms. Clinical exome sequencing revealed a novel heterozygous missense variant: Chr17(GRCh37):g.68171593A>T, NM_000891.2:c.413A>T, p.(Glu138Val) in *KCNJ2* in the proband and the proband's father.

**Keywords: *KCNJ2* gene, KIR2.1, periodic paralysis, long QT syndrome, potassium channel, genetic disorder, dysmorphic feature, Andersen–Tawil syndrome**

OPEN ACCESS

***Edited by:***
*Natália Duarte Linhares, Genuity Science, Ireland*

***Reviewed by:***
*Aleksandra Jezela-Stanek, National Institute of Tuberculosis and Lung Diseases, Poland*
*Ronald Wilders, University of Amsterdam, Netherlands*

***Correspondence:**
*Myungshin Kim*
*microkim@catholic.ac.kr*

***Specialty section:***
*This article was submitted to Genetics of Common and Rare Diseases, a section of the journal Frontiers in Pediatrics*

***Received:*** *06 October 2021*
***Accepted:*** *22 November 2021*
***Published:*** *31 January 2022*

***Citation:***
*Yim J, Kim KB, Kim M, Lee GD and Kim M (2022) Andersen–Tawil Syndrome With Novel Mutation in KCNJ2: Case Report. Front. Pediatr. 9:790075. doi: 10.3389/fped.2021.790075*

## INTRODUCTION

Andersen–Tawil syndrome (ATS; MIM 170390) is a rare autosomal dominant disorder characterized by a classic triad of recurrent flaccid muscle weakness (periodic paralysis), cardiac arrhythmias, and distinctive skeletal and facial features (1). Since the first ATS case presenting with muscle weakness, extrasystoles, and multiple developmental anomalies was reported (2), ATS prevalence is estimated at 0.8–1 in 1,000,000 (3, 4). However, the actual prevalence is expected to be

higher than previously reported because the reported prevalence is from a specific population and subgroup of ATS patients (5) and ATS is a rare disease with a broad phenotypic heterogeneity that is difficult to diagnose. Affected individuals present with episodes of periodic paralysis or cardiac symptoms in their first or second decade (6); however, highly variable phenotypes and incomplete penetrance make diagnosis challenging. Cardiac arrhythmias include ventricular arrhythmias associated with prolonged QT interval and prominent U waves. Skeletal and facial features include low-set ears, broad forehead, ocular hypertelorism, small mandible, dental abnormalities, fifth digit clinodactyly, syndactyly, short stature, and scoliosis (7). The *KCNJ2* and *KCNJ5* are causative genes of ATS. *KCNJ2* on chromosome 17q24.3 encodes an inward rectifier potassium channel protein (Kir2.1), which plays an important role in setting and stabilizing the resting membrane potential primarily in the skeletal muscles, heart, and brain (8). *KCNJ5* (11q24.3) encodes a G-protein-activated inward rectifier potassium channel 4 protein (Kir3.4) that induces an inhibitory effect on the inward rectifier potassium current (9). Therefore, the diagnosis of ATS is established in individuals with two or more symptoms of the classic triad or individuals with one of the classic triad and at least one family member with ATS and/or identification of pathogenic variants in *KCNJ2* and *KCNJ5*. Here, we report a family with ATS with a novel causative missense variant in *KCNJ2* c.413A>T, p.(Glu138Val) and compare their phenotypes along with a literature review.

## CASE REPORT

The proband is a 15-year-old girl who presented with recurrent episodes of leg weakness and paralysis lasting several days beginning 2–3 months prior. These episodes recently became more severe and frequent, occurring at least twice a week. The severity of the weakness during the episodes varied from mild (weakness and difficulty climbing stairs) to severe (inability to walk). Triggers were intensive exercise or feeling tired. Clinical evaluation included medical and family history and previous medical record review where applicable. She had been taking flecainide for an arrhythmia associated with a long QT interval since she was 7 years old and had been receiving regular echocardiogram (ECG) follow-up. Her serum potassium levels tested during hospitalization were normal. Thyroid function test results were also within the reference interval. Electromyography (EMG) showed no specific findings. Brain magnetic resonance imaging (MRI) revealed suspicious congenital absence or hypoplasia of both internal carotid arteries (ICA) with supply to the anterior circulation via carotid-vertebrobasilar anastomoses and hypertrophy of the posterior communicating artery (PCOM). No intellectual disability was observed during development. Chest X-rays showed mild scoliosis in the thoracolumbar region. During physical examination, she was 142 cm tall (<3rd percentile) and weighed 40 kg (3rd percentile). Additional physical examination during genetic counseling revealed that the proband had dysmorphic features, including orbital hypertelorism, mandibular hypoplasia, dental crowding, clinodactyly of the fifth finger, and small hands.

Family history revealed that her 37-year-old father had similar episodes that began during adolescence. He was hospitalized for three quadriplegia attacks, and he took potassium supplement 600 mg once or twice a year when he experienced generalized weakness attacks. His potassium levels were low (lowest: 2.2 mmol/L) during his acute episodes. His medical history included transient hyperthyroidism but no cardiac symptoms. Since age 25, he had not experienced a paralysis severe enough to be hospitalized but continued to have periodic paralysis once every few years. His paralysis events were mild and improved without going to the hospital by taking a potassium supplement, applying hot compresses, and receiving massages. However, from March to May 2020, at the age of 35, he was hospitalized almost every 2 weeks with severe general paralysis during which he could not even move his fingers. His blood thyroid function test indicated hyperthyroidism. After taking thyroid medication for 2–3 months, he underwent three thyroid function tests, all of which were within the normal range. He remained well without any uncomfortable symptoms after that. However, detailed records (i.e., thyroid hormone levels) were not obtained as data by the hospital at that time. His height and weight were 163.5 cm (≈3rd percentile) and 51.25 kg, respectively. Hypertelorism, small mandible, tooth crowding, and small hands were also observed. **Table 1** provides a summary of the clinical characteristics of our patients.

## GENETIC TESTING

Genomic DNA was extracted from peripheral blood samples using the QIAsymphony DSP DNA mini kit (Qiagen, Hilden, Germany). Sequencing was performed using a TruSight One Expanded sequencing panel (Illumina Inc., San Diego, CA, USA) consisting of 6,700 genes associated with known Mendelian genetic disorders on a NextSeq 550 (Illumina, CA, USA). The obtained sequence reads were aligned to the human reference genome sequence 19 (hg19) from University of California, Santa Cruz (UCSC) [original GRCh37 from National Center for Biotechnology Information (NCBI), February 2009] using Burrows–Wheeler Aligner (BWA version 0.7.7-isis-1.0.0). The BAM files were then processed by base quality recalibration, duplicate removal, and local realignment following the Genome Analysis Toolkit (GATK version 1.6-23-gf0210b3) for variant calling. Annotation and filtering were performed with Illumina Annotation Service (IAS version 3). Clinical exome sequencing revealed a novel causative heterozygous missense Chr17(GRCh37):g.68171593A>T, NM_000891.2:c.413A>T, p.(Glu138Val) variant in *KCNJ2*, which was thought to be causal of ATS syndrome. This domain (H5, intramembrane pore loop) is highly preserved, and the mutation is predicted to be damaging by the protein function prediction software SIFT and PolyPhen-2. It was not present in the general population databases such as the Genome Aggregation Database (gnomAD; https://gnomad.broadinstitute.org/), the Single Nucleotide Polymorphism Database (dbSNP; https://www.

**TABLE 1 |** Clinical details of ATS patients with pathogenic/likely pathogenic variants in Kir2.1 at position 138.

| Patients | Causative variant *KCNJ2* in amino acid 138 | Sex/age at diagnosis (years) | Age with first symptoms (years) | Typical triad of Andersen–Tawil syndrome | | | Other additional features |
|---|---|---|---|---|---|---|---|
| | | | | Periodic paralysis | Cardiac arrhythmia | Dysmorphic features | |
| Proband | E138V | F/15 | 7 | Yes | Yes (long QT interval) | Short stature (<3 percentile), hypertelorism, micrognathia, dental crowding, clinodactyly of the fiftth finger, small hands | Cerebrovascular abnormality (congenital hypoplasia of both internal carotid arteries) |
| Proband's father | E138V | M/37 | 19 | Yes | No | Short stature (163.5 cm), hypertelorism, micrognathia, dental crowding, small hands | Transient hyperthyroidism and ictal exacerbation of periodic paralysis |
| Kostera-Pruszczyk et al. (10), same position variant case as our patients | E138K | M/5 | *In utero* | Yes | Yes (long QT interval) | Short stature (<3 percentile), hypertelorism, micrognathia, clinodactyly of the fifth toe, syndactyly of the second and third toes | Not described |

ncbi.nlm.nih.gov/snp/), ClinVar (http://www.ncbi.nlm.nih.gov/clinvar/), Human Gene Mutation Database (HGMD; http://www.hgmd.cf.ac.uk/ac/index.php), a public archive of the relationship between human variants and phenotypes, and the Korean Reference Genome Database (KRGDB; http://coda.nih.go.kr/coda/KRGDB/index.jsp) as an ethnic-specific Korean population database. We could not find the variant in our case in any previous reports. The identified variants were assessed according to the standards and guidelines for the interpretation of sequence variants in Mendelian disorders of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology (ACMG/AMP) (11), with PM2 (absent from population database), PM5 (novel missense change at an amino acid residue where a different missense change determined to be pathogenic has been seen before), PP2 (missense variants are a common mechanism of disease), PP3 (pathogenic computational evidence), and PP4 (phenotype specific for gene) relevant evidence. Therefore, we classified the variant as likely pathogenic. It was validated by Sanger sequencing (**Figure 1A**) and then confirmed diagnosed as ATS. Because a causative variant was identified in the proband, Sanger sequencing of the causative variant was also performed on her symptomatic father, and the same variant that was consistent with his history of periodic paralysis were identified (**Figure 1B**). **Figure 1C** shows the Sanger sequencing result of an unaffected control on this position. No other pathogenic variants in genes associated with other periodic paralysis disorders (*CACNA1S*, *SCN4A*, or *KCNJ18*) were identified in the proband. The father's parents and two brothers were asymptomatic until now. Although the proband's 11-year-old brother and 4-year-old sister had no symptoms, genetic testing was recommended to all at-risk family members during genetic counseling.

## DISCUSSION

In this case report, we identified two ATS patients in a family with the novel *KCNJ2* c.413A>T variant. Our proband had all triad symptoms, while her father had only two symptoms and no cardiac symptoms. It is very challenging to diagnose ATS early due to the rarity and high degree of phenotypic variability. Fifty-eight to seventy-eight percent of ATS patients present with all three symptoms (12). Moreover, patients can present with only one of the triad symptoms (5). Dysmorphic features are usually subtle in ATS patients (13), as shown in our patients (we recognized retrospectively), and are easy to overlook unless suspected of ATS. Therefore, ATS should always be considered even in patients with only periodic paralysis or cardiac arrhythmias.

Approximately 60% of ATS cases have causal variants in *KCNJ2* (6, 14). Most *KCNJ2* pathogenic variants in ATS are missense variants (1, 6) that result in loss of function of the Kir2.1 channel and is expressed in skeletal muscle, heart, and bone. This channel plays a role in prolonged depolarization of the action potential, and thereby cause periodic paralysis and cardiac arrhythmia (14). Additionally, *KCNJ2* variants can affect skeletal dysmorphic features because *KCNJ2* is expressed during the early stages of craniofacial development in *Xenopus* and mice (15). However, as reports by variable penetrance of ATS cases suggest, a subset of individuals with pathogenic variants do not exhibit any associated symptoms (16–18). A subset of *KCNJ2* pathogenic variants that cause ATS is identified at the phosphatidylinositol-4,5-biphosphate (PIP2) binding site, a known regulator of the Kir2.1 channel. Another subset of *KCNJ2* pathogenic variants is found in the pore-forming loop of Kir2.1. In addition, pathogenic variants in regions such as slide helix have also been reported (5, 19–21), suggesting various mechanisms. The p.(Glu138Val) variant in our patients was located in the pore region of the Kir2.1 channel. Another study showed the same position but other missense change (c.412G>A, p.E138K) as our case variant (c.413A>T, p.E138V) (10). Although not able to perform physiological studies to identify the nature of the alteration in channel function, other studies showed that this subunit pathogenic variant was unable to form functional

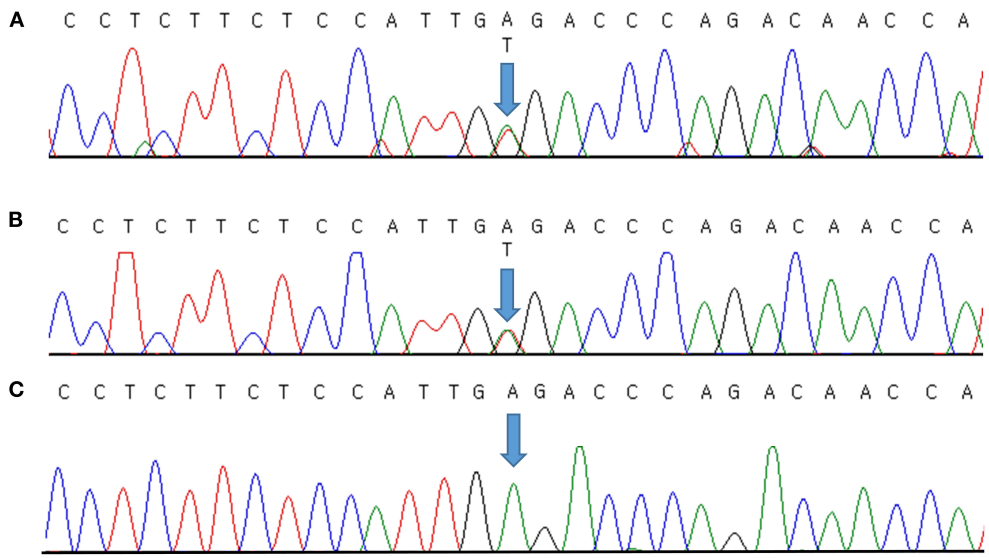



**FIGURE 1 |** Molecular characterization of a family with Andersen–Tawil syndrome. Sequence chromatogram showing the position of the c.413A>T transversion (NM_000891.2) that led to the p.(Glu138Val) missense variant. The variant was heterozygous in the proband **(A)** and her father **(B)**. Sanger sequencing result of an unaffected control **(C)**. Variants are annotated according to The Human Genome Variation Society nomenclature.

homomultimeric channels, abolishing channel function (10, 18). And Yang et al. also described that position 138, a negatively charged residue in the pore loop of Kir2.1, had a strong effect on the stability of the pore loop (22, 23). Additional *in silico* analyses predicted deleterious (SIFT score: 0.05), probably damaging (polyphen-2 score: 1.000), pathogenic (REVEL score: 0.972), and disease-causing (MutationTaster) effects; thus, this variant may cause complete loss of channel function.

We found additional features in our patients. The proband had congenital hypoplasia of the ICA, identified by brain MRI. Although we did not find any reports that also reported vascular abnormalities in ATS, central nervous system involvement including mild learning disability was frequently observed in ATS patients. Because potassium channels contribute to angiogenesis in cancer (24), it is worth investigating a possible association between cerebrovascular development and *KCNJ2* variant. Meanwhile, the proband's father had a history of transient hyperthyroidism with exacerbated periodic paralysis that was consistent with previous studies' findings. They demonstrated that the paralysis symptoms were aggravated in patients with ATS secondary to thyrotoxicosis or co-existence of both disorders (25, 26). The heterogeneous phenotypes between the proband and her father demonstrated the weak genotype–phenotype correlation in ATS. This is further evidenced by a previously reported case with the missense variant of same position 138 who had earlier onset of more severe periodic paralysis (1 year old) and arrhythmia (*in utero*) than our proband (7 years old).

In summary, we identified a novel causative variant in *KCNJ2* in ATS patients and provided detailed clinical findings to expand the genotype–phenotype correlation. The heterogeneous phenotype of ATS should alert physicians to perform genetic analysis for ATS even in patients with only one or two among the typical triad symptoms.

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Seoul St. Mary's Hospital Institutional Review Board. Written informed consent for participation was not provided by the participants' legal guardians/next of kin: Because the patient did not return to our hospital, written consent could not be obtained. However, it has been approved by the Ethics Committee of our hospital, and the patient's identification information is not included in this case.

## AUTHOR CONTRIBUTIONS

JY and MyK wrote the manuscript. JY, KK, MiK, and GL conducted genetic analyses. All authors contributed to the manuscript.

## REFERENCES


1. Veerapandiyan A, Statland JM, Tawil R. Andersen-Tawil syndrome. In: Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJH, Mirzaa G, Amemiya A, editors. *GeneReviews*. Seattle (WA) (1993). Available at: http://www.ncbi.nlm.nih.gov/books/nbk1264/ (Last updated 2018).
2. Andersen ED, Krasilnikoff PA, Overvad H. Intermittent muscular weakness, extrasystoles, and multiple developmental anomalies. A new syndrome? *Acta Paediatr Scand.* (1971) 60:559–64. doi: 10.1111/j.1651-2227.1971.tb06990.x
3. Stunnenberg BC, Raaphorst J, Deenen JCW, Links TP, Wilde AA, Verbove DJ, et al. Prevalence and mutation spectrum of skeletal muscle channelopathies in the Netherlands. *Neuromuscul Disord.* (2018) 28:402–7. doi: 10.1016/j.nmd.2018.03.006
4. Horga A, Raja Rayan DL, Matthews E, Sud R, Fialho D, Durran SC, et al. Prevalence study of genetically defined skeletal muscle channelopathies in England. *Neurology.* (2013) 80:1472–5. doi: 10.1212/WNL.0b013e31828cf8d0
5. Nguyen HL, Pieper GH, Wilders R. Andersen-Tawil syndrome: clinical and molecular aspects. *Int J Cardiol.* (2013) 170:1–16. doi: 10.1016/j.ijcard.2013.10.010
6. Perez-Riera AR, Barbosa-Barros R, Samesina N, Pastore CA, Scanavacca M, Daminello-Raimundo R, et al. Andersen-Tawil syndrome: a comprehensive review. *Cardiol Rev.* (2021) 29:165–77. doi: 10.1097/CRD.0000000000000326
7. Yoon G, Oberoi S, Tristani-Firouzi M, Etheridge SP, Quitania L, Kramer JH, et al. Andersen-Tawil syndrome: prospective cohort analysis and expansion of the phenotype. *Am J Med Genet A.* (2006) 140:312–21. doi: 10.1002/ajmg.a.31092
8. Plaster NM, Tawil R, Tristani-Firouzi M, Canun S, Bendahhou S, Tsunoda A. et al. Mutations in Kir21 cause the developmental and episodic electrical phenotypes of Andersen's syndrome. *Cell.* (2001) 105:511–9. doi: 10.1016/S0092-8674(01)00342-7
9. Kokunai Y, Nakata T, Furuta M, Sakata S, Kimura H, Aiba T. et al. A Kir34 mutation causes Andersen-Tawil syndrome by an inhibitory effect on Kir21. *Neurology.* (2014) 82:1058–64. doi: 10.1212/WNL.0000000000000239
10. Kostera-Pruszczyk A, Potulska-Chromik A, Pruszczyk P, Bieganowska K, Miszczak-Knecht M, Bienias P, et al. Andersen-Tawil syndrome: report of 3 novel mutations and high risk of symptomatic cardiac involvement. *Muscle Nerve.* (2015) 51:192–6. doi: 10.1002/mus.24293
11. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med.* (2015) 17:405–24. doi: 10.1038/gim.2015.30
12. Statland JM, Fontaine B, Hanna MG, Johnson NE, Kissel JT, Sansone VA, et al. Review of the diagnosis and treatment of periodic paralysis. *Muscle Nerve.* (2018) 57:522–30. doi: 10.1002/mus.26009
13. Almuqbil M, Srour M. Child neurology: Andersen-Tawil syndrome. *Neurology.* (2015) 84:e78–80. doi: 10.1212/WNL.0000000000001377
14. Maggi L, Bonanno S, Altamura C, Desaphy JF. Ion channel gene mutations causing skeletal muscle disorders: pathomechanisms and opportunities for therapy. *Cells.* (2021) 10:1521. doi: 10.3390/cells10061521
15. Adams DS, Uzel SG, Akagi J, Wlodkowic D, Andreeva V, Yelick PC, et al. Bioelectric signalling *via* potassium channels: a mechanism for craniofacial dysmorphogenesis in KCNJ2-associated Andersen-Tawil Syndrome. *J Physiol.* (2016) 594:3245–70. doi: 10.1113/JP271930
16. Andelfinger G, Tapper AR, Welch RC, Vanoye CG, George AL, Benson DW. KCNJ2 mutation results in Andersen syndrome with sex-specific cardiac and skeletal muscle phenotypes. *Am J Hum Genet.* (2002) 71:663–8. doi: 10.1086/342360
17. Deeb R, Veerapandiyan A, Tawil R, Treidler S. Variable penetrance of Andersen-Tawil syndrome in a family with a rare missense KCNJ2 mutation. *Neurol Genet.* (2018) 4:e284. doi: 10.1212/NXG.0000000000000284
18. Tristani-Firouzi M, Jensen JL, Donaldson MR, Sansone V, Meola G, Hahn A, et al. Functional and clinical characterization of KCNJ2 mutations associated with LQT7 (Andersen syndrome). *J Clin Invest.* (2002) 110:381–8. doi: 10.1172/JCI15183
19. Donaldson MR, Jensen JL, Tristani-Firouzi M, Tawil R, Bendahhou S, Suarez WA, et al. PIP2 binding residues of Kir21 are common targets of mutations causing Andersen syndrome. *Neurology.* (2003) 60:1811–6. doi: 10.1212/01.WNL.0000072261.14060.47
20. Bendahhou S, Donaldson MR, Plaster NM, Tristani-Firouzi M, Fu YH. Ptacek LJ. Defective potassium channel Kir21 trafficking underlies Andersen-Tawil syndrome. *J Biol Chem.* (2003) 278:51779–85. doi: 10.1074/jbc.M310278200
21. Barron-Diaz DR, Totomoch-Serra A, Escobar-Cedillo RE, Garcia-Gutierrez A, Reyes-Quintero AE, Villegas Daviran SE, et al. Andersen-Tawil syndrome with high risk of sudden cardiac death in four mexican patients. Cardiac and extra-cardiac phenotypes. *Rev Invest Clin.* (2021) 73::145–53. doi: 10.24875/RIC.20000310
22. Zangerl-Plessl EM, Qile M, Bloothooft M, Stary-Weinzinger A, van der Heyden MAG. Disease associated mutations in KIR proteins linked to aberrant inward rectifier channel trafficking. *Biomolecules.* (2019) 9:650. doi: 10.3390/biom9110650
23. Yang J, Yu M, Jan YN, Jan LY. Stabilization of ion selectivity filter by pore loop ion pairs in an inwardly rectifying potassium channel. *Proc Natl Acad Sci USA.* (1997) 94:1568–72. doi: 10.1073/pnas.94.4.1568
24. Umaru B, Pyriochou A, Kotsikoris V, Papapetropoulos A, Topouzis S. ATP-sensitive potassium channel activation induces angiogenesis *in vitro* and *in vivo*. *J Pharmacol Exp Ther.* (2015) 354:79–87. doi: 10.1124/jpet.114.222000
25. Diaz-Manera J, Querol L, Alejaldre A, Rojas-Garcia R, Ramos-Fransi A, Gallardo E, et al. Severe exacerbation of Andersen-Tawil syndrome secondary to thyrotoxicosis. *J Hum Genet.* (2014) 59:465–6. doi: 10.1038/jhg.2014.43
26. Iskander B, Malik BH, Cancarevic I. Thyrotoxic periodic paralysis with features of Andersen-Tawil syndrome: a case report and literature review. *Cureus.* (2020) 12:e8169. doi: 10.7759/cureus.8169


**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.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

*Copyright © 2022 Yim, Kim, Kim, Lee and Kim. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*

frontiers in Genetics

CASE REPORT
published: 03 February 2022
doi: 10.3389/fgene.2021.780363

# Case Report: A Novel Homozygous Mutation in *MYF5* Due to Paternal Uniparental Isodisomy of Chromosome 12 in a Case of External Ophthalmoplegia With Rib and Vertebral Anomalies

*Qianqian Li[1]*, Xiaofan Zhu[1], Chenguang Yu[2], Lin Shang[3], Ranran Li[4], Xia Wang[5], Yaping Yang[5], Jingjing Meng[1] and Xiangdong Kong[1]**

[1]*Genetics and Prenatal Diagnosis Center, Department of Obstetrics and Gynecology, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, China,* [2]*Key Laboratory of Molecular Biophysics of the Ministry of Education, College of Life Science and Technology and Center for Human Genome Research, Huazhong University of Science and Technology, Wuhan, China,* [3]*Department of Foot and Ankle Surgery, Zhengzhou Orthopedic Hospital, Zhengzhou, China,* [4]*School of Life Science and Technology, Xinxiang Medical University, Xinxiang, China,* [5]*AiLife Diagnostics, Inc., Houston, TX, United States*

OPEN ACCESS

**Edited by:**
*Kathleen M. Gorman, Temple Street Children's University Hospital, Ireland*

**Reviewed by:**
*Mehmet Ali Ergün, Gazi University, Turkey*
*Ludmila Matyakhina, GeneDx, United States*

***Correspondence:**
*Qianqian Li*
*37030810@qq.com*
*Xiangdong Kong*
*kongxd@263.net*

**Specialty section:**
*This article was submitted to Genetics of Common and Rare Diseases, a section of the journal Frontiers in Genetics*

**Received:** *21 September 2021*
**Accepted:** *30 December 2021*
**Published:** *03 February 2022*

**Citation:**
*Li Q, Zhu X, Yu C, Shang L, Li R, Wang X, Yang Y, Meng J and Kong X (2022) Case Report: A Novel Homozygous Mutation in* MYF5 *Due to Paternal Uniparental Isodisomy of Chromosome 12 in a Case of External Ophthalmoplegia With Rib and Vertebral Anomalies. Front. Genet. 12:780363. doi: 10.3389/fgene.2021.780363*

External ophthalmoplegia with rib and vertebral anomalies (EORVA) is characterized by congenital nonprogressive external ophthalmoplegia, ptosis, scoliosis, torticollis, vertebral, and rib anomalies, caused by homozygous mutations in the myogenic factor 5 gene (*MYF5*) located on chromosome 12q21.31. Uniparental disomy (UPD) is a rare inheritance of a pair of chromosomes originating from only one parent. This study describes a case of an 8-year-old boy with ptosis, scoliosis, and dysmorphic hypoplastic ribs with fusion anomalies. Trio-based exome sequencing (trio-ES) identified a novel homozygous mutation c.191delC (p.Ala64Valfs*33) in *MYF5* in the proband, with the father being heterozygous and the mother wild-type, as verified by Sanger sequencing. UPD identified from trio-ES variant call format data suggested the possibility of paternal UPD of chromosome 12 (UPD12pat) in the proband, further confirmed to be a complete isodisomy type of UPD by genome-wide single nucleotide polymorphism array. MYF5 was significantly downregulated by 69.14% (**$p < 0.01$) in HeLa cells transfected with mutant *MYF5* containing c.191delC compared to those transfected with the wild-type *MYF5*, resulting in a truncated protein with a size of ~20 kDa. In conclusion, this study identified a novel homozygous mutation in *MYF5*, broadening the genetic spectrum of EORVA and further deepening the understanding of this rare disease.

**Keywords: external ophthalmoplegia with rib and vertebral anomalies, trio-based exome sequencing, paternal uniparental isodisomy, chromosome 12, the myogenic factor 5 gene, homozygous mutation**

## INTRODUCTION

External ophthalmoplegia with rib and vertebral anomalies (EORVA; MIM# 618155), an extremely rare autosomal recessive disorder, is characterized by congenital nonprogressive external ophthalmoplegia and ptosis, with torticollis and scoliosis developing during childhood. In addition, patients may present with hypoplastic or missing ribs with fusion anomalies (Di Gioia et al., 2018). EORVA is caused by homozygous mutations in the myogenic factor 5 gene (*MYF5*; MIM* 159990) located on chromosomal 12q21.31. MYF5, a transcriptional activator encoded by *MYF5*, is a member of the Myc-like basic helix-loop-helix transcription factor family that plays an important role in promoting the transcription of muscle-specific target genes and hence, muscle differentiation (Wang et al., 1996). To date, only two homozygous variations in *MYF5* have been reported (Di Gioia et al., 2018).

Uniparental disomy (UPD), first introduced by Engel (1980), is a non-traditional Mendelian inheritance pattern, in which a pair of chromosomes are inherited from only one parent. It mainly consists of two subtypes: heterodisomy (hUPD), where a pair of non-identical chromosomes are contributed by one parent due to an error during meiosis I, and isodisomy (iUPD), where a chromosome from one parent is duplicated due to an error during meiosis II (Benn, 2021). Recently, Nakka et al. identified 675 instances of UPD across 4,400,363 consented research participants from the personal genetics company 23andMe, Inc., and 431,094 UK Biobank participants, and estimated that the prevalence of UPD on all chromosomes was 1:2000 (0.1%) (Nakka et al., 2019). The clinical consequences of UPD depend on the chromosomes involved, including imprinting diseases, recessive Mendelian diseases, and mosaic aneuploidy associated with diseases (Scuffins et al., 2021).

In this study, an 8-year-old boy with ptosis, scoliosis, and dysmorphic hypoplastic ribs with fusion anomalies was identified with a novel homozygous variation in *MYF5* due to paternal iUPD of chromosome 12 (iUPD12pat).

## Case Presentation

The proband was an 8-year-old boy with ptosis since birth (**Figure 1A**). At the age of 1 year, the proband was diagnosed with ptosis by magnetic resonance imaging of the ocular motor nerve. At the age of 6 years, the proband underwent reconstructive surgery, and the symptom of ptosis was improved (**Figure 1B**, before surgery; **Figure 1C**, after surgery). And at the age of 6 years, the proband was further diagnosed with scoliosis by computed tomography at a local hospital (data not shown) and further diagnosed with scoliosis and dysmorphic hypoplastic ribs with fusion anomalies, however, with no vertebral anomalies, by digital radiography (DR) at Zhengzhou Orthopedic Hospital at

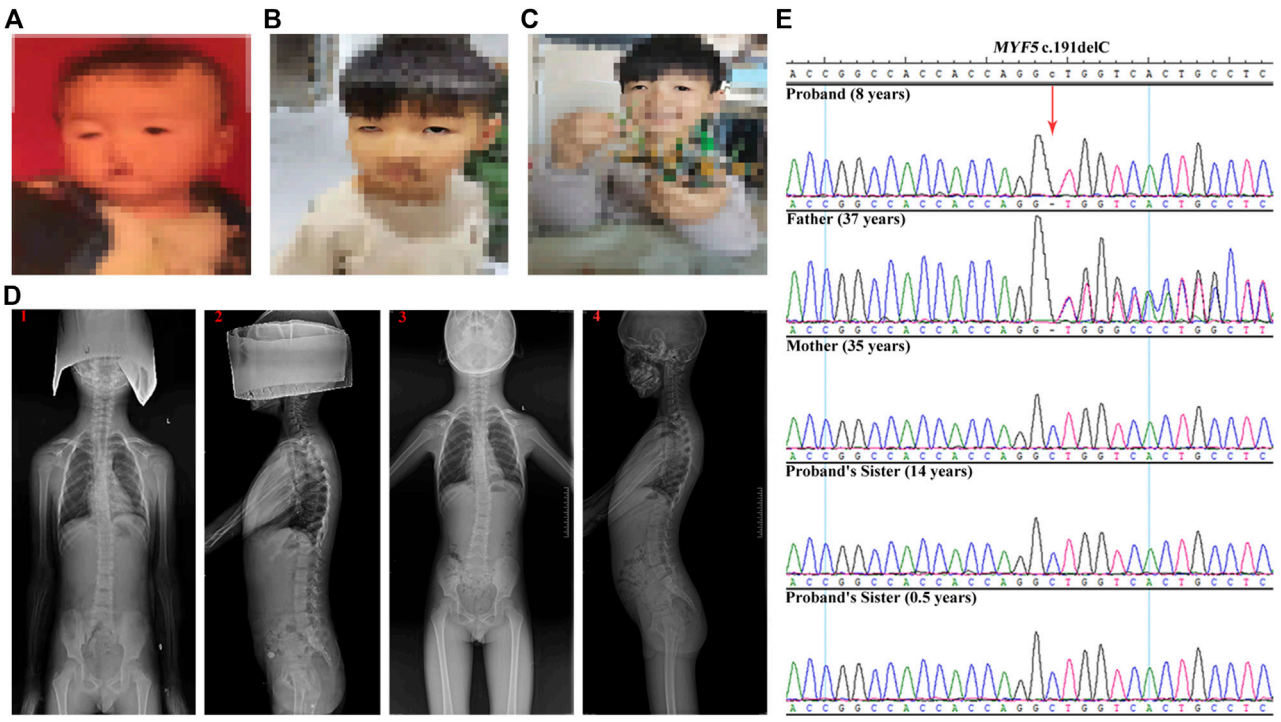



**FIGURE 1 |** Clinical features of the proband and Sanger sequencing results of the family. **(A–C)** The ptosis phenotype at birth, at the age of 6 years before the reconstructive surgery, and at the age of 6 years after the reconstructive surgery, respectively, showing the improvement in ptosis symptom for the proband. **(D)** DR images acquired at the age of 7 years (1 and 2) and at the age of 8 years (3 and 4), showing the significant aggravation of the scoliosis condition for the proband. **(E)** Sanger sequencing results of c.191delC for the family. The proband is homozygous. The father is heterozygous, whereas the mother and the two sisters are wild-type.

the age of 7 years (**Figure 1D**, **1** and **2**). At the age of 8 years, the DR images suggested a significant aggravation of scoliosis for the proband (**Figure 1D**, **3** and **4**).

# METHODS

## Ethics Approval

The family provided written informed consent to carry out this study. This study was approved by the appropriate local institutional review board on human subject research at the First Affiliated Hospital of Zhengzhou University.

## Quantitative Fluorescent Polymerase Chain Reaction

Genomic DNA was extracted from 500 μL of peripheral blood using the Lab-Aid Nucleic Acid (DNA) Isolation Kit (Zeesan, Xiamen, China). The genetic relationship of the proband and the parents was confirmed by QF-PCR using the Goldeneye™ DNA ID System 20A Kit (Peoplespot, Beijing, China).

## Trio-Based Exome Sequencing

The methods of experiment and data analysis used for trio-ES have been described in detail in our previous study (Li et al., 2021). The candidate mutation in *MYF5* was verified by Sanger sequencing using the primer pairs provided in **Supplementary Table S1**.

## Bioinformatics Analysis

The conservation analysis was performed using the MultAlin tool (http://multalin.toulouse.inra.fr/multalin/multalin.html). The variant frequency in different populations was obtained from 1000G (https://www.internationalgenome.org/), gnomAD (http://gnomad-sg.org/), and ExAC databases (http://exac.broadinstitute.org). SIFT (http://provean.jcvi.org/protein_batch_submit.php?species=human), PolyPhen2 (http://genetics.bwh.harvard.edu/pph2/index.shtml), and MutationTaster (https://www.mutationtaster.org/) were used to predict the deleterious effects of the variant on the protein structure and function.

UPD was identified online through AilisNGS® ("https://www.ailifeus.com/#/products/ngs". Houston, USA) from the trio-ES variant call format (vcf) data as per the method previously described (King et al., 2014).

## Genome-Wide Single Nucleotide Polymorphism Array

Genome-wide SNP array, including fragmentation, labelling, and hybridization, was carried out using DNA samples from the proband and the father using Illumina HumanCytoSNP-12 v2.1 BeadChip (California, USA) on Illumina iScan. The methods of experiment and data analysis have been previously described (Bruno et al., 2011).

## Cell Line, Plasmids, and Transfection

HeLa cells (human epithelial cervix carcinoma cells, American Type Culture Collection) were cultured in Dulbecco's modified Eagle's medium (Gibco Life Technologies, Gaithersburg, MD, USA) supplemented with 10% fetal bovine serum (Gibco Life Technologies) and 5% $CO_2$ at 37°C.

The full-length coding region of *MYF5* was sub-cloned into p3×FLAG-CMV-10 (p3×FLAG-CMV-10-*MYF5*-WT), whereas the mutant plasmid was constructed using the PCR-based site-directed mutagenesis (p3×FLAG-CMV-10-*MYF5*-MU) using the primers given in **Supplementary Table S1**.

HeLa cells were cultured for 24 h and then transfected with 2 μg of p3×FLAG-CMV-10-*MYF5*-WT, p3×FLAG-CMV-10-*MYF5*-MU, or p3×FLAG-CMV-10.

## Western Blotting

Following 48 h incubation with 5% $CO_2$ at 37°C, the cells were collected and lysed in lysis buffer. Then, the cell lysate was separated using 10% sodium dodecyl sulfate-polyacrylamide gel electrophoresis. After blocking with 5% skimmed milk powder in 1 × TBST buffer (10 mM Tris, 150 mM NaCl, and 0.05% Tween-20, pH7.5), the membrane was incubated with the primary antibody (anti-DDDDK tag, M185-3L, Medical & Biological Laboratories Co., Ltd, Nagoya, Japan) at 4°C overnight. Following washing with 1 × TBST buffer, the membrane was incubated with the secondary antibody (BL001A, Biosharp, Anhui, China). β-Actin (GB11001, Servicebio, Wuhan, China) was used as a loading control.

## Statistical Analysis

Data are presented as the mean ± SEM. Statistical analysis was performed using GraphPad Prism 6. Student's two-tailed *t*-test was used for between-group comparisons. Statistically significant differences were considered at $p < 0.05$ ($^{**}p < 0.01$).

# RESULTS

The genetic relationship between the proband and the parents was confirmed by QF-PCR (**Supplementary Table S2** and **Supplementary Figure S1**). The quality control of the trio-ES data is summarized in **Supplementary Table S3**. After filtering, only the homozygous candidate variant c.191delC (p.Ala64Valfs*33) in *MYF5* (NM_005593.3) was identified in the proband (reference allele/alternative allele, ref/alt: 0/78), with the father heterozygous (ref/alt: 40/32) and the mother wild-type (**Supplementary Figures S2–4**), as confirmed by Sanger sequencing (**Figure 1E**).

Online prediction using MultAlin indicated that p.Ala64 was highly conservative across different species (**Figure 2A**, black box), and c.191delC was absent in 1000G, ExAC, and gnomAD databases (**Supplementary Table S4**). UPD identified based on the trio-ES vcf data suggested the possibility of UPD12pat in the proband (**Supplementary Figure S5**).

For the proband, log R ratio and B allelic frequency confirmed that UPD was a complete paternal UPD12 spanning 12p13.33-q24.33, named arr [GRCh37/hg19]12p13.33q24.33 (413,635-133,272,968)×2 hmz according to the International System for Human Cytogenomic Nomenclature (ISCN 2016), including the paternal mutant allele of *MYF5* (**Figure 2D**, the proband; **2E**, the

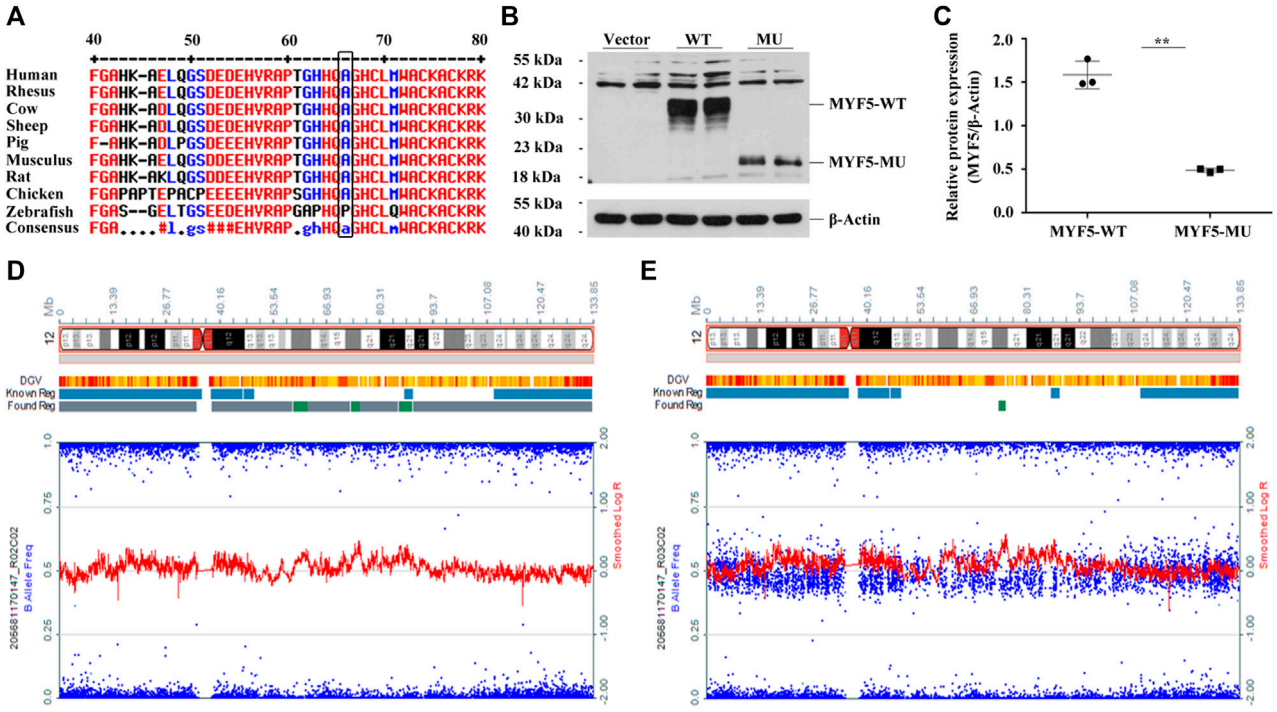



**FIGURE 2 |** Functional study of c.191delC and genome-wide SNP array results of the proband and the father. **(A)** p.Ala64 is highly conservative across different species. **(B)**, **(C)** Western blotting indicated the significant downregulation of MYF5 by 69.14% (**$p < 0.01$) in cells transfected with p3×FLAG-CMV-10-*MYF5*-MU compared to those transfected with p3×FLAG-CMV-10-*MYF5*-WT, resulting in a truncated protein with a size of ~20 kDa. Vector: p3×FLAG-CMV-10. **(D)**, **(E)** Genome-wide SNP array results of the proband and the father, respectively (only chromosome 12 shown).

father). All SNPs are listed in **Supplementary Excel S1** (205681170147_R02C02, the proband; 205681170147_R03C02, the father).

Western blotting confirmed the significant downregulation of MYF5 by 69.14% (**$p < 0.01$) in cells transfected with p3×FLAG-CMV-10-*MYF5*-MU compared to those transfected with p3×FLAG-CMV-10-*MYF5*-WT, resulting in a truncated protein with a size of ~20 kDa (**Figures 2B,C**). These findings suggested that the mutation of c.191delC drastically affects the biological function of MYF5.

## DISCUSSION

External ophthalmoplegia with rib and vertebral anomalies is a rare autosomal recessive disorder characterized by congenital ophthalmoplegia with scoliosis and vertebral and rib anomalies, with only five individuals (three families) reported to date (Di Gioia et al., 2018). In this study, we report a novel homozygous variant in *MYF5* (c.191delC) due to iUPD.

The proband in this study was an 8-year-old boy diagnosed with ptosis, scoliosis, and dysmorphic hypoplastic ribs with fusion anomalies. To the best of our knowledge, individuals with EORVA reported previously (Traboulsi et al., 2000; Di Gioia et al., 2018) and the patient in this study had similar phenotypes regarding external ophthalmoplegia, rib, and vertebral anomalies. In most cases, unilateral or bilateral ptosis was present, including our current case. Additionally, delayed motor development can occur, which was previously reported in one EORVA case but not observed in our case. Cognition was normal in all cases.

For the proband, trio-ES suggested a novel homozygous variant c.191delC in exon 1 of *MYF5*. This variant was heterozygous in the father and absent in the mother (**Supplementary Figures S2–4**), as further confirmed by Sanger sequencing (**Figure 1E**). To identify this non-Mendelian inheritance, UPD was analyzed based on the trio-ES vcf data. The result indicated that the homozygous state of c.191delC in the proband may be presumably attributed to UPD12pat (**Supplementary Figure S5A**), which was finally confirmed to be a complete iUPD by genome-wide SNP array (**Figures 2D,E** and **Supplementary Excel S1**).

The novel variant c.191delC results in p.Ala64Valfs*33, causing a frameshift mutation and premature termination of translation at codon 96. To our knowledge, this is the third pathogenic homozygous mutation reported in EORVA patients and the first mutation caused by UPD in *MYF5*. The biological function of MYF5 was severely affected because of the mutant-type (**Figures 2B,C**). Therefore, regular clinical observations need to be conducted on the proband to evaluate if the clinical symptoms of EOVAR will worsen in the future, though undergoing reconstructive surgery for ptosis.

UPD is the inheritance in which both homologous chromosomes are from one parent with no representative copy from the other, occurring as iUPD, hUPD, a combination of both, or only

chromosome segment(s) (Schroeder et al., 2014). In 2020, a UPD study of trios in a clinical cohort of suspected genetic diseases revealed that the incidence of UPD was about 2:1,000 (0.2%) (Yauy et al., 2020). Later in 2021, in a population of 32,067 clinical exome trios, 16 cases were related to a positive test result through homozygous sequence variations (Scuffins et al., 2021).

Complete iUPD, an underestimated cause of recessive Mendelian disorders, is the inheritance of two identical chromosomes that are the same at all polymorphic sites along their length (Middleton et al., 2006), and is the predominant UPD type observed in the largest chromosomes (Del Gaudio et al., 2020). Roberts et al. identified a homozygous nonsense mutation c.1618A > T (p.Lys540*) in *CD45* (NM_002838) in a severe combined immunodeficiency patient caused by maternal iUPD1 (iUPD1mat) (Roberts et al., 2012). Chen et al. represented a case of epileptic disorder associated with a novel mutation c.2873_c.2874delCT (p.Thr958Thrfs*17) in homozygosity in *CNTN2* (NM_005076) due to UPD1mat (Chen et al., 2021). Several autosomal recessive disorders have been reported to be owing to the homozygous mutations on chromosome 2 disclosed by iUPD2pat (Kantarci et al., 2008; Dasi et al., 2016; Hara-Isono et al., 2021). Sasaki et al. reported a 3M syndrome patient with a homozygous mutation c.2975G > C (p.Arg992Pro) in *CUL7* (NM_014780) attributed to iUPD6mat (Sasaki et al., 2011). Cho et al. unmasked a homozygous mutation c.1120C > G (p.Tyr400*) in *SUOX* (NC_000012.11) in a patient with isolated sulfite oxidase deficiency resulting from UPD12pat (Cho et al., 2013). Wiszniewski et al. revealed a homozygous mutation c.1148delC (p.Thr383Ilefs*13) in *CNGB3* (NM_019098) in a case with achromatopsia associated with UPD14mat (Wiszniewski et al., 2007). SoehnSoehn et al. described four novel homozygous mutations in *FA2H* (NM_024306) in four unrelated families with spastic paraplegia type 35 because of UPD16 (SoehnSoehn et al., 2016).

In conclusion, to the best of our knowledge, for the first time, we identified a novel homozygous mutation in *MYF5* due to iUPD12pat in a case of EORVA, a rare event and enriching *MYF5* gene variation spectrum, further suggesting that UPD of any chromosome is associated with an increased risk of recessive disease because it may affect children when only one parent is a carrier of a pathogenic variant. Additionally, the results re-emphasized the clinical importance of UPD in case the results are inconsistent with recessive inheritance.

## DATA AVAILABILITY STATEMENT

The datasets for this article are not publicly available due to concerns regarding participant/patient anonymity. Requests to access the datasets should be directed to the corresponding authors.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the guidelines set forth by the Declaration of Helsinki. This study was also conducted in accordance with the International Ethical Guidelines for Biomedical Research Involving Human Subjects (CIOMS). The family provided written informed consent to participate in this research. 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 minor(s)' legal guardian/next of kin for the publication of any potentially identifiable images or data included in this article.

## AUTHOR CONTRIBUTIONS

QL, XZ, CY, and XK designed the research. QL, XZ, CY, and JM performed the experiments. LS supported the DR images of the proband. QL, XZ, CY, XW, and YY analyzed the data. QL wrote the manuscript. XZ, RL, XW, YY, and XK revised the manuscript. All authors contributed to the article and approved the submitted version.

## FUNDING

This study was supported by the National Natural Science Foundation of China (no. 82000321) and the Henan Educational Committee Program for Science and Technology Development of Universities (22A310022).

## ACKNOWLEDGMENTS

The authors were very grateful to the family for participating in this study.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.780363/full#supplementary-material

## REFERENCES

Benn, P. (2021). Uniparental Disomy: Origin, Frequency, and Clinical Significance. *Prenatal Diagn.* 41 (5), 564–572. doi:10.1002/pd.5837

Bruno, D. L., White, S. M., Ganesamoorthy, D., Burgess, T., Butler, K., Corrie, S., et al. (2011). Pathogenic Aberrations Revealed Exclusively by Single Nucleotide Polymorphism (SNP) Genotyping Data in 5000 Samples Tested by Molecular Karyotyping. *J. Med. Genet.* 48 (12), 831–839. doi:10.1136/jmedgenet-2011-100372

Chen, W., Chen, F., Shen, Y., Yang, Z., and Qin, J. (2021). Case Report: A Case of Epileptic Disorder Associated with a Novel CNTN2 Frameshift Variant in Homozygosity Due to Maternal Uniparental Disomy. *Front. Genet.* 12, 743833. doi:10.3389/fgene.2021.743833

Cho, S. Y., Goh, D. L.-M., Lau, K.-C., Ong, H. T., and Lam, C.-w. (2013). Microarray Analysis Unmasked Paternal Uniparental Disomy of

Chromosome 12 in a Patient with Isolated Sulfite Oxidase Deficiency. *Clinica Chim. Acta* 426, 13–17. doi:10.1016/j.cca.2013.08.013

Dasi, M. A., Gonzalez-Conejero, R., Izquierdo, S., Padilla, J., Garcia, J. L., Garcia-Barberá, N., et al. (2016). Uniparental Disomy Causes Deficiencies of Vitamin K-dependent Proteins. *J. Thromb. Haemost.* 14 (12), 2410–2418. doi:10.1111/jth.13517

Del Gaudio, D., Shinawi, M., Astbury, C., Tayeh, M. K., Deak, K. L., and Raca, G. (2020). Diagnostic Testing for Uniparental Disomy: a Points to Consider Statement from the American College of Medical Genetics and Genomics (ACMG). *Genet. Med.* 22 (7), 1133–1141. doi:10.1038/s41436-020-0782-9

Di Gioia, S. A., Shaaban, S., Elcioglu, B. N. H., Chan, W.-M., Robson, C. D., Ecklund, K., et al. (2018). Recessive MYF5 Mutations Cause External Ophthalmoplegia, Rib, and Vertebral Anomalies. *Am. J. Hum. Genet.* 103 (1), 115–124. doi:10.1016/j.ajhg.2018.05.003

Engel, E. (1980). A New Genetic Concept: Uniparental Disomy and its Potential Effect, Isodisomy. *Am. J. Med. Genet.* 6 (2), 137–143. doi:10.1002/ajmg.1320060207

Hara-Isono, K., Matsubara, K., Hamada, R., Shimada, S., Yamaguchi, T., Wakui, K., et al. (2021). A Patient with Silver-Russell Syndrome with Multilocus Imprinting Disturbance, and Schimke Immuno-Osseous Dysplasia Unmasked by Uniparental Isodisomy of Chromosome 2. *J. Hum. Genet.* 66 (11), 1121–1126. doi:10.1038/s10038-021-00937-7

Kantarci, S., RaggeRagge, N. K., Thomas, N. S., Robinson, D. O., Noonan, K. M., Russell, M. K., et al. (2008). Donnai-Barrow Syndrome (DBS/FOAR) in a Child with a homozygousLRP2mutation Due to Complete Chromosome 2 Paternal Isodisomy. *Am. J. Med. Genet.* 146A (14), 1842–1847. doi:10.1002/ajmg.a.32381

King, D. A., Fitzgerald, T. W., Miller, R., Canham, N., Clayton-Smith, J., Johnson, D., et al. (2014). A Novel Method for Detecting Uniparental Disomy from Trio Genotypes Identifies a Significant Excess in Children with Developmental Disorders. *Genome Res.* 24 (4), 673–687. doi:10.1101/gr.160465.113

Li, Q., Zhu, X., Wang, C., Meng, J., Chen, D., and Kong, X. (2021). Identification of a Rare Case with Nagashima-type Palmoplantar Keratoderma and 18q Deletion Syndrome via Exome Sequencing and Low-Coverage Whole-Genome Sequencing. *Front. Genet.* 12, 707411. doi:10.3389/fgene.2021.707411

Middleton, F. A., Trauzzi, M. G., Shrimpton, A. E., Gentile, K. L., Morley, C. P., Medeiros, H., et al. (2006). Complete Maternal Uniparental Isodisomy of Chromosome 4 in a Subject with Major Depressive Disorder Detected by High Density SNP Genotyping Arrays. *Am. J. Med. Genet.* 141B (1), 28–32. doi:10.1002/ajmg.b.30250

Nakka, P., Pattillo Smith, S., O'Donnell-LuriaDonnell-Luria, A. H., McManus, K. F., Mountain, J. L., Ramachandran, S., et al. (2019). Characterization of Prevalence and Health Consequences of Uniparental Disomy in Four Million Individuals from the General Population. *Am. J. Hum. Genet.* 105 (5), 921–932. doi:10.1016/j.ajhg.2019.09.016

Roberts, J. L., Buckley, R. H., Luo, B., Pei, J., Lapidus, A., Peri, S., et al. (2012). CD45-deficient Severe Combined Immunodeficiency Caused by Uniparental Disomy. *Proc. Natl. Acad. Sci.* 109 (26), 10456–10461. doi:10.1073/pnas.1202249109

Sasaki, K., Okamoto, N., Kosaki, K., Yorifuji, T., Shimokawa, O., Mishima, H., et al. (2011). Maternal Uniparental Isodisomy and Heterodisomy on Chromosome 6 Encompassing a CUL7 Gene Mutation Causing 3M Syndrome. *Clin. Genet.* 80 (5), 478–483. doi:10.1111/j.1399-0004.2010.01599.x

Schroeder, C., Ekici, A. B., Ekici, A. B., Moog, U., Grasshoff, U., Mau-Holzmann, U., et al. (2014). Genome-wide UPD Screening in Patients with Intellectual Disability. *Eur. J. Hum. Genet.* 22 (10), 1233–1235. doi:10.1038/ejhg.2014.63

Scuffins, J., Keller-Ramey, J., Dyer, L., Douglas, G., Torene, R., Gainullin, V., et al. (2021). Uniparental Disomy in a Population of 32,067 Clinical Exome Trios. *Genet. Med.* 23 (6), 1101–1107. doi:10.1038/s41436-020-01092-8

SoehnSoehn, A. S., RattayRattay, T. W., Beck-Wödl, S., Schäferhoff, K., Monk, D., Döbler-Neumann, M., et al. (2016). Uniparental Disomy of Chromosome 16 Unmasks Recessive Mutations of FA2H/SPG35 in 4 Families. *Neurology* 87 (2), 186–191. doi:10.1212/wnl.0000000000002843

Traboulsi, E. I., Lee, B. A., Mousawi, A., Khamis, A. R., and Engle, E. C. (2000). Evidence of Genetic Heterogeneity in Autosomal Recessive Congenital Fibrosis of the Extraocular Muscles. *Am. J. Ophthalmol.* 129 (5), 658–662. doi:10.1016/s0002-9394(99)00467-5

Wang, Y., Schnegelsberg, P. N. J., Dausman, J., and Jaenisch, R. (1996). Functional Redundancy of the Muscle-specific Transcription Factors Myf5 and Myogenin. *Nature* 379 (6568), 823–825. doi:10.1038/379823a0

Wiszniewski, W., Lewis, R. A., and Lupski, J. R. (2007). Achromatopsia: the CNGB3 p.T383fsX Mutation Results from a Founder Effect and Is Responsible for the Visual Phenotype in the Original Report of Uniparental Disomy 14. *Hum. Genet.* 121 (3-4), 433–439. doi:10.1007/s00439-006-0314-y

Yauy, K., de Leeuw, N., Yntema, H. G., Pfundt, R., and Gilissen, C. (2020). Accurate Detection of Clinically Relevant Uniparental Disomy from Exome Sequencing Data. *Genet. Med.* 22 (4), 803–808. doi:10.1038/s41436-019-0704-x

**Conflict of Interest:** XW and YY were employed by AiLife Diagnostics, Inc. 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.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

*Copyright © 2022 Li, Zhu, Yu, Shang, Li, Wang, Yang, Meng and Kong. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*

frontiers in Genetics

CASE REPORT
published: 04 February 2022
doi: 10.3389/fgene.2021.752272

# Prenatal Diagnosis of Fetus With Transaldolase Deficiency Identifies Compound Heterozygous Variants: A Case Report

*Jiaxin Xue[1,2], Jin Han[1]*, Xiaopeng Zhao[3], Li Zhen[1], Shanshan Mei[4], Zhiyang Hu[5] and Xiuzhen Li[6]*

*[1]Prenatal Diagnosis Center, Guangzhou Women and Children's Medical Center, Guangzhou, China, [2]Department of Obstetrics and Gynecology, Guangzhou Medical University, Guangzhou, China, [3]Division of Neonatology, Guangzhou Women and Children's Medical Center, Guangzhou, China, [4]Division of Obstetrics, Guangzhou Women and Children's Medical Center, Guangzhou, China, [5]Shenzhen People's Hospital, Shenzhen, China, [6]Division of Endocrinology, Guangzhou Women and Children's Medical Center, Guangzhou, China*

Transaldolase (TALDO) deficiency is a rare autosomal recessive disorder caused by variants in the *TALDO1* gene that commonly results in multisystem dysfunction. Herein, we reported compound heterozygous variants in a Chinese prenatal case with TALDO deficiency using whole-exome sequencing (WES) for trios and Sanger sequencing. The heterozygous variants were located on the *TALDO1* gene: NM_006755.2:c.574C > T(Chr11:g.763456C > T), a missense variant in exon 5 paternally inherited; NM_006755.2:c.462-2A > G(Chr11:g.763342A > G), a splicing aberration in intron 4 maternally inherited. The qualitative analysis of urinary polyols in neonatal urine indicated that xylitol + arabitol and ribitol in the proband's urine were significantly increased. These findings expand the variation spectrum of the *TALDO1* gene, provide solid evidence for the counseling of the family in regard to future pregnancies, strongly support the application of WES in prenatal diagnosis, and further prove that effective postpartum treatments could improve prognosis.

**Keywords: Transaldolase deficiency, pentose phosphate pathway, *TALDO1*, prenatal diagnosis, whole-exome sequencing (WES)**

OPEN ACCESS

***Edited by:***
*Natália Duarte Linhares, Genuity Science, Ireland*

***Reviewed by:***
*Vassili Valayannopoulos, Ultragenyx Pharmaceutical, United States*
*Mirjam Wamelink, Amsterdam University Medical Center, Netherlands*

***Correspondence:**
*Jin Han*
*hanjin1123@163.com*

***Specialty section:***
*This article was submitted to Genetics of Common and Rare Diseases, a section of the journal Frontiers in Genetics*

***Received:*** *02 August 2021*
***Accepted:*** *14 December 2021*
***Published:*** *04 February 2022*

***Citation:***
*Xue J, Han J, Zhao X, Zhen L, Mei S, Hu Z and Li X (2022) Prenatal Diagnosis of Fetus With Transaldolase Deficiency Identifies Compound Heterozygous Variants: A Case Report. Front. Genet. 12:752272. doi: 10.3389/fgene.2021.752272*

## INTRODUCTION

Transaldolase (TALDO) deficiency (OMIM 606003), a rare metabolic congenital defect of the pentose phosphate pathway (PPP), is caused by homozygous or compound heterozygous variants of the *TALDO1* gene (Wamelink et al., 2008) located on chromosome 11p15. Its main clinical manifestations usually appear in the neonatal period, while they are relatively rare in the antenatal period. The typical symptoms include coagulopathy, thrombocytopenia, liver dysfunction, hepatosplenomegaly, hepatic fibrosis, hemolytic anemia, generalized edema, dysmorphic features, and renal dysfunction that rarely occurs. Prolonged activated partial thromboplastin time (APTT), prothrombin time (PT) and thrombin time (TT), low cholesterol, high alkaline phosphatase (AKP), as well as elevated total bilirubin (TBIL), direct bilirubin (DBIL), total bile acid (TBA), and β2-microglobulin (β2-MG), can indicate liver and renal dysfunction in some reported cases (Valayannopoulos et al., 2006).

The PPP has two main functions: 1) It provides reduced nicotine adenine dinucleotide phosphate (the cofactor of redox reaction for organism biosynthesis), and 2) it offers ribose-5-phosphate to the nucleic acid. The PPP is divided into oxidative (nonreversible) and nonoxidative (reversible) enzymatic reactions/parts. Moreover, TALDO is the second enzyme of the nonoxidative part tightly linking the PPP and glycolysis pathway (Verhoeven et al., 2005).

To date, approximately 39 cases diagnosed with TALDO deficiency have been reported, but the incidence is unclear (Verhoeven et al., 2001; Eyaid et al., 2013; Rodan and Berry, 2017; Halabi et al., 2019; Lee-Barber et al., 2019; Williams et al., 2019; Lafci et al., 2021) (**Table 3**). Yet, the pathophysiology leading to TALDO deficiency remains unclear due to the low number of reported cases. TALDO deficiency can also have high variability in clinical manifestations and outcomes, even within the same family (Tylki-Szymanska et al., 2009; Leduc et al., 2014). Herein, we reported a novel compound heterozygous variant in a Chinese prenatal case with multiorgan dysfunction confirmed as TALDO deficiency by prenatal molecular diagnosis.

# MATERIALS AND METHODS

## Ethics Approval

After receiving written informed consent from both of the parents, WES (trio analysis of the proband, mother, and father) was carried out. Our study was approved by the Ethics Committee of Guangzhou Women and Children's Medical Center and Guangzhou Medical University, and it conformed with the ethical standards of experiments on human subjects.

## Case Presentation

A 33-year-old pregnant woman, gravida 2, para 1, was referred to our hospital at 34 weeks because of ultrasonic abnormalities. Fetal middle cerebral artery peak systolic velocity (MCA-PSV) kept increasing from 24 gestational weeks, reaching 93.97 cm/s [>1.5 MoM (Multiples of the Median)] at 33 gestational weeks. Additional anomalies included a slightly high echo of the right lobe of the liver, cardiomegaly with the cardiothoracic ratio of 0.61, a small amount of pericardial effusion, and placental thickness of 46 mm.

Similar manifestations, including cardiac enlargement, hepatosplenomegaly, placental thickness, elevated MCA-PSV, high umbilical artery resistance, and intrauterine growth restriction (IUGR), were observed during the first pregnancy (II:1). Following fetal distress, at 36 weeks, a baby boy was born by cesarean section weighing 1,860 g (<10th). His Apgar scores were normal (9′-10′-10′), while the neonatal peripheral blood test detected that hemoglobin (HGB) and platelet (PLT) were low. Repeated examinations of coagulation showed extended APTT, PT, and TT. Brain ultrasound suggested a head injury with subependymal hemorrhage. Therefore, II:1 received human immunoglobulin and blood transfusion to prevent infection and improve blood coagulation. The neonate did not recover and consequently died of disseminated intravascular coagulation (DIC), a low-birth-weight, and hypoproteinemia at 18 days. Clinical findings of the two affected fetuses (II:1 and II:2) are summarized in **Tables 1**, **2**.

To assess the risk of recurrence, cordocentesis was performed for genetic diagnosis, including karyotype analysis and chromosomal microarray analysis (CMA), to clarify the potential cause of the disease two times in another hospital, but the results were negative. Consequently, WES was performed on the proband and his healthy parents (**Figures 1A,B**) to search for potential variants. The detailed examinations during pregnancy are listed in **Table 1**.

## Metabolite Analyses

Urine xylitol + arabitol and ribitol were measured using gas chromatography-mass spectrometry (GC-MS). Urine sample preparation was based on urease pretreatment methods. Samples were standardized to 0.25 mg creatinine. Derivatization was performed with 100 µl bis-(trimethylsilyl) trifluoracetamide + 1% trimethylchlorosilane and was allowed to react at 60°C for 10 min. The metabolites were chromatographically analyzed as trimethylsilyl compounds.

## Whole-Exome Sequencing

Genomic DNA was randomly fragmented and purified using the magnetic particle method. WES was performed on an IIIumina HiSeq 2,500 sequencer (Illumina, San Diego, CA, United States) for a minimal of 10.14 Gb read-depth per case. Sequencing reads after quality control were aligned to the human reference genome by BWA (hg19). Nucleotide changes of aligned reads were reviewed using NextGENe software (Version 2.4.1.2) (SofGenetics, State College, PA, United States). Novel variants were filtered against the 1,000 Genomes database (http://www.1000genomes.org/), dbSNP database (http://www.ncbi.nlm.nih.gov/projects/SNP/snp_summary.cgi), and the Genome Aggregation database (gnomad.broadinstitute.org). Databases, including ClinVar (version: #372716), OMIM (version: #602063.0005), ClinGen (version: #CA5788214), and Human Gene Mutation database, were used. In addition, software (SIFT, Polyphen, MutationTaster, PROVEAN and REVEL) was used to predict the impact of missense variants. For the splicing variant, the *in silico* prediction tools were dbscSNV and MaxEntScan. Common variants (with high minor allele frequency in normal population; gnomAD) were eliminated. Finally, polymerase chain reaction (PCR) was performed to amplify the affected fragment of *TALDO1* gene using specific primers, and the purified PCR products were applied to Sanger sequencing to affirm the variant(s).

# RESULTS

The umbilical cord blood samples of the fetus (II:2) in 24 and 28 gestational weeks in another hospital showed fetal anemia, thrombocytopenia, coagulation dysfunction, and elevated liver

**TABLE 1 |** Ultrasound findings of II:2 at different gestational weeks.

| Test time(weeks) | NT/ NF[a] (mm) | MCA-PSV[b] (cm/s) | Reference interval of MCA-PSV (cm/s)[c] | Cardiothoracic ratio | Pericardial effusion (mm) | Placenta thickening (mm) | Right lobe length of liver (mm) | Reference interval of liver (mm)[d] |
|---|---|---|---|---|---|---|---|---|
| 12+ | 1.1 | – | – | – | – | 16 | – | – |
| 17+ | 4.9 | – | – | – | – | 22 | – | – |
| 22+ | – | – | – | <0.50 | – | 26 | – | – |
| 25+ | – | 44.0 | 23.6–40.8 | <0.50 | – | 32 | – | – |
| 27+ | – | 56.7 | 26.8–46.1 | <0.50 | – | 33 | – | – |
| 33+ | – | 93.0 | 36.0–46.3 | 0.57 | 2.8 | 46 | 57.7 | 40.6–52.3 |
| 37+ | – | 93.9 | 38.9–75.4 | 0.64 | 4.0 | 48 | 64.0 | 47.0–58.7 |

[a]*Note. NT, nuchal translucency (normal <3.0 mm); NF, neck fold (normal <6.0 mm).*
[b]*MCA-PSV, middle cerebral artery peak systolic velocity.*
[c]*Ebbing et al., 2007.*
[d]*Tongprasert et al., 2011.*

**TABLE 2 |** The laboratory results of II:2 as determined directly in 24+ and 28+ weeks through the umbilical cord blood tests and after birth between II:1 and II:2.

| Time test | II:2[a] | | | II:2 | II:1[b] | Reference interval |
|---|---|---|---|---|---|---|
| | 24 Gestational weeks | 28 Gestational weeks | Reference interval | Newborn | Newborn | |
| TT[c](s) | – | – | – | 23.6 | 34.8 | 14–21 |
| PT(s) | – | – | – | 30.7 | 31.7 | 11–15 |
| APTT(s) | – | – | – | 80.5 | 124.7 | 28–45 |
| AKP(U/L) | 289 | 359 | 15–121 | 354 | - | 118–390 |
| TBIL(μmol/L) | 23.1 | 39.3 | 1.7–20.0 | 83.3 | 157.52 | 2–17 |
| DBIL(μmol/L) | 2.76 | 3.89 | 0–6 | 10.0 | 96.3 | 0–7 |
| TBA(μmol/L) | – | – | – | 18.4 | 75.41 | 0.5–10.0 |
| β2-MG (mg/L) | 10.85 | 4.71 | 0.7–1.8 | – | – | – |
| HGB (g/L) | 82 | 104 | 110–150 | 95 | 104 | 135–195 |
| PLT (*$10^{9/L}$) | 123 | 137 | 100–300 | 79 | 48 | 140–440 |
| HbA (%) | 4.0 | 5.3 | 96.8%–97.8% | – | – | – |
| LDH (U/L) | 263 | 256 | 110–240 | 989 | 1,679 | 159–322 |
| AST (U/L) | 24 | 0–45 | 26 | 89 | 63 | 5–60 |
| ALB (g/L) | – | – | – | 23.1 | 16.5 | 40–55 |

[a]*Note. The second offspring.*
[b]*The first offspring.*
[c]*TT, thrombin time; PT, prothrombin time; APTT, activated partial thromboplastin time; AKP, alkaline phosphatase; TBIL, total bilirubin; DBIL, direct bilirubin; TBA, total bile acid; β2-MG, β2-microglobulin; HGB, hemoglobin; PLT, platelet; HbA, hemoglobin A; LDH, lactate dehydrogenase; AST, aspartate aminotransferase; ALB, albumin.*

enzymes [lactate dehydrogenase (LDH) and β2-MG] **(Table 2)**. Urine test for metabolic compounds using GC-MS showed an elevation of xylitol + arabitol at 170,388 mmol/mol creatinine (normal 0–1,151 mmol/mol creatinine) and ribitol at 193,301 mmol/mol creatinine (normal 0–886 mmol/mol creatinine) (**Figure 2**).

WES revealed compound heterozygosity of variants in the *TALDO1* gene in the proband: maternally inherited likely splicing aberration NM_006755.2(*TALDO1*):c.462-2A > G(Chr11:g.763342A > G), and paternally inherited missense variant NM_006755.2(*TALDO1*):c.574C > T(Chr11:g.763456C > T). Both parents were heterozygous carriers and phenotypically normal. Sanger sequencing of the patient and his family members further validated these results (**Figures 1A,B**). According to the ACMG standards, both variants were defined as likely pathogenic (c.462-2A > G: PVS1 + PM2, c.574C > T: PM2 + PM3 + PM5 + PP3). The two variants had a very low carrying rate in some databases: c.462-2A > G is not recorded in gnomAD, dbSNP, or 1,000 Genomes databases; the minor allele frequency of c.574C > T is .00001591 (4 heterozygotes) in gnomAD, and 0.000008243 (1 heterozygote) in ExAC. c.574C > T is a missense variant located in exon 5, which is also described on dbSNP (rs751425603) and reported as pathogenicity in ClinVar (variation ID: 381,759). c.462-2A > G is a splicing variant located in intron 4 and may lead to abnormal mRNA splicing that affects protein expression. The splicing variant c.462-2A > G in MaxEntScan score was from 10.76 in wild type to 2.824 in mutant type. This variant is predicted to cause a loss of function of the protein.

## OUTCOME

The couple decided to continue the pregnancy after genetic counseling. A baby girl was born at 38 weeks, with a

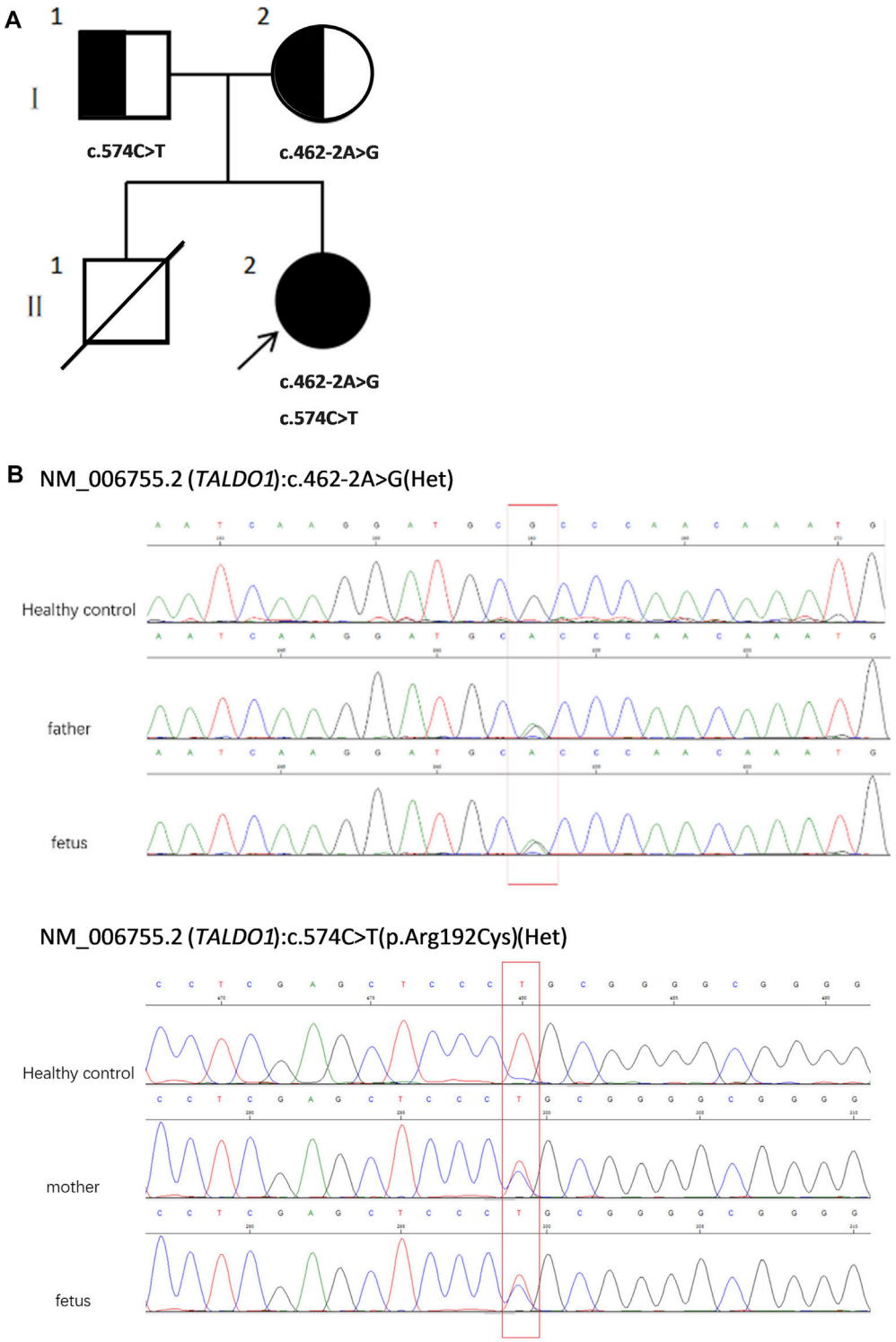



**FIGURE 1 | (A)** Pedigree of the family. **(B)** Sequencing of *TALDO1* gene (reference cDNA sequence, NM_006755.2) revealed two heterozygous variations, resulting in A to G splicing at nucleotide position 462–2 (c.462-2A > G) and C to T substitution at nucleotide position 574 [c.574C > T(p.Arg192Cyrs)]. Het, heterozygous.

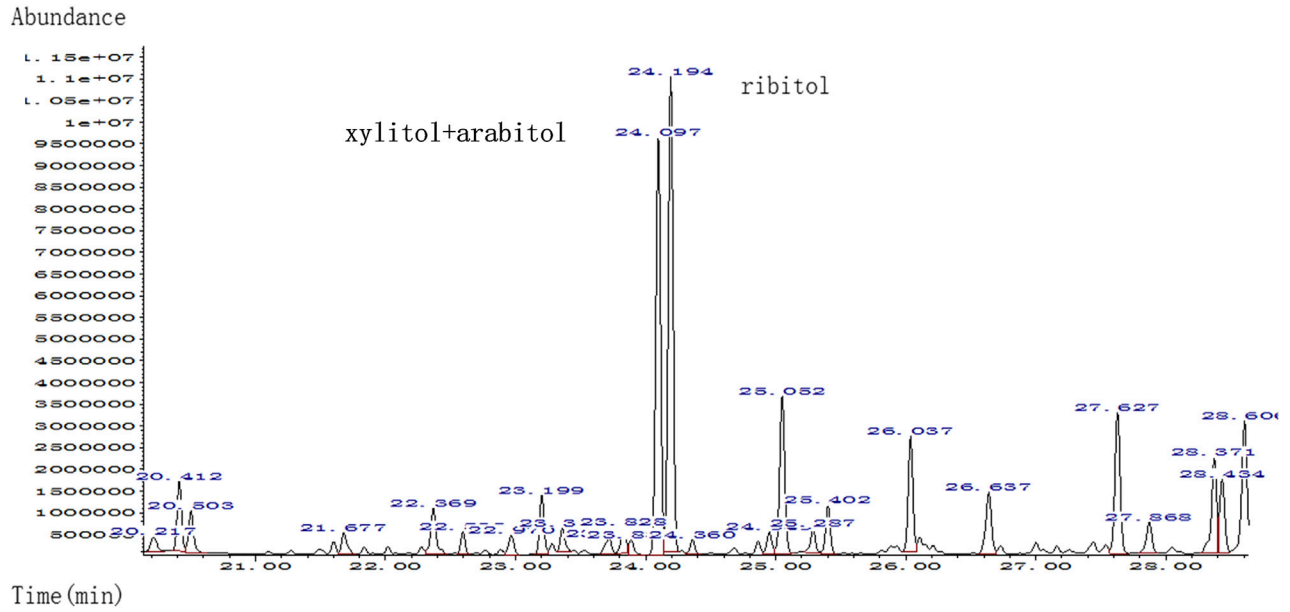



**FIGURE 2 |** Quality analysis of polyols in urine sample of patient by urine gas chromatography-mass spectrometry (GC-MS); markedly elevated xylitol + arabitol and ribitol were detected.

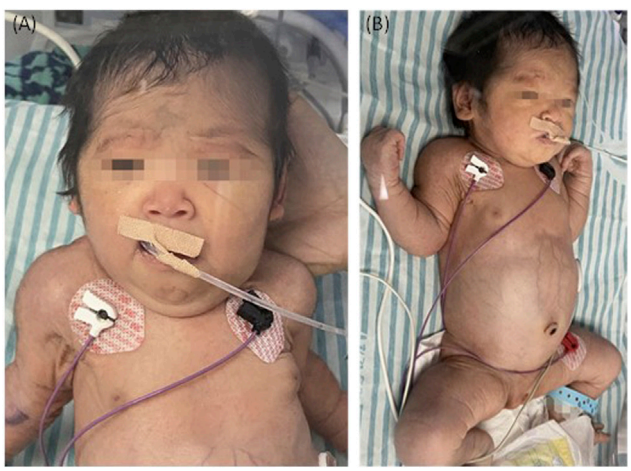



**FIGURE 3 |** The TALDO-deficient dysmorphic feature-hirsutism (forehead), low hair implantation, mild pallor, and cuties laxa with visible vascular network of the patient. The abdomen was also grossly distended with dilated visible veins [**(A)** frontal view, **(B)** side view].

weight of 2,760 g. Apgar scores were normal (8′-8′-8′) after delivery. At birth, the baby had dysmorphic features (hirsutism, low hair implantation), mild pallor, and cutis laxa. She also presented low skin temperature, quick breath with groaning, thick breath sounds in both lungs with moist rales, and abdominal distention with the visible vascular network. The baby was hospitalized at the Neonatal Department in our hospital for 9 days (**Figures 3A,B**). She had hepatosplenomegaly and developed jaundice. A peripheral blood test showed HGB of 95 g/L (normal 110–150 g/L) with fragmented red cells on film, thrombocytopenia, and mild neutropenia, consistent with those *in utero*. Serum TBIL was 83.3 μmol/L (normal 2–17 μmol/L) with DBIL of 10.0 μmol/L (normal 0–7 μmol/L). LDH was also increased to 989 U/L (normal 159–322 U/L) together with marginally elevated transaminases, bile acids and alkaline phosphatase (ALP). Albumin was 23.1 g/L (normal 40–55 g/L) and PT was 31.7 s (normal 11–15 s). The infant received continuous ventilation for 9 days. Fresh frozen plasma and fibrinogen infusion were given to improve thrombocytopenia and coagulation. Blood glucose level was stable and was closely monitored. GC-MS indicated elevated urinary xylitol + arabitol and ribitol levels.

After her condition gradually improved, the patient was discharged from the hospital and was regularly followed up. At the age of 9 months, HGB was still slightly decreased (100 g/L), while red blood cell and PLT were both increased to $4.7 \times 10^{12/L}$ (normal $3.5–5.0 \times 10^{12/L}$) and $517 \times 10^{9/L}$ (normal $100–300 \times 10^{9/L}$), respectively. DBIL, TBIL, TBA, LDH, and AKP levels were normal, whereas aspartate aminotransferase (AST) mildly elevated to 84 U/L. The dysmorphic features and cutis laxa were not observed. Thus far, the child has shown normal physical and cognitive development.

## DISCUSSION

TALDO deficiency is a rare autosomal recessive error of the PPP caused by a variant in the *TALDO1* gene (Williams et al., 2019). TALDO1 gene encodes TALDO implicated as a major modulator between the PPP and glycolysis in a reversible reaction. TALDO catalyzes the conversion of glyceraldehyde-3-phosphate and sedoheptulose-7-phosphate into fructose-6-phosphate and erythrose-4-phosphate, which are also considered targets for the treatments of this condition. In addition, its absence can result in the accumulation of intermediate products (e.g., sedoheptulose, erythritol, and ribitol) and eventually cause lesions of the patent. TALDO deficiency has been associated with a range of phenotypes, including intrauterine lethality together with fetal multimalformation syndrome and hydrops fetalis. The most common clinical manifestations in neonates are cirrhosis, liver failure, hepatosplenomegaly, anemia, thrombocytopenia, dysmorphia, congenital heart defects, and tubulopathy (Verhoeven et al., 2001; Verhoeven et al.,2005; Valayannopoulos et al., 2006; Wamelink et al., 2008; Tylki-Szymanska et al., 2009; Balasubramaniam et al., 2011; Eyaid et al., 2013). Yet, prenatal diagnosis is very challenging. Abnormal findings in the fetus are rare. Some of the common manifestations in the antenatal period are IUGR (Verhoeven et al., 2001; Valayannopoulos et al., 2006; Wamelink et al., 2008), oligohydramnios, fetal splenomegaly, fetal distress (Wamelink et al., 2008), and hyperechogenic bowel (Banne et al., 2015). Also, TALDO deficiency can be easily misdiagnosed with gestational alloimmune liver disease (GALD). GALD is the result of maternal alloimmune injury, which includes neonatal liver failure (coagulation disorders, ascites, and hypoalbuminemia), intrahepatic, and extrahepatic iron accumulation (hemosiderosis).

In this study, a family that had experienced neonatal death following IUGR, hepatosplenomegaly, anemia with thrombopenia, and abnormal coagulation tests in a previous pregnancy (II:1) and recurrent fetal anemia, hepatosplenomegaly in the second pregnancy (II:2), was recruited for WES. A prenatal diagnosis of the fetus confirmed heterozygous variants in the *TALDO1* gene in II:2. Yet, prenatal findings were different between II:1 and II:2. Fetal MCA-PSV increased from 24 gestational weeks, which reflected fetal anemia *in utero*. Additional ultrasound anomalies identified at 33 gestational weeks included slightly high echo of the right lobe of the liver, cardiomegaly with increased cardiothoracic ratio, a small amount of pericardial effusion and placental thickness, all of which suggested a progressive development of fetal anemia. After birth, the postpartum symptoms were clearer and more obvious, including dysmorphic features, liver dysfunction and hemolytic anemia. This case is consistent with the range of phenotypes most commonly observed; however, fetal anemia, liver dysfunction, and coagulopathy are the main manifestations.

The accumulation of sugars and polyols [e.g., sedoheptulose-7-phosphate, ribose-5-phosphate, ribulose-5-phosphate, xylulose-5-phosphate, and C5-polyols (i.e., D-ribitol and D-arabitol)] are believed to be the cause of liver involvement in TALDO deficiency. Higher concentrations of the polyols xylitol + arabitol and ribitol in the urine of the proband could be relevant of the phenotypes in II.2, but could also be related to the younger age, since the polyol concentrations were higher in the neonatal period in other patients and accumulated less when they were older (Wamelink et al., 2008). Although from the same family, patient II:1 had IUGR, anemia, hepatosplenomegaly, DIC, a low-birth-weight, and secondary hemorrhage (subependymal hemorrhage), yet, even considering that molecular analysis was not performed for patient II:1, it was likely that these phenotypes were associated with TALDO deficiency.

To the best of our knowledge, this case is the first prenatal diagnosis of TALDO deficiency in a Chinese population (Verhoeven et al., 2001; Eyaid et al., 2013; Rodan and Berry, 2017; Lee-Barber et al., 2019; Williams et al., 2019; Halabi et al., 2019; Lafci et al., 2021). Both variants of this case were defined as likely pathogenic. One of the variants [c.574C > T p.(Arg192Cys)], reported as pathogenicity in ClinVar (Variation ID: 381,759), was previously reported in an Arab patient, suggesting a founder effect in Arab populations (Wamelink et al., 2008). The other is a novel splicing variant (c.462-2A > G), which is predicted to affect splicing while not exon skipping. The in-silico tools are dbscSNV and MaxEntScan. They all predict altering *TALDO1* exon splicing. To date, there have been 13 variants reported to cause this condition worldwide (**Table 3**). Individuals with the same variant show different clinical manifestations.

The prenatal diagnosis of TALDO deficiency remains a challenge and is usually confirmed by gene analysis. Thus far, there is still no effective treatment for TALDO deficiency. Yet, early and accurate prenatal diagnosis can lead to a better outcome and can provide better aid for prenatal management, including fetal surveillance strategy and appropriate postpartum treatment, as was the case in the present study. In particular, a higher frequency of fetal surveillance with targeted ultrasound can help identify early signs of clinical manifestations (e.g., elevated MCA-PSV, cardiomegaly and placental thickness), which are important prognostic indicators. Most important of all, it is inseparable from the joint efforts of multi-disciplinary team. Currently, there is only one gene known to cause TALDO deficiency. Further studies are warranted to comprehensively characterize the genetic contributions.

In conclusion, our data suggests that TALDO deficiency is a pleiotropic disorder that should be considered when investigating a prenatal case with unexplained hepatosplenomegaly or fetal anemia. Although no specific treatment is currently available, targeted molecular analysis of the *TALDO1* gene in amniotic fluid or chorionic villi

**TABLE 3 |** Summary of clinical manifestations in the current patients with TALDO deficiency.

| Case | Variant | Gender | Ethnicity | Consanguinity | Pregnancy | Dysmorphism | Liver dysfunction | Hepatosplenomegaly | Anemia | Thrombocytopenia | Impaired coagulation | Cardiac abnormalities | Neonatal edema | Renal | Respiratory | Developmental delay | Abnormal genitalia | Clinical course |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1[a] | NM_006755.2: c.512_514del | F | Turkey | + | IUGR | - | + | + | + | + | + | Aortic coarctation | - | - | - | + | + | Hepatosplenomegaly, telangiectasias of her skin, enlarged clitoris |
| 2[b]* | NM_006755.2: c.575G > A | F | Turkey | + | HELPP syndrome | + | + | + | + | + | + | Cardiomyopathy | + | Glomerular proteinuria | + | - | - | * |
| 3[c]* | NM_006755.2: c.512_514del | F | Turkey | + | *n* | + | + | + | + | + | + | large venous duct ASD, MVP | - | Nephrocalcinosis | - | - | - | * |
| 4[c]* (fetus) | NM_006755.2: c.512_514del | M | Turkey | + | *n* | + | + | - (Splenic fibrosis) | + | - | - | Cardiomegaly | + | - | - | - | - | * |
| 5[c] | NM_006755.2: c.512_514del | M | Turkey | + | *n* | +→*n* | + | + | + | + | + | PFO | - | Chronic renal failure hypoplastic kidney | - | - | - | n/- |
| 6[c] | NM_006755.2: c.512_514del | M | Turkey | + | Oligohydramnion splenomegaly fetal distress | +→n | + | + | + | + | + | Cardiomegaly PFO | - | Transient renal failure | - | - | - | n |
| 7[d] | NM_006755.2: c.574C > T | M | Arab | + | IUGR | + | + | + | + | + | + | Small patent ductus | - | Tubulopathy | - | Mild delay | - | Speech delay (deaf) |
| 8[e] | NM_006755.2: c.575G > A | M | Pakistani | + | *n* | +→n | + | + | - | + | + | - | - | - | - | Mild delay | - | Speech delay |
| 9[f] | NM_006755.2: c.575G > A | M | Poland | + | *n* | - | + | + | + | + | + | - | - | - | - | - | - | Hepatosplenomegaly |
| 10[f] | NM_006755.2: c.575G > A | M | Poland | + | *n* | - | + | + | + | + | + | - | - | - | - | - | - | Unilateral cryptorchidism, hepatosplenomegaly |
| 11[g]* | NM_006755.1: c.895_897del; NM_006755.1: c.931G > A | M | China | - | *n* | + | - | + | + | + | + | + | + | - | + | - | - | *n* |
| 12[h-1] | NM_006755.2: c.793del | F | Saudi Arabia | + | A dilated left ventricle | + | + | + | + | + | - | PFO,PDA | - | - | - | + | - | *n* |
| 13[h-1]* | NM_006755.2: c.793del | M | Saudi Arabia | + | Polyhydramnios | + | + | + | + | + | - | PFO, ASD | - | - | - | - | - | * |
| 14[h-1] | NM_006755.2: c.793del | M | Saudi Arabia | + | n | + | + | + | + | + | - | + | - | - | - | + | - | *n* |
| 15[h-2] | NM_006755.2: c.793del | M | Saudi Arabia | + | — | + | + | + | + | + | - | PFO, ASD | - | - | + | - | - | *n* |
| 16[h-2] | NM_006755.2: c.793del | F | Saudi Arabia | + | — | + | + | + | + | + | - | PFO | - | - | - | - | - | *n* |
| 17[h-2] | NM_006755.2: c.793del | M | Saudi Arabia | + | Mild pericardial effusion, cardiomegaly, and echogenic bowel | + | + | + | + | + | - | PDA, VSDs, and ASD | - | - | - | - | - | *n* |
| 18[h-3] | NM_006755.2: c.793del | M | Saudi Arabia | + | *n* | + | + | + | + | + | - | ASD | - | - | + | - | - | *n* |
| 19[h-3] | NM_006755.2: c.793del | | Saudi Arabia | + | IUGR, oligohydramnios, situs inversus totalis, thick nuchal skin, slightly enlarged right heart, and hepatosplenomegaly | + | + | + | + | + | - | PDA | + | - | - | - | - | *n* |
| 20[h-3] | NM_006755.2: c.793del | M | Saudi Arabia | + | *n* | + | + | + | + | + | - | ASD | - | - | - | - | - | *n* |
| 21[h-4] | NM_006755.2: c.793del | F | Saudi Arabia | + | IUGR | + | + | + | + | + | - | + | - | - | - | - | + | *n* |
| 22[h-5] | NM_006755.2: c.793del | F | Saudi Arabia | + | *n* | - | + | + | + | + | + | + | - | + | - | - | - | *n* |
| 23[h-6] | NM_006755.2: c.793del | F | Saudi Arabia | + | *n* | + | - | + | + | + | - | + | - | - | - | - | - | *n* |
| 24[i] | — | M | Saudi Arabia | + | *n* | + | + | + | + | + | - | PDA, VSDs, and ASD | - | - | + | - | - | *n* |
| 25[j] | c.462-174_981 + 53del | M | Poland | - | IUGR, ascites, and oligoamnios | + | + | + | + | + | - | - | - | + | - | - | - | *n* |
| 26[j] | | M | Poland | – | | + | + | + | + | + | - | - | - | Renal calculus | - | - | + | *n* |

(Continued on following page)

# Table of Contents

can be valuable in helping those suffering families to make informed reproductive choices.

## DATA AVAILABILITY STATEMENT

TThe original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by The ethics committee of Guangzhou Women and Children's Medical Center and Guangzhou Medical University. 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) for the publication of any potentially identifiable images or data included in this article.

## AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

## FUNDING

This study was supported (or partially supported) by the Science and Technology Department of Guangdong Province (2016A020218003), Guangzhou Science Technology and Innovation Commission (201607010341), and Guangzhou Women and Children's Medical Center (GWCMC2020-6-007).

## ACKNOWLEDGMENTS

The authors kindly acknowledge the participation of the family members and the staff members in the Prenatal Diagnosis Center of Guangzhou Women and Children's Medical Center and Shenzhen People's Hospital, Guangdong Province, China. The authors would also like to thank the Genome Aggregation database (gnomAD) for providing genome variant data.

## REFERENCES

Al-Shamsi, A. M., Ben-Salem, S., Hertecant, J., and Al-Jasmi, F. (2015). Transaldolase Deficiency Caused by the Homozygous p.R192C Mutation of the TALDO1 Gene in Four Emirati Patients With Considerable Phenotypic Variability. *Eur. J. Pediatr.* 174 (5), 661–668. doi:10.1007/s00431-014-2449-5

Balasubramaniam, S., Wamelink, M. M., Ngu, L.-H., Talib, A., Salomons, G. S., Jakobs, C., et al. (2011). Novel Heterozygous Mutations in TALDO1 Gene Causing Transaldolase Deficiency and Early Infantile Liver Failure. *J. Pediatr. Gastroenterol. Nutr.* 52, 113–116. doi:10.1097/MPG.0b013e3181f50388

Banne, E., Meiner, V., Shaag, A., Katz-Brull, R., Gamliel, A., and Korman, S. (2015). Transaldolase Deficiency: A New Case Expands the Phenotypic Spectrum. *JIMD Rep.* 26, 31–36. doi:10.1007/8904_2015_474

Ebbing, C., Rasmussen, S., and Kiserud, T. (2007). Middle Cerebral Artery Blood Flow Velocities and Pulsatility Index and the Cerebroplacental Pulsatility Ratio: Longitudinal Reference Ranges and Terms for Serial Measurements. *Ultrasound. Obstet. Gynecol.* 30 (3), 287–296. doi:10.1002/uog.4088

Eyaid, W., Al Harbi, T., Anazi, S., Wamelink, M. M. C., Jakobs, C., Al Salammah, M., et al. (2013). Transaldolase Deficiency: Report of 12 New Cases and Further Delineation of the Phenotype. *J. Inherit. Metab. Dis.* 36, 997–1004. doi:10.1007/s10545-012-9577-8

Fung, C. W., Siu, S., and Mak, C. (2007). A Rare Cause of Hepatosplenomegaly-Transaldolase Deficiency. *J. Inherit. Metab. Dis.* 30 (Supplement 1), 62.

Halabi, M., Alqoaer, K., and Asaad, Z. (2019). Severe Infantile Transaldolase Deficiency. *Ajpr*, 1–5. doi:10.9734/ajpr/2019/v2i430112

Jassim, N., Alghaihab, M., Saleh, S. A., Alfadhel, M., Wamelink, M. M., and Eyaid, W. (2014). Pulmonary Manifestations in a Patient With Transaldolase Deficiency. *JIMD Rep.* 12, 47–50. doi:10.1007/8904_2013_243

Lafci, N. G., Colak, F. K., Sahin, G., Sakar, M., Çetinkaya, S., and Savas-Erdeve, S. (2021). Hypergonadotrophic Hypogonadism in a Patient with Transaldolase Deficiency: Novel Mutation in the Pentose Phosphate Pathway. *Hormones* 20, 581–585. doi:10.1007/s42000-020-00252-4

Leduc, C. A., Crouch, E. E., Wilson, A., Lefkowitch, J., Wamelink, M. M. C., Jakobs, C., et al. (2013). Novel Association of Early Onset Hepatocellular Carcinoma with Transaldolase Deficiency. *JIMD Rep.* 12, 121–127. doi:10.1007/8904_2013_254

Lee-Barber, J., English, T. E., Britton, J. F., Sobreira, N., Goldstein, J., Valle, D., et al. (2019). Apparent Acetaminophen Toxicity in a Patient with Transaldolase Deficiency. *JIMD Rep.* 44, 9–15. doi:10.1007/8904_2018_116

Rodan, L. H., and Berry, G. T. (2016). N-Acetylcysteine Therapy in an Infant with Transaldolase Deficiency Is Well Tolerated and Associated with Normalization of Alpha Fetoprotein Levels. *JIMD Rep.* 31, 73–77. doi:10.1007/8904_2016_555

Tongprasert, F., Srisupundit, K., Luewan, S., and Tongsong, T. (2011). Normal Length of the Fetal Liver From 14 to 40 Weeks of Gestational Age. *J. Clin. Ultrasound.* 39 (2), 74–77. doi:10.1002/jcu.20756

Tylki-Szymańska, A., Stradomska, T. J., Wamelink, M. M. C., Salomons, G. S., Taybert, J., Pawłowska, J., et al. (2009). Transaldolase Deficiency in Two New Patients with a Relative Mild Phenotype. *Mol. Genet. Metab.* 97, 15–17. doi:10.1016/j.ymgme.2009.01.016

Valayannopoulos, V., Verhoeven, N. M., Mention, K., Salomons, G. S., Sommelet, D., Gonzales, M., et al. (2006). Transaldolase Deficiency: a New Cause of Hydrops Fetalis and Neonatal Multi-Organ Disease. *J. Pediatr.* 149, 713–717. doi:10.1016/j.jpeds.2006.08.016

Verhoeven, N. M., Huck, J. H. J., Roos, B., Struys, E. A., Salomons, G. S., Douwes, A. C., et al. (2001). Transaldolase Deficiency: Liver Cirrhosis Associated with a New Inborn Error in the Pentose Phosphate Pathway. *Am. J. Hum. Genet.* 68, 1086–1092. doi:10.1086/320108

Verhoeven, N. M., Wallot, M., Huck, J. H. J., Dirsch, O., Ballauf, A., Neudorf, U., et al. (2005). A Newborn with Severe Liver Failure, Cardiomyopathy and Transaldolase Deficiency. *J. Inherit. Metab. Dis.* 28, 169–179. doi:10.1007/s10545-005-5261-6

Wamelink, M. M., Struys, E. A., Salomons, G. S., Fowler, D., Jakobs, C., and Clayton, P. T. (2008). Transaldolase Deficiency in a Two-Year-Old Boy with Cirrhosis. *Mol. Genet. Metab.* 94, 255–258. doi:10.1016/j.ymgme.2008.01.011

Williams, M., Valayannopoulos, V., Altassan, R., Chung, W. K., Heijboer, A. C., Keng, W. T., et al. (2019). Clinical, Biochemical, and Molecular Overview of Transaldolase Deficiency and Evaluation of the Endocrine Function:

Update of 34 Patients. *Jrnl Inher Metab. Disea.* 42, 147–158. doi:10.1002/jimd.12036

**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.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed nor endorsed by the publisher.

*Copyright © 2022 Xue, Han, Zhao, Zhen, Mei, Hu and Li. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*

frontiers in Genetics

ORIGINAL RESEARCH
published: 25 February 2022
doi: 10.3389/fgene.2022.796759

# Mild Phenotype of Arthrogryposis, Renal Dysfunction, and Cholestasis Syndrome 1 Caused by a Novel *VPS33B* Variant

*Natália Duarte Linhares*[1,2], *Eleonora Druve Tavares Fagundes*[3,4], *Alexandre Rodrigues Ferreira*[3,4], *Thaís Costa Nascentes Queiroz*[4], *Luiz Roberto da Silva*[5] *and Sergio D. J. Pena*[1,2,6]*

[1]Laboratório de Genômica Clínica, Faculdade de Medicina, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil, [2]Departamento de Bioquímica e Imunologia, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil, [3]Departamento de Pediatria, Faculdade de Medicina, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil, [4]Hospital das Clínicas, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil, [5]Hospital de Clínicas—EBSERH, Universidade Federal de Uberlândia, Uberlândia, Brazil, [6]Laboratório Gene—Núcleo de Genética Médica, Belo Horizonte, Brazil

The arthrogryposis, renal dysfunction, and cholestasis syndrome (ARCS) is an autosomal recessive multisystem disease caused by variants in *VPS33B* or *VIPAS39*. The classical presentation includes congenital joint contractures, renal tubular dysfunction, cholestasis, and early death. Additional features include ichthyosis, central nervous system malformations, platelet dysfunction, and severe failure to thrive. We studied three patients with cholestasis, increased aminotransferases, normal gamma-glutamyl transferase, and developmental and language delay. Whole exome sequencing analysis identified *VPS33B* variants in all patients: patients 1 and 2 presented a novel homozygous variant at position c.1148T>A. p.(Ile383Asn), and patient 3 was compound heterozygous for the same c.1148T>A. variant, in addition to the c.940-2A>G. variant. ARCS is compatible with the symptomatology presented by the studied patients. However, most patients that have been described in the literature with ARCS had severe failure to thrive and died in the first 6 months of life. The three patients studied here have a mild ARCS phenotype with prolonged survival. Consequently, we believe that the molecular analysis of the *VPS33B* and *VIPAS39* should be considered in patients with normal gamma-glutamyl transferase cholestasis.

**Keywords: whole exome sequencing, VPS33B gene, cholestasis, arthrogryposis, renal dysfunction**

OPEN ACCESS

***Edited by:***
*Babak Behnam, National Sanitation Foundation International, United States*

***Reviewed by:***
*Bixia Zheng, Nanjing Children's Hospital, China*
*Fred Pluthero, Hospital for Sick Children, Canada*

***Correspondence:**
*Sergio D. J. Pena*
*spena@dcc.ufmg.br*

***Specialty section:***
*This article was submitted to Genetics of Common and Rare Diseases, a section of the journal Frontiers in Genetics*

***Received:*** *17 October 2021*
***Accepted:*** *21 January 2022*
***Published:*** *25 February 2022*

***Citation:***
*Linhares ND, Fagundes EDT, Ferreira AR, Queiroz TCN, da Silva LR and Pena SDJ (2022) Mild Phenotype of Arthrogryposis, Renal Dysfunction, and Cholestasis Syndrome 1 Caused by a Novel VPS33B Variant. Front. Genet. 13:796759. doi: 10.3389/fgene.2022.796759*

## INTRODUCTION

Arthrogryposis, renal dysfunction, and cholestasis syndrome (ARCS) is a rare autosomal recessive multisystem disorder that has been named because of its three cardinal features (Nezelof et al., 1979; Horslen et al., 1994). There are two forms of the disease, with similar clinical symptoms: ARCS type 1 (ARCS1, OMIM 208085) is caused by pathogenic variants in the vacuolar protein sorting 33 homolog B (*VPS33B* gene; OMIM 608552), whereas ARCS type 2 (ARCS2, OMIM 613404) is caused by pathogenic variants in the VPS33B-interacting protein apical-basolateral polarity regulator spe-39 homolog (*VIPAS39* gene, also known as *VIPAR*, OMIM 613401). Germline variants in the *VPS33B* gene have been found in approximately 75% of patients with ARCS (Gissen et al., 2006).

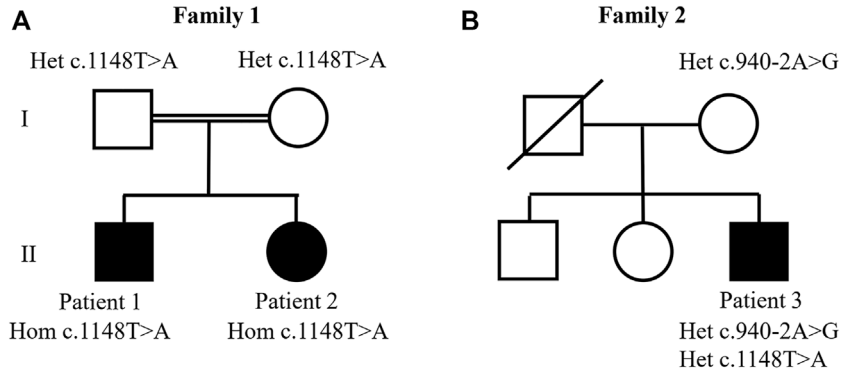



**FIGURE 1 |** Studied families' pedigrees. **(A)** Family 1: whole exome sequencing analysis identified the homozygous c.1148T>A. p.(Ile383Asn) variant in patients 1 and 2. **(B)** Family 2: compound heterozygous variants were identified in patient 3 at positions c.1148T>A. p.(Ile383Asn) and c. 940-2A>G.

Consistent with the widespread organ dysfunction in ARCS, VPS33B has a role in the regulation of intracellular protein trafficking, particularly with abnormal organelle biogenesis on the liver and on the kidney that may ultimately result in cholestasis and tubular disfunction (Gissen et al., 2004). VPS33B interacts with soluble N-ethylmaleimide–sensitive factor attachment protein receptors (SNAREs), which are involved in synaptic vesicle fusion, vesicular exocytosis, and neurosecretion (Lobingier and Merz, 2012; Han et al., 2017). In addition, mouse knockout studies have shown that *VPS33B* and *VIPAS39* are essential for epidermal lamellar body biogenesis and function (Rogerson and Gissen, 2018).

Clinical presentation, together with *VPS33B* and *VIPAS39* sequencing analysis, has been considered the recommended diagnostic procedure (Gissen et al., 2006; Cullinane et al., 2009; Zhou and Zhang, 2014). Organ biopsy, which was used in the past in combination with clinical diagnosis, has largely been replaced by molecular analysis, because more than 50% of patients are vulnerable to coagulation defects and kidney and/or liver biopsies may result in risk of fatal hemorrhage (Gissen et al., 2006; Zhou and Zhang, 2014). Nine of 11 patients that were reported with bleeding episodes had normal platelet morphology and count (Gissen et al., 2006). Clinical diagnosis of ARCS consists on identifying the triad conditions of arthrogryposis, renal tubular acidosis, and neonatal cholestatic jaundice with normal gamma-glutamyl transferase (GGT) activity (Gissen et al., 2006). No specific treatment currently exists for this syndrome. Rather, supportive care should be administered to patients with the aim of improving the quality of life (Zhou and Zhang, 2014). As additional features have been described, it has become evident that the phenotype is variable. For instance, renal disease can range from renal tubular acidosis to Fanconi syndrome or nephrogenic diabetes insipidus (Del Brio Castillo et al., 2019).

Comprehensive reviews have analyzed the clinical phenotype of more than 62 patients with ARCS1 and showed that the three cardinal features are sometimes accompanied by other phenotypic features, including ichthyosis, mild dysmorphic signs, platelet anomalies, agenesis of the corpus callosum, hypotonia, structural cardiac defects, deafness, recurrent infection, and severe failure to thrive (Abu-Sa'da et al., 2005; Gissen et al., 2006; Zhou and Zhang, 2014). Most patients have failed to survive beyond the first year of life because of recurrent infections, acidosis, or severe hemorrhage (Gissen et al., 2006; Zhou and Zhang, 2014). However, recently, some reported patients with milder phenotypes survived infancy, including cases with isolated liver disease (Agawu et al., 2019; Qiu et al., 2019).

Here, we report three patients with variants in *VPS33B* identified by whole exome sequencing (WES). They presented a mild ARCS phenotype with prolonged survival.

## CASE REPORT

### Family 1—Patients 1 and 2

Patients 1 and 2 were siblings born from a consanguineous healthy couple (their parent's grandparents were siblings) (**Figure 1**). Patient 1 is a boy who was born at term with weight of 2,850 g (10th centile). Jaundice was noticed on first month and resolved spontaneously without additional inquiry. He started to have intermittent pruritus at 8 months of age. Neurodevelopmental delay was noticed at the second year of life. He acquired independent walking at 1 year and 6 months. He was evaluated with a clinical report of syndromic face, mainly characterized by the lack of hair on the eyebrows, low-set ears, discrete ptosis, discrete camptodactyly of fingers, tissue excess in the hands, hyperreflexia in the lower limbs, bilateral short hallux, gait with *equinovarus* on the left, and dry and scaling skin. Brain magnetic resonance imaging showed hypoplasia of the corpus callosum and dysgenesis in a small area of the left cerebellar hemisphere, suggesting an abnormality of neuronal migration. There was no growth delay. He was first seen at the Clínica de Hepatologia Pediátrica of Hospital das Clínicas of the Universidade Federal de Minas Gerais (UFMG) at age 9 due to severe and difficult control pruritus, which was persistent despite topic skin care and use of ursodeoxycholic acid and rifampicin. He had mild elevated aminotransferases with normal GGT. Bilirubin, albumin, prothrombin time, and partial thromboplastin time were normal. Other causes of liver disease, such as Wilson's disease, hepatitis B and C,

α1-antitrypsin deficiency, and autoimmune hepatitis, were excluded by laboratory tests. Liver biopsy showed preserved lobular architecture and mild portal mononuclear inflammatory infiltrate without fibrosis and cholestasis. He had normal platelet number and function without bleeding episodes and no renal dysfunction. He was studied at the Laboratório de Genômica Clínica at age 11. He was also referred to the Departamento de Neurologia Pediátrica, where he was followed due to neurodevelopmental delay, right hemiplegia, cognitive impairment, and behavioral abnormalities. Currently, at age 17, he maintains the cognitive and language delay.

Patient 2 is a girl who was born at term by normal delivery with weight 3,600 g (75th centile). However, she had low weight gain after 6 months of life. She presented jaundice, with spontaneous resolution and pruritus, although receiving ursodeoxycholic acid and rifampicin. She had mild elevated aminotransferases with normal GGT, hepatomegaly and decreased body weight and height, and normal liver function. She also presented facial dysmorphisms, dry and scaling skin, and neurodevelopmental delay with slow language development. She had no bleeding episodes and no renal dysfunction. Currently, at age 10, she presents mild pruritus, short stature, and normal liver and renal function.

## Family 2—Patient 3

Patient 3 is a boy who was born from non-consanguineous parents (**Figure 1**) and had healthy older brother and sister. He was born at term by cesarean delivery without complications with weight 2.910 g (10th centile). Intermittent jaundice started with 15 days of life. His mother noticed pruritus in the first months of life. He was followed at the Clínica de Hepatologia Pediátrica of Hospital das Clínicas of UFMG from the age of 4 years old. He presented cholestasis with intense pruritus, elevated aminotransferases, normal GGT, dry and scaling skin, developmental and language delay, sensorineural hearing loss, and syndromic facies characterized by low-set ears, broad forehead, brachycephaly, and short nasolabial filter, but he had no joint contractures. Other causes of cholestasis were excluded. Albumin, prothrombin time, and partial thromboplastin time were normal. Blood tests showed large platelet, however in normal number. He had no hemorrhage episodes. There was no renal dysfunction. Currently, at age 11, he presents mild pruritus and jaundice (total bilirubin, 5.6 mg/dl; direct bilirubin, 4.2 mg/dl) but normal liver function (normal albumin and prothrombin and partial thromboplastin time). He has normal growth but keeps developmental and language delay.

# METHODS

## Samples, DNA Isolation, and WES Analysis

The Research Ethics Committee of the Hospital das Clínicas of UFMG approved the study protocol. Informed consent was obtained according to current ethical and legal guidelines. The study was conducted in accordance with the Declaration of Helsinki.

Genomic DNA was isolated from whole peripheral blood using a modified salting out procedure (Miller et al., 1988). It was not possible to collect patient 3's father sample, because he was deceased.

WES was performed using patient 2' sample by the Centre for Applied Genomics, Hospital for Sick Children, Toronto, Canada, using the SureSelect Human All Exon kit V5 (Agilent Technologies, Santa Clara, CA, USA), which targeted 21,522 genes and 357,999 exons, with a total size of 50 Mb. Enriched genomic DNA was sequenced on a HiSeq 2,500 Sequencer (Illumina, San Diego, CA, USA). The average coverage was 119,71X, with circa 95% of the target bases being covered at least at 20X.

WES of patient 3 was performed by Theragen Etex, Seoul, South Korea, using the SureSelect Human All Exon kit V6 (Agilent Technologies, Santa Clara, CA, USA), with a total target size of 58 Mb. Enriched genomic DNA was massively parallel sequenced on HiSeq 2,500 Sequencer (Illumina, San Diego, CA, USA). The average coverage was higher than 60X, with circa 70.9% of the target bases being covered at least at 20X.

All data were aligned to the GRCh37/hg19 reference genome build using the Burrows–Wheeler Aligner (BWA) aligner. Variants were called and quality trimmed using Genome Analysis Toolkit (GATK), and they were annotated for functional effect by SnpEff (Cingolani et al., 2012). Variants were filtered for rare variants (allele frequency < 0.005) utilizing databases such as 1,000 Genomes phase 3, NHLBI Exome Sequencing Project (ESP6500), Single Nucleotide Polymorphism database (dbSNP141), and gnomAD database using the Mendel, MD software developed in-house (Cardenas et al., 2017) and the ENLIS Genome Research software (Enlis Genomics, Berkeley, CA, USA). Only variants with impact moderate or high according to SNPeff were taken into account (Cingolani et al., 2012). To analyze the impact of the candidate variants, the software Alamut Visual version 2.15.0 (Interactive Biosoftware) was used, which showed the alignment of orthologous genes, displayed protein domains information from InterPro, and hosted protein function prediction tools such as SIFT, PolyPhen-2, MutationTaster, and Align GVGD (Tavtigian et al., 2006; Adzhubei et al., 2010; Schwarz et al., 2010; Sim et al., 2012). CADD and REVEL scores were also evaluated (Ioannidis et al., 2016; Rentzsch et al., 2019). Splice site predictions were performed using tools on Alamut Visual: MaxEntScan, SpliceSiteFinder-like, and NNSPLICE (Reese et al., 1997; Zhang, 1998; Yeo and Burge, 2004). Because the probands were Brazilians, the allele frequencies of the candidate variants were also investigated on the Online Archive of Brazilian Mutations (ABraOM), a repository containing genomic variants from 1,171 unrelated Brazilian individuals (Naslavsky et al., 2020).

## Sanger Sequencing

Sanger sequencing was performed for validation of the variants of interest identified by exome analysis using the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) and the Applied Biosystems (ABI) 3,130 Genetic Analyzer. Sequencing data were analyzed using the software Sequencher version 4.1.4 (Gene Codes Corporation).

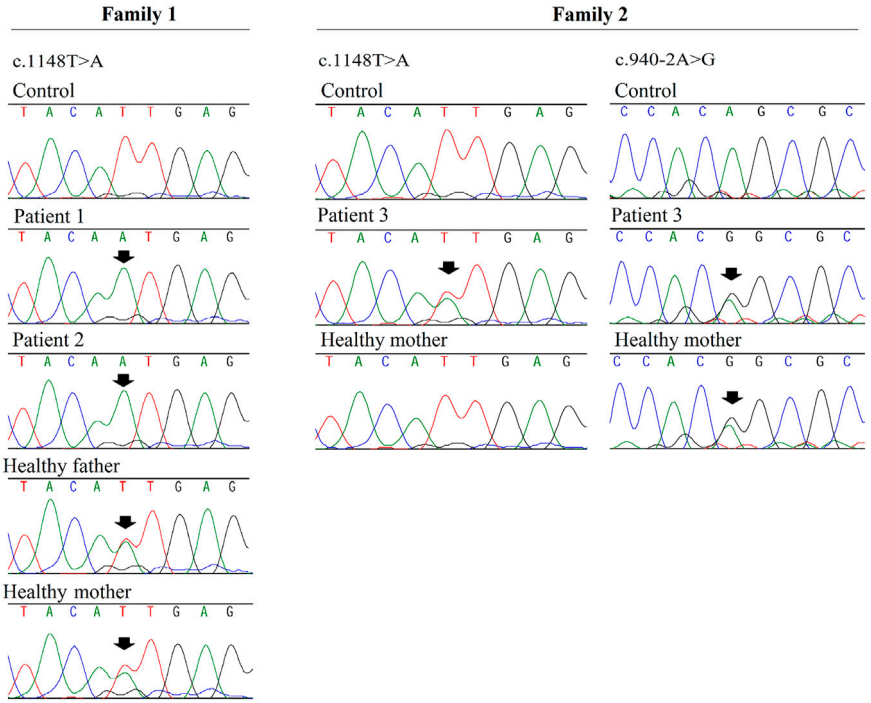



**FIGURE 2 |** Fragments of Sanger sequencing chromatograms are shown for *VPS33B* gene. The black arrows indicate the variants positions. Patients 1 and 2 (siblings) are homozygous for the c.1148T>A variant and their parents are heterozygous; patient 3 is compound heterozygous for variants c.1148T>A and c.9 40-2A>G (which was inherited from the proband's mother). DNA sample from patient 3's father was not available.

## RESULTS

We evaluated different inheritance models to filter the variants detected by WES. The fact that all parents were healthy and that one family had two children affected rendered dominant inheritance (autosomal or X-linked) unlikely. In addition, X-linked recessive inheritance was unlikely, as family 2 had one affected female proband. Consequently, we tested the autosomal recessive model of inheritance, which received support from the presence of distant consanguinity on family 1. It resulted in the identification of *VPS33B* variants in all patients.

Analysis of patients 1 and 2 identified a novel homozygous variant in exon 15 of *VPS33B*, at position chr15 (GRCh37): g.91548307A>T, NM_018,668.5(VPS33B):c.1148T>A. p.(Ile383Asn). The c.1148T>A. variant was classified as likely pathogenic (scores PM1, PM2, PM3, and PP3) according to the American College of Medical Genetics and Genomics (ACMG) and Association for Clinical Genomic Science (ACGS) Best Practice Guidelines (Richards et al., 2015; Ellard et al., 2020). This variant has not been previously described in patients with ARCS1 or in healthy individuals from worldwide populations according to the gnomAD database (Karczewski et al., 2019) and the ABraOM database (Naslavsky et al., 2020). The c.1148T>C variant (same position, but different nucleotide) has been previously registered in dbSNP under the code rs149121639, and it has been reported as having "uncertain significance" in ClinVar (RCV000372554.2). According to Alamut Visual software, the c.1148T>A. variant is located in well-established functional domains (Sec1-like, domain 2, and Vacuolar protein sorting-associated protein 33, domain 3b), and it has *in silico* pathogenic characteristics as assessed by the prediction programs SIFT ("deleterious"; score = 0.00), PolyPhen-2 ("probably damaging"; score = 0.995), MutationTaster ("disease causing"; p-value = 1), Align GVGD (Class C0), CADD ("deleterious"; score = 27.5), and REVEL ("deleterious"; score = 0.815). Altogether, the PM1 score was assigned because the variant is located in a well-established functional domain, PM2 was assigned because the variant was novel, PM3 was assigned because the variant was detected *in trans* with a pathogenic variant in patient 3, and PP3 was assigned because the variant was predicted to be pathogenic by computational tools. Sanger sequencing validated the homozygous variant in patients 1 and 2 and showed that their parents were heterozygous (**Figure 2**).

Compound heterozygous variants were identified in patient 3: he shared the same c.1148T>A. variant also present in patients 1 and 2, and additionally, he had a splice site variant located in the acceptor splice site of intron 12 of *VPS33B*, at position chr15 (GRCh37):g.91549016T>C, NM_018,668.5(VPS33B):c.940-2A>G. (HGMD accession CS041133) (Xiong et al., 2015). A skip of exon 13 was very likely; the predicted change at the acceptor site 2 bps downstream was 100% according to all splicing predictors shown on Alamut Visual (PVS1 score). This variant has been previously registered in the dbSNP dataset under the number rs774529051. Its allelic frequency was 0.00040% in healthy worldwide populations according to

**TABLE 1 |** Summarized clinical characteristics of the patients described here, and comparison with ARCS reported clinical features.

| | Reported clinical features Gissen et al. (2006) | Patient 1 | Patient 2 | Patient 3 |
|---|---|---|---|---|
| Nucleotide alterations | — | c.1148T>A. | c.1148T>A. | c.1148T>A. c.940-2A>G. |
| Alterations in coding sequence | — | p.(Ile383Asn) | p.(Ile383Asn) | p.(Ile383Asn) p.? |
| Zygosity | — | Hom | Hom | Het Het |
| Gender | — | Male | Female | Male |
| Age | — | 17 years | 10 years | 11 years |
| Classical clinical features | — | — | — | — |
| Congenital joint contractures-Arthrogryposis | + | + Mild flexion contractures | – | – |
| Renal tubular dysfunction | + | – | – | – |
| Cholestasis with normal GGT | + | + Neonatal cholestatic jaundice | + Neonatal cholestatic jaundice | + Neonatal cholestatic jaundice |
| Additional clinical features | — | — | — | — |
| Failure to thrive | + | – | + Short stature | – |
| Neurodevelopmental delay | + | + | + | + |
| Dysmorphic features | + | + | + | + |
| Hypotonia | + | – | – | – |
| Diarrhea | + | – | – | – |
| Cardiovascular anomalies | + | – | – | – |
| Ichthyosis | + | Dry and scaling skin | Dry and scaling skin | Dry and scaling skin |
| Recurrent sepsis | + | – | – | – |
| Dysgenesis of the corpus callosum | + | + | – | – |
| Sensorial hearing loss | + | – | – | + |
| Platelet alteration | + | – | – | + Large platelet |
| Hemorrhage | + | – | – | – |
| Hypothyroidism | + | – | – | – |

*+, present; –, absent.*

gnomAD database v2.1.1. This variant was classified as pathogenic according to ACMG and ACGS Best Practice Guidelines (scores PVS1, PM2, and PS4_Moderate). According to the gnomAD database, it has been reported in only one heterozygous individual from African population (PM2 score). This variant is absent from ABraOM database. It has been classified as "pathogenic" in the ClinVar database (accession RCV000730889.1), and it has previously been reported in one French individual (Gissen et al., 2006). Consequently, a PS4_Moderate score was assigned. Sanger sequencing showed that only the c.940-2A>G. variant was inherited from the mother (**Figure 2**). Thus, the c.1148T>A. variant could have been inherited from the deceased father, or it could have occurred *de novo* in the patient.

## DISCUSSION

We described here three patients with mild phenotype of ARCS with *VPS33B* variants. Patients 1 and 2, who are siblings, presented a novel homozygous c.1148T>A. p.(Ile383Asn) variant, and patient 3 was compound heterozygous for the same c.1148T>A. variant and c.940-2A>G. variant.

To date, a total of 49 unique pathogenic *VPS33B* variants and 14 pathogenic *VIPAS39* variants are listed in the Leiden Open-Source Variation Database (LOVD) for ARCS (Smith et al., 2012; Zhou and Zhang, 2014). The variant distribution is relatively uniform within *VPS33B*, with no obvious mutational hotspots (Smith et al., 2012). The c.940-2A>G. variant, present in patient 3, has been reported in one French individual (pedigree 22); it was described in compound heterozygosity with a c.240-13delTT variant in the 2 months old proband who had cholestasis with normal GGT, renal Fanconi syndrome, arthrogryposis multiplex congenita, and failure to thrive (Gissen et al., 2006).

The phenotype of ARCS is compatible with the symptoms shown by our patients: all presented cholestasis, patient 1 had discrete camptodactyly of fingers, march with *equinovarus* on the left and suggestion of dysgenesis of the corpus callosum (**Table 1**). Gissen et al. (2006) analyzed the phenotype of 62 patients with ARCS and, similarly, to our patient 1, dysgenesis of the corpus callosum and other intracranial abnormalities were reported in nine patients. In addition, arthrogryposis was present in 60 patients, and its severity ranged from isolated talipes to congenital hip dysplasia (Gissen et al., 2006).

All patients analyzed by Gissen et al. (2006) presented difficulties in gaining weight and most of the patients died within the first 6 months of life. Possibly, due to an investigation bias, the diagnosis of ARCS only would be suggested by a severe phenotype with the three cardinal signals and then confirmed by molecular analysis of the *VPS33B* and *VIPAS39* genes.

The patients studied here shared the novel missense c.1148T>A. p.(Ile383Asn) variant and presented a mild phenotype of the ARCS, with cholestasis as a main feature and without arthrogryposis or renal dysfunction. Moreover, these patients are surviving much longer than the ones with typical

ARCS phenotype. Currently, they have 17, 10, and 11 years old, respectively, and to our knowledge, patient 1 is one of the oldest patients described with ARCS to date. Other similar patients with milder phenotypes have been reported indicating the possibility of incomplete ARCS phenotype (Bull et al., 2006; Smith et al., 2012; Agawu et al., 2019; Del Brio Castillo et al., 2019; Qiu et al., 2019; Agakidou et al., 2020).

We then hypothesized that perhaps patients with missense variants in *VPS33B* gene might have an attenuated incomplete phenotype when compared with the ones with loss-of-function variants. On **Supplementary Table S1**, we compared our patient's phenotype with other patients reported in the literature with missense variants (Gissen et al., 2004; Cullinane et al., 2009; Tornieri et al., 2013; Gruber et al., 2017; Del Brio Castillo et al., 2019; Lee et al., 2019; Qiu et al., 2019; Seidl-Philipp et al., 2020).

Only 11 patients have been reported with pathogenic *VPS33B* missense variants in the literature and two of them had no detailed phenotypic data described (Gissen et al., 2004; Cullinane et al., 2009; Tornieri et al., 2013) (**Supplementary Table S1**). With the exception of the patient described by Lee et al. (2019), all patients that were reported with missense variants had a milder phenotype. However, further patients with missense variants are needed to validate this hypothesis. The proband described by Lee et al. (2019) had a missense variant in one allele and a splice site variant in the other allele, which could be influencing her phenotype (she carried p.Asp236Val and c.239+5G>A variants). Interestingly, three patients with the same missense p.(Gly131Glu) variant were described as having the phenotype of Keratoderma-ichthyosis-deafness (ARKID) syndrome, a rare multisystem disorder also caused by biallelic mutations in *VPS33B* (Gruber et al., 2017; Seidl-Philipp et al., 2020). It is important to note that at least six patients with incomplete phenotype have been reported with loss of functions variants, which shows that an incomplete phenotype is not always caused by missense variants—these patient's phenotypes were also summarized on the **Supplementary Table S1** (Bull et al., 2006; Smith et al., 2012; Agawu et al., 2019; Agakidou et al., 2020; Duong et al., 2020b; a).

The genetic and clinical features of the previously reported patients were reviewed by Smith et al. (2012), and they provided the first evidence of genotype-phenotype correlation in ARCS. They reported two patients with an attenuated ARCS phenotype, who were compound heterozygous for the same c.1225+5G>C variant, resulting in the expression of a shorter VPS33B protein product that retained some ability to interact with VIPAS39. Other previous studies also suggested that variants in patients with complete ARCS phenotype caused absent VPS33B protein expression or abolished the interaction with VIPAS39, whereas variants in patients with attenuated phenotype would be less severe with partially preserved protein expression and function (Cullinane et al., 2009; Smith et al., 2012; Qiu et al., 2019).

Here, we report three patients with the same novel c.1148T>A. variant, and we believe that this variant could be associated with a mild phenotype. Further cell-based assays would be necessary to analyze if this variant would retain the VPS33B protein ability to interact with VIPAS39, similarly to the previously discussed studies (Cullinane et al., 2009; Smith et al., 2012; Qiu et al., 2019).

Those subtypes/incomplete phenotypes make it difficult to differentiate through routine clinical investigations without genetic analysis. Our patients with low GGT cholestasis would remain undiagnosed without genetic tests. Thus, we raised awareness of the mild clinical picture of ARCS, and we propose that molecular analysis of the *VPS33B* and *VIPAS39* should be considered in patients with normal GGT cholestasis, and not only for patients with the complete ARCS phenotype. Other potential cause of normal GGT cholestasis that should be considered is progressive familial intrahepatic cholestasis (PFIC), which is a heterogeneous group of autosomal recessive disorders that accounts for 10%–15% of the cholestasis cases in children (Davit-Spraul et al., 2009). Because PFIC has a higher incidence than ARCS, it should be initially considered. As a consequence of this approach, it is expected that the number of atypical ARCS diagnoses may increase as well.

## CONCLUSION

In conclusion, we described here three patients with ARCS diagnosed by WES analysis. They carry the same novel c.1148T>A. variant, and we believe that this variant could be associated with a mild phenotype. Classical clinical diagnosis would not be appropriate for patients with mild phenotype of this syndrome. We propose here that *VPS33B* and *VIPAS39* mutation screening in patients with normal GGT cholestasis could facilitate accurate diagnosis and the administration of supportive care at early stage, in addition to provide genetic counseling for the affected families. No specific treatment currently exists for ARCS, but advances in knowledge or ARCS pathogenesis may lead to novel therapies and improved management, which are valuable in patients with prolonged survival.

## DATA AVAILABILITY STATEMENT

The datasets for this article are not publicly available due to concerns regarding participant/patient anonymity. Requests to access the datasets should be directed to the corresponding author.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Research Ethics Committee of the Hospital das Clinicas of UFMG. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

## AUTHOR CONTRIBUTIONS

SP conceived the study. NL conducted bioinformatics analysis of patients' exomes and Sanger sequencing experiments. NL wrote the manuscript with the help of all authors, EF, AF, TQ, and LdS

enrolled the patients' families and were involved in the clinical aspect of the study. All authors read and approved the final version.

## FUNDING

NL was supported by a fellowship from Conselho Nacional de Desenvolvimento Científco e Tecnológico (CNPq). SP received support as a Scientist 1A of the CNPq. This study was funded by the CNPq and the Fundação de Amparo à Pesquisa do Estado Minas Gerais (FAPEMIG).

## ACKNOWLEDGMENTS

We are grateful to the patients studied here and their families. Their cooperation made this work possible.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.796759/full#supplementary-material

## REFERENCES

Abu-Sa'da, O., Barbar, M., Al-Harbi, N., and Taha, D. (2005). Arthrogryposis, Renal Tubular Acidosis and Cholestasis (ARC) Syndrome: Two New Cases and Review. *Clin. Dysmorphol.* 14 (4), 191–196. doi:10.1097/00019605-200510000-00005

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 (4), 248–249. doi:10.1038/nmeth0410-248

Agakidou, E., Agakidis, C., Kambouris, M., Printza, N., Farini, M., Vourda, E., et al. (2020). A Novel Mutation of VPS33B Gene Associated with Incomplete Arthrogryposis-Renal Dysfunction-Cholestasis Phenotype. *Case Rep. Genet.* 2020, 1–8. doi:10.1155/2020/8872294

Agawu, A., Sheppard, S., and Lin, H. C. (2019). A Novel VPS33B Mutation Causing a Mild Phenotype of Arthrogryposis, Renal Dysfunction, and Cholestasis Syndrome. *J. Pediatr. Gastroenterol. Nutr.* 69 (2), e55–e56. doi:10.1097/MPG.0000000000002306

Bull, L. N., Mahmoodi, V., Baker, A. J., Jones, R., Strautnieks, S. S., Thompson, R. J., et al. (2006). VPS33B Mutation with Ichthyosis, Cholestasis, and Renal Dysfunction but without Arthrogryposis: Incomplete ARC Syndrome Phenotype. *J. Pediatr.* 148 (2), 269–271. doi:10.1016/j.jpeds.2005.10.005

Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., et al. (2012). A Program for Annotating and Predicting the Effects of Single Nucleotide Polymorphisms, SnpEff. *Fly* 6 (2), 80–92. doi:10.4161/fly.19695

Cullinane, A. R., Straatman-Iwanowska, A., Seo, J. K., Ko, J. S., Song, K. S., Gizewska, M., et al. (2009). Molecular Investigations to Improve Diagnostic Accuracy in Patients with ARC Syndrome. *Hum. Mutat.* 30 (2), E330–E337. doi:10.1002/humu.20900

Davit-Spraul, A., Gonzales, E., Baussan, C., and Jacquemin, E. (2009). Progressive Familial Intrahepatic Cholestasis. *Orphanet J. Rare Dis.* 4, 1. doi:10.1186/1750-1172-4-1

Del Brio Castillo, R., Squires, J. E., and McKiernan, P. J. (2019). A Novel Mutation in VPS33B Gene Causing a Milder ARC Syndrome Phenotype with Prolonged Survival. *JIMD Rep.* 47 (1), 4–8. doi:10.1002/jmd2.12027

Duong, M. D., Rose, C. M., Reidy, K. J., and Del Rio, M. (2020a). An Uncommon Case of Arthrogryposis, Renal Dysfunction, and Cholestasis (ARC) Syndrome and Review of the Renal Involvement: Answers. *Pediatr. Nephrol.* 35 (2), 249–251. doi:10.1007/s00467-019-04338-z

Duong, M. D., Rose, C. M., Reidy, K. J., and Del Rio, M. (2020b). An Uncommon Case of Arthrogryposis, Renal Dysfunction, and Cholestasis (ARC) Syndrome and Review of the Renal Involvement: Questions. *Pediatr. Nephrol.* 35 (2), 247–248. doi:10.1007/s00467-019-04336-1

Ellard, S., Baple, E. L., Callaway, A., Berry, I., Forrester, N., Turnbull, C., et al. (2020). ACGS Best Practice Guidelines for Variant Classification in Rare Disease 2020, [Online]. *Assoc. Clin. Genet. Sci.*. Available: https://www.acgs.uk.com/media/11631/uk-practice-guidelines-for-variant-classification-v4-01-2020.pdf (Accessed May 05, 2021).

Cardenas, R. G. C. C. L., D. Linhares, N., L. Ferreira, R., and Pena, S. D. J. (2017). Mendel,MD: A User-Friendly Open-Source Web Tool for Analyzing WES and WGS in the Diagnosis of Patients with Mendelian Disorders. *Plos Comput. Biol.* 13 (6), e1005520. doi:10.1371/journal.pcbi.1005520

Gissen, P., Johnson, C. A., Morgan, N. V., Stapelbroek, J. M., Forshew, T., Cooper, W. N., et al. (2004). Mutations in VPS33B, Encoding a Regulator of SNARE-dependent Membrane Fusion, Cause Arthrogryposis-Renal Dysfunction-Cholestasis (ARC) Syndrome. *Nat. Genet.* 36 (4), 400–404. doi:10.1038/ng1325

Gissen, P., Tee, L., Johnson, C. A., Genin, E., Caliebe, A., Chitayat, D., et al. (2006). Clinical and Molecular Genetic Features of ARC Syndrome. *Hum. Genet.* 120 (3), 396–409. doi:10.1007/s00439-006-0232-z

Gruber, R., Rogerson, C., Windpassinger, C., Banushi, B., Straatman-Iwanowska, A., Hanley, J., et al. (2017). Autosomal Recessive Keratoderma-Ichthyosis-Deafness (ARKID) Syndrome Is Caused by VPS33B Mutations Affecting Rab Protein Interaction and Collagen Modification. *J. Invest. Dermatol.* 137 (4), 845–854. doi:10.1016/j.jid.2016.12.010

Han, J., Pluhackova, K., and Böckmann, R. A. (2017). The Multifaceted Role of SNARE Proteins in Membrane Fusion. *Front. Physiol.* 8, 5. doi:10.3389/fphys.2017.00005

Horslen, S. P., Quarrell, O. W., and Tanner, M. S. (1994). Liver Histology in the Arthrogryposis Multiplex Congenita, Renal Dysfunction, and Cholestasis (ARC) Syndrome: Report of Three New Cases and Review. *J. Med. Genet.* 31 (1), 62–64. doi:10.1136/jmg.31.1.62

Ioannidis, N. M., Rothstein, J. H., Pejaver, V., Middha, S., McDonnell, S. K., Baheti, S., et al. (2016). REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. *Am. J. Hum. Genet.* 99 (4), 877–885. doi:10.1016/j.ajhg.2016.08.016

Karczewski, K. J., Francioli, L. C., Tiao, G., Cummings, B. B., Alföldi, J., Wang, Q., et al. (2019). Variation across 141,456 Human Exomes and Genomes Reveals the Spectrum of Loss-Of-Function Intolerance across Human Protein-Coding Genes. *bioRxiv*, 531210. doi:10.1101/531210

Lee, M. J., Suh, C. R., Shin, J. H., Lee, J. H., Lee, Y., Eun, B.-L., et al. (2019). A Novel VPS33B Variant Identified by Exome Sequencing in a Patient with Arthrogryposis-Renal Dysfunction-Cholestasis Syndrome. *Pediatr. Gastroenterol. Hepatol. Nutr.* 22 (6), 581–587. doi:10.5223/pghn.2019.22.6.581

Lobingier, B. T., and Merz, A. J. (2012). Sec1/Munc18 Protein Vps33 Binds to SNARE Domains and the Quaternary SNARE Complex. *MBoC* 23 (23), 4611–4622. doi:10.1091/mbc.E12-05-0343

Miller, S. A., Dykes, D. D., and Polesky, H. F. (1988). A Simple Salting Out Procedure for Extracting DNA from Human Nucleated Cells. *Nucl. Acids Res.* 16 (3), 1215. doi:10.1093/nar/16.3.1215

Naslavsky, M. S. Scliar, M. O. Yamamoto, G. L. Wang, J. Y. T. Zverinova, S. Karp, T., et al. (2020202020092). *Whole-genome Sequencing of 1,171 Elderly Admixed Individuals from the Largest Latin American metropolis (São Paulo, Brazil)*, 298026. doi:10.1101/2020.09.15.298026*bioRxiv*Whole-genome Sequencing of 1,171 Elderly Admixed Individuals from the Largest Latin American metropolis (São Paulo, Brazil)

Nezelof, C., Dupart, M. C., Jaubert, F., and Eliachar, E. (1979). A Lethal Familial Syndrome Associating Arthrogryposis Multiplex Congenita, Renal Dysfunction, and a Cholestatic and Pigmentary Liver Disease. *J. Pediatr.* 94 (2), 258–260. doi:10.1016/S0022-3476(79)80839-2

Qiu, Y. L., Liu, T., Abuduxikuer, K., Hao, C. Z., Gong, J. Y., Zhang, M. H., et al. (2019). Novel Missense Mutation in VPS33B Is Associated with Isolated Low Gamma-glutamyltransferase Cholestasis: Attenuated, Incomplete Phenotype of Arthrogryposis, Renal Dysfunction, and Cholestasis Syndrome. *Hum. Mutat.* 40 (12), 2247–2257. doi:10.1002/humu.23770

Reese, M. G., Eeckman, F. H., Kulp, D., and Haussler, D. (1997). Improved Splice Site Detection in Genie. *J. Comput. Biol.* 4 (3), 311–323. doi:10.1089/cmb.1997.4.311

Rentzsch, P., Witten, D., Cooper, G. M., Shendure, J., and Kircher, M. (2019). CADD: Predicting the Deleteriousness of Variants throughout the Human Genome. *Nucleic Acids Res.* 47 (D1), D886–D894. doi:10.1093/nar/gky1016

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

Rogerson, C., and Gissen, P. (2018). VPS33B and VIPAR Are Essential for Epidermal Lamellar Body Biogenesis and Function. *Biochim. Biophys. Acta (Bba) - Mol. Basis Dis.* 1864 (5 Pt A), 1609–1621. doi:10.1016/j.bbadis.2018.01.028

Schwarz, J. M., Rödelsperger, C., Schuelke, M., and Seelow, D. (2010). MutationTaster Evaluates Disease-Causing Potential of Sequence Alterations. *Nat. Methods* 7 (8), 575–576. doi:10.1038/nmeth0810-575

Seidl-Philipp, M., Schatz, U. A., Gasslitter, I., Moosbrugger-Martinz, V., Blunder, S., Schossig, A. S., et al. (2020). Spectrum of Ichthyoses in an Austrian Ichthyosis Cohort from 2004 to 2017. *JDDG: J. der Deutschen Dermatologischen Gesellschaft* 18 (1), 17–25. doi:10.1111/ddg.13968

Sim, N.-L., Kumar, P., Hu, J., Henikoff, S., Schneider, G., and Ng, P. C. (2012). SIFT Web Server: Predicting Effects of Amino Acid Substitutions on Proteins. *Nucleic Acids Res.* 40, W452–W457. doi:10.1093/nar/gks539

Smith, H., Galmes, R., Gogolina, E., Straatman-Iwanowska, A., Reay, K., Banushi, B., et al. (2012). Associations Among Genotype, Clinical Phenotype, and Intracellular Localization of Trafficking Proteins in ARC Syndrome. *Hum. Mutat.* 33 (12), 1656–1664. doi:10.1002/humu.22155

Tavtigian, S. V., Deffenbaugh, A. M., Yin, L., Judkins, T., Scholl, T., Samollow, P. B., et al. (2006). Comprehensive Statistical Study of 452 BRCA1 Missense Substitutions with Classification of Eight Recurrent Substitutions as Neutral. *J. Med. Genet.* 43 (4), 295–305. doi:10.1136/jmg.2005.033878

Tornieri, K., Zlatic, S. A., Mullin, A. P., Werner, E., Harrison, R., L'Hernault, S. W., et al. (2013). Vps33b Pathogenic Mutations Preferentially Affect VIPAS39/SPE-39-Positive Endosomes. *Hum. Mol. Genet.* 22 (25), 5215–5228. doi:10.1093/hmg/ddt378

Xiong, H. Y., Alipanahi, B., Lee, L. J., Bretschneider, H., Merico, D., Yuen, R. K. C., et al. (2015). The Human Splicing Code Reveals New Insights into the Genetic Determinants of Disease. *Science* 347 (6218), 1254806. doi:10.1126/science.1254806

Yeo, G., and Burge, C. B. (2004). Maximum Entropy Modeling of Short Sequence Motifs with Applications to RNA Splicing Signals. *J. Comput. Biol.* 11 (2-3), 377–394. doi:10.1089/1066527041410418

Zhang, M. (1998). Statistical Features of Human Exons and Their Flanking Regions. *Hum. Mol. Genet.* 7 (5), 919–932. doi:10.1093/hmg/7.5.919

Zhou, Y., and Zhang, J. (2014). Arthrogryposis-renal Dysfunction-Cholestasis (ARC) Syndrome: from Molecular Genetics to Clinical Features. *Ital. J. Pediatr.* 40 (1), 1–7. doi:10.1186/s13052-014-0077-3

**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.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

*Copyright © 2022 Linhares, Fagundes, Ferreira, Queiroz, da Silva and Pena. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*

frontiers in Genetics

ORIGINAL RESEARCH
published: 02 March 2022
doi: 10.3389/fgene.2022.785570

# Copy Number Variations in Genetic Diagnosis of Congenital Adrenal Hyperplasia Children

*Aisha Tolba[1†], Iman Mandour[1], Noha Musa[2], Fatma Elmougy[1], Mona Hafez[2], Sahar Abdelatty[1], Amany Ibrahim[2], Hend Soliman[2], Bahaaeldin Labib[3], Yasmine Elshiwy[1], Tarek Ramzy[1] and Marwa Elsharkawy[1*†]*

[1]Clinical and Chemical Pathology Department, Cairo University, Giza, Egypt, [2]Diabetes, Endocrine and Metabolism Pediatric Unit, Cairo University, Giza, Egypt, [3]Royal College of Surgeons in Ireland, Medical University of Bahrain, Giza, Egypt

**Background:** Congenital adrenal hyperplasia (CAH) is a monogenic disorder caused by genetic diversity in the *CYP21A2* gene, with 21-hydroxylase deficiency (21-OHD) as the most common type. Early sex assignment and early diagnosis of different genetic variations with a proper technique are important to reduce mortality and morbidity. Proper early sex identification reduces emotional, social, and psychological stress.

**Aim:** Detection of a spectrum of aberrations in the *CYP21A2* gene, including copy number variations, gene conversion, chimeric genes, and point variations.

**Methods:** The *CYP21A2* gene was screened using MLPA assay in 112 unrelated Egyptian children with 21-OHD CAH (33 males and 79 females).

**Results:** In the studied group, 79.5% were diagnosed within the first month of life. 46.8% of the genetic females were misdiagnosed as males. Among the copy number variation results, large deletions in 15.4% and three types of chimeric genes in 9% (CH-1, CH-7, and CAH-X CH-1) were detected. Regarding gene dosage, one copy of *CYP21A2* was found in 5 cases (4.5%), three copies were detected in 7 cases (6.3%), and one case (0.9%) showed four copies. Eight common genetic variants were identified, I2G, large deletions, large gene conversion (LGC), I172N, F306 + T, -113 SNP, 8bp Del, and exon 6 cluster (V237E and M239K) with an allelic frequency of 32.62%, 15.45%, 7.30%, 3.00%, 2.58%, 2.15%, 0.86%, and 0.86%, respectively.

**Conclusion:** High prevalence of copy number variations highlights the added value of using MLPA in routine laboratory diagnosis of CAH patients.

**Keywords: 21-Hydroxylase deficiency, CYP21A2, multiplex ligation-dependent probe amplification, copy number variations, congenital adrenal hyperplasia**

OPEN ACCESS

***Edited by:***
*Kathleen M Gorman, Temple Street Children's University Hospital, Ireland*

***Reviewed by:***
*Fatma Savran OGUZ, Istanbul University, Turkey*
*Enrique Medina-Acosta, State University of the North Fluminense Darcy Ribeiro, Brazil*

***Correspondence:**
*Marwa Elsharkawy*
*marwaelsharkawy@yahoo.com*

*[†]These authors have contributed equally to this work and share first authorship*

***Specialty section:***
*This article was submitted to Genetics of Common and Rare Diseases, a section of the journal Frontiers in Genetics*

***Received:*** *29 September 2021*
***Accepted:*** *31 January 2022*
***Published:*** *02 March 2022*

***Citation:***
*Tolba A, Mandour I, Musa N, Elmougy F, Hafez M, Abdelatty S, Ibrahim A, Soliman H, Labib B, Elshiwy Y, Ramzy T and Elsharkawy M (2022) Copy Number Variations in Genetic Diagnosis of Congenital Adrenal Hyperplasia Children. Front. Genet. 13:785570. doi: 10.3389/fgene.2022.785570*

## INTRODUCTION

Congenital adrenal hyperplasia (CAH) represents a group of common defects in the adrenal gland steroidogenesis; the main form is 21-hydroxylase enzyme deficiency (21-OHD; OMIM #201910) (Chan et al., 2011). Patients with the classic phenotype can present with life-threatening salt-wasting (SW) and simple virilizing (SV) types depending on the enzymatic activity of the 21-OH enzyme (Rabbani et al., 2011). There is a higher prevalence of CAH in Egypt (1 in 1,209 live births) and Arab

countries due to a higher rate of consanguinity (Temtamy and Aglan, 2012) (Tayel et al., 2011) (Alshabab et al., 2015).

In males, CAH is a medical emergency, with an urgent need for early diagnosis as it can be fatal in cases with SW-CAH. In females, CAH-SV can also be considered a social emergency (Kutney et al., 2016). The sex of a newborn is usually determined at birth on the basis of external genital appearance. Thus, children with atypical genitalia frequently require reassignment of sex because of incorrect original identification. As sex is a fundamental attribute of human life, its reversal after the original assignment is usually associated with marked emotional, social, and psychological stress with accompanying shame (Wisniewski, 2017).

First-line diagnostic tests for CAH include karyotype analysis, hormonal evaluation (from 48 h after birth) for 17-hydroxyprogesterone (17OHP), androgens, basal cortisol, and adrenocorticotropic hormone (ACTH), and ultrasound (for detecting Mullerian structures). The definitive diagnosis of 21OHD CAH is through the assessment of serum 17OHP levels (levels above 10 ng/dl are diagnostic). Additional imaging tests may be recommended, including genogram, retrograde urethrogram, or cystoscopy/vaginoscopy (in females with atypical genitalia). When the gonads are not detected by ultrasound, magnetic resonance imaging is indicated. Molecular studies are performed to check for gene mutations (Guerrero-Fernández et al., 2018).

Management of CAH cases includes lifelong hormonal therapy and reconstructive surgery for female cases with atypical genitalia (feminizing genitoplasty) (Savanelli et al., 2008). This is best done in early infancy to confer an early physical appearance consistent with the gender/sex of rearing and to cause less psychological trauma than if delayed (Speiser et al., 2018; Musa et al., 2020).

While CAH-21OH is a monogenic disorder, the human 21-OH gene (cytochrome P450 family 21 subfamily A member 2, CYP21A2) encodes the 21-OH enzyme, and a highly homologous inactive pseudogene, CYP21A1P, is located closely adjacent to the CYP21A2 gene, approximately 30 kb apart in tandem arrangement with the C4A and C4B genes (Blaskó et al., 2009; Choi et al., 2016; Daae et al., 2018). These arrangements comprise the most frequent bimodular RCCX of the RP1-C4-CYP21A1P-TNXA-RP2-C4-CYP21A2-TNXB gene sequence in 69% of alleles in Caucasians (Lee et al., 2010; Tsai et al., 2012). Recombination events commonly occur when a loop forms in the linear DNA of the chromosome. The active CYP21A2 and inactive CYP21A1P genes lie against each other, allowing the transfer of genetic material between them. The advantage of this physiological event is that it increases immunological diversity over time. However, the CYP21A2 gene sits adjacent to the HLA locus and becomes a passive participant in this process (Huynh et al., 2009).

The high degree of sequence similarity between active and pseudogene allows two types of intergenic recombination events that are responsible for about 95% of the mutations associated with 21-OHD (Vrzalová et al., 2011). The first is gene conversion events: approximately 65%–70% of the deleterious mutations are derived from pseudogene CYP21A1P due to small gene conversions, including I2G (28%), I172N (9%), V281L (9%), Q318X (4%), R356W (4%), E6 cluster [I235N, V236E, and M238K] (4%), 8bp Del (3%), P30L (2%), and F306 + T (1%) (Xu et al., 2013). If these transfers comprise a large region, it is called a "large-scale conversion" (Toraman et al., 2013).

When targeted mutation analysis detects multiple mutations, it is possible that the mutations are most likely arising from gene conversion (Nimkarn et al., 1993). The second is unequal crossing over during meiosis: the remaining 20%–25% of the intergenic recombination is represented by CYP21A2 gene deletions or CYP21A1P/CYP21A2 chimeric genes (Concolino et al., 2009). Nine different CYP21A1P/CYP21A2 chimeric genes were described according to their breakpoints (CH-1 to CH-9) (Chen et al., 2012; Veldhuisen et al., 2015). The third common variation in CYP21A2 is spontaneous point mutations (Xu et al., 2013).

Multiplex Ligation-dependent Probe Amplification (MLPA) assay with the addition of common point-mutation (PMS) MLPA probes represents the gold standard in CAH diagnosis and a sensitive tool for identification of chimerical genes (Choi et al., 2016) (Erlandson et al., 2008).

The aim of this study is to perform genetic analysis of CAH 21-OHD Egyptian children to detect a spectrum of aberrations in the CYP21A2 gene, including copy number variations, gene conversion, gene deletion, chimeric genes, and point variations using MLPA analysis.

## MATERIALS AND METHODS

Editorial policies and ethical considerations: children with CAH were recruited from the Diabetes Endocrine Metabolism Pediatric Unit (DEMPU), Children's Hospital, Cairo University, after obtaining written informed consent from all participants or their legal guardians. The study was approved by the Research Ethics Committee of Cairo University. The diagnosis of CAH due to 21-OHD (salt wasting and simple virilizing) was based on both clinical and biochemical examinations (elevated 17-hydroxyprogesterone).

Three millilitres of blood were collected from each subject in a sterile EDTA vacutainer for the genotyping technique. DNA was extracted using the DNA extraction kit QIAamp (DNA Blood Mini Kit) supplied by QIAGEN (Thermo Fisher Scientific, USA) according to the manufacturer's instructions. Then Selective Adaptor Ligation, Selective Amplification (SALSA) MLPA Probemix P050-C1 CAH (Catalog no. P050-100R) in combination with the SALSA MLPA EK1 reagent kit (EK1-FAM) supplied by MRC-Holland was used.

Multiplex ligation-dependent probe amplification analysis of the CYP21A2 gene included the following steps according to the manufacturer's instructions: DNA denaturation, hybridization reaction, ligation reaction, PCR reaction, and fragment separation by capillary electrophoresis, ABI 3500 Genetic Analyzer (Applied Biosystems, Thermo Fisher Scientific, USA).

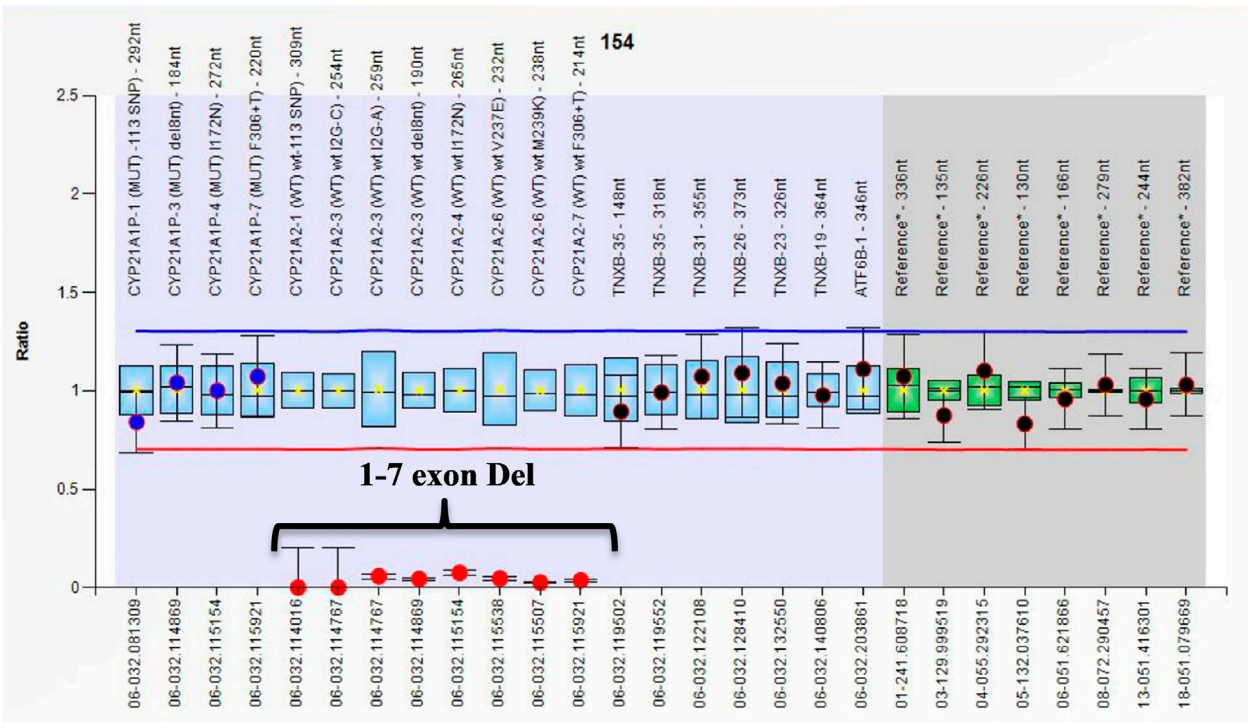



**FIGURE 1 |** Ratio chart showing the homozygous 1–7 exon Del CYP21A2 gene.

## Bioinformatics and Statistical Analysis

The raw data were analysed using Coffalyser.net software, the MLPA analysis tool developed at MRC-Holland. It was used for data analysis in combination with the appropriate lot-specific MLPA Coffalyser sheet. They were available on www.mlpa.com.

The nomenclature of the variants reported in the current study are according to the Human Genome Variation Society (HGVS) (denDunnen et al., 2016) as follows (NM_000500.8): 113 SNP (rs1246774295, c.-113G > A), I2G (rs6467, c.293–13C > G), 8bp Del (rs387906510, c.332_339delGAGACTAC, c.332_339del), I172N (rs6475, c.518T > A), Exon 6 cluster variants [V237E (rs12530380, c.713T > A), M239K (rs6476), c.719T > A], and F306 + T (rs267606756, c.923dup).

Data were analysed using IBM SPSS Advanced Statistics version 21.0 (IBM Corp., Armonk, NY). Qualitative data were expressed as frequency and percentage.

## RESULTS

One hundred and twelve probands with 21-OHD were recruited for the study. The majority (79.5%, 89/122) were diagnosed in the first month of life. 50/112 (44.6%) were diagnosed at birth (day 0) and 39/112 (34.8%) between1–30 days. Seventy-six (67.9%) patients were born to consanguineous parents. By karyotyping, the genetically confirmed females constituted 79/112 (70.5%) and 33/112 (29.5%) were males. The female-to-male ratio was 2.4:1, and 37/79 (46.8%) of genetic females were misdiagnosed as males.

**TABLE 1 |** The agreement between MLPA and karyotype among studied cases (n = 112).

| Karyotype | MLPA | | Total |
|---|---|---|---|
| | Male, *n* (%) | Female, *n* (%) | |
| 46XY | 32 (97) | 1 (3) | 33 |
| 46XX | 7 (8.9) | 72 (91.1) | 79 |
| Total | 39 | 73 | 112 |
| Kappa value* (*p* value) | 0.837 (<0.001) | | |

**Kappa value: < 0: less than chance agreement; 0.01–0.2: slight agreement; 0.21–0.4: fair agreement; 0.41–0.6: moderate agreement; 0.61–0.8: substantial agreement; 0.81–0.99: almost perfect.*

The majority of females presented with ambiguous genitalia (78/79, 98.7%). In most cases, 49/79 (62%) presented with Prader stage III, followed by stage IV (16.5%). The remaining cases presented with Prader stage II (10.1%), I (7.6%), and V (2.5%), while 4/33 males (12.1%) presented with precocious puberty.

All cases in the current study had the classical form of 21-OHD. Eighty-eight out of 112 patients (78.6%) were SW, and 24/112 (21.4%) cases had the SV form. The phenotype of all cases with large gene conversions and/or deletions is SW except for one case that was SV (LGC 1-3 exon on one allele and Del 1-7 exon Del on the other allele). **Figure 1** shows the ratio chart result of 1–7 exon deletion (Del 1-7).

Regarding the gene dosage, 67% of cases (75/112) had two copies of the active gene. Three copies and one copy of CYP21A2 were found in 7/112 (6.3%) and 5/112 (4.5%) cases, respectively. Zero copies were reported in 24/112 (21.4%), and four copies

**TABLE 2 |** Genotype frequency among different phenotypes (n = 112).

| Variant | Zygosity | Genotype | Phenotype | | *n* (%) |
|---|---|---|---|---|---|
| | | | SW, *n* (%) | SV, *n* (%) | |
| I2G | Hetero | I2G/N | 8 (80) | 2 (20) | 10 (8.9) |
| | Homo | I2G/I2G | 18 (72) | 7 (28) | 25 (22.3) |
| I172N | Hetero | I172 N/N | 1 (100) | 0 (0) | 1 (0.9) |
| | Homo | I172N/I172N | 0 (0) | 2 (100) | 2 (1.8) |
| F306+T | Homo | F306 + T/F306 + T | 2 (100) | 0 (0) | 2 (1.8) |
| LGC (1–3 exon) | Hetero | LGC/N | 1 (100) | 0 (0) | 1 (0.9) |
| | Homo | LGC/LGC | 3 (100) | 0 (0) | 3 (2.7) |
| 1–3 exon Del | Homo | Del 1–3/Del 1–3 | 1 (100) | 0 (0) | 1 (0.9) |
| 1–7 exon Del | Homo | Del 1–7/Del 1–7 | 4 (100) | 0 (0) | 4 (3.6) |
| -113 SNP, I2G | Hetero | (-113 SNP, I2G)/N or* -113 SNP/I2G | 1 (100) | 0 (0) | 1 (0.9) |
| | Homo | (-113 SNP, I2G)/(-113 SNP, I2G) | 2 (100) | 0 (0) | 2 (1.8) |
| I2G, 8bp Del | Hetero | (I2G, 8bp Del)/N or* I2G/8bp Del | 2 (100) | 0 (0) | 2 (1.8) |
| I2G, I172N | Hetero | (I2G, I172N)/N or* I2G/I172N | 1 (100) | 0 (0) | 1 (0.9) |
| I2G, E6† | Hetero | (I2G, E6)/N or* I2G/E6 | 2 (100) | 0 (0) | 2 (1.8) |
| I2G, F306+T | Hetero | (I2G, F306 + T)/N or* I2G/F306 + T | 1 (100) | 0 (0) | 1 (0.9) |
| I172N, F306+T | Hetero | (I172N, F306 + T)/N or* I172N/F306 + T | 0 (0) | 1 (100) | 1 (0.9) |
| LGC (1–3 exon), I2G | Hetero | LGC/I2G | 2 (100) | 0 (0) | 2 (1.8) |
| LGC (1–4 exon), I2G | Hetero | LGC/I2G | 1 (100) | 0 (0) | 1 (0.9) |
| 1–3 exon Del, I2G | Hetero | Del 1–3/I2G | 1 (100) | 0 (0) | 1 (0.9) |
| LGC (1–3 exon), 1–3 exon Del | Hetero | LGC/Del 1–3 | 4 (100) | 0 (0) | 4 (3.6) |
| LGC (1–3 exon), 1–7 exon Del | Hetero | LGC/Del 1–7 | 0 (0) | 1 (100) | 1 (0.9) |
| LGC (1–7 exon), 1–7 exon Del | Hetero | LGC/Del 1–7 | 1 (100) | 0 (0) | 1 (0.9) |
| LGC (1–4 exon), 30-KB Del (CH-7) | Hetero | LGC/Del | 1 (100) | 0 (0) | 1 (0.9) |
| 30-KB Del (CH-1) | Homo | Del/Del | 6 (100) | 0 (0) | 6 (5.4) |
| Large gene Del (CAH-X CH-1) | Homo | Del[a]/Del[a] | 3 (100) | 0 (0) | 3 (2.7) |
| No identified variant | Homo | N/N | 22 (66.7) | 11 (33.3) | 33 (29.5) |
| Total | | | 88 | 24 | 112 |

*LGC, large gene conversion; Del, deletion; N, no identified variant; Homo, homozygous; Hetero, heterozygous; CH-1, chimeric gene produced by 30-KB, Del extending from exon 4 of the* CYP21A1P *pseudogene to exon 3 of the active* CYP21A2 *gene; CH-7, chimeric gene produced by 30-KB, Del extending from exon 7 of the* CYP21A1P *pseudogene to exon 6 of the active* CYP21A2 *gene; Del[a], large gene deletion from the* CYP21A2 *gene extending to exon 35 of the* TNXB *gene.*

*†E6: exon 6 cluster variants (V237E and M239K), SW: salt wasting, SV: simple virilizing.*

**Compound heterozygous variants may be in* cis *or* trans *configuration and needs further investigations for both parents.*

were reported in only one patient (0.9%). Genotype results of cases with zero copies of the active gene show homozygous LGC or large deletions.

The majority of single nucleotide variants (SNV) were in a homozygous state, 29/112 (25.9%). Heterozygous SNV was found in 11/112 (9.8%) cases. Compound heterozygous SNV and compound homozygous single nucleotide variants were present in 8/112 (7.1%) and 2/112 (1.8%) cases, respectively. The frequency of genotype distribution of CYP21A2 gene variants among cases is presented in **Table 1**.

## DISCUSSION

To the best of our knowledge, the study includes the largest sample size among Arab countries and the first one to use MLPA assay for the diagnosis of copy number variations in a group of 21-OHD CAH children in Egypt. One hundred and twelve consecutive unrelated 21-OHD CAH patients were enrolled in the present study. The results revealed that 89/112 (79.5%) were diagnosed within the first month of life and highlight the necessity of a neonatal screening program for CAH in Egypt.

All cases in the current study had the classical form of 21-OHD. Eighty-eight out of 112 patients (78.6%) were salt wasting (SW) and 24 out of 112 (21.4%) cases had the simple virilizing (SV) form. Fortunately, affected males are fewer than females, as males with the SW form are at higher risk of adrenal crises and may pass away unnoticed without proper diagnosis. In our study, the female-to-male ratio was 2:1. This is in accordance with previous studies that found the female: male ratio was 3:1 (Al-obaidi et al., 2016).

The wrong sex identification of affected females reared as males in 46.8% of cases can be explained by the high prevalence of female ambiguous genitalia. In the present study, all female cases presented with ambiguous genitalia, except for one case that received prenatal treatment. Reassignment of gender to the female sex in cases with 46XX 21OHD CAH is recommended (Freire et al., 2018) and occurred in 36/37 of the genetic females that were misdiagnosed as males. One genetic female only continued as a male after surgical operations at the request of the parents.

One of the potential roles of MLPA is genetic sex identification due to the presence of probes for y and x chromosomes. The results of sex determination by MLPA showed agreement with the karyotyping results, so MLPA can be used for confirmation of karyotyping (**Table 1**). In cases with atypical genitalia (SW or SV) and CAH cases with a male external phenotype, karyotype is the only way to prove the genetic sex (whether males or virilized

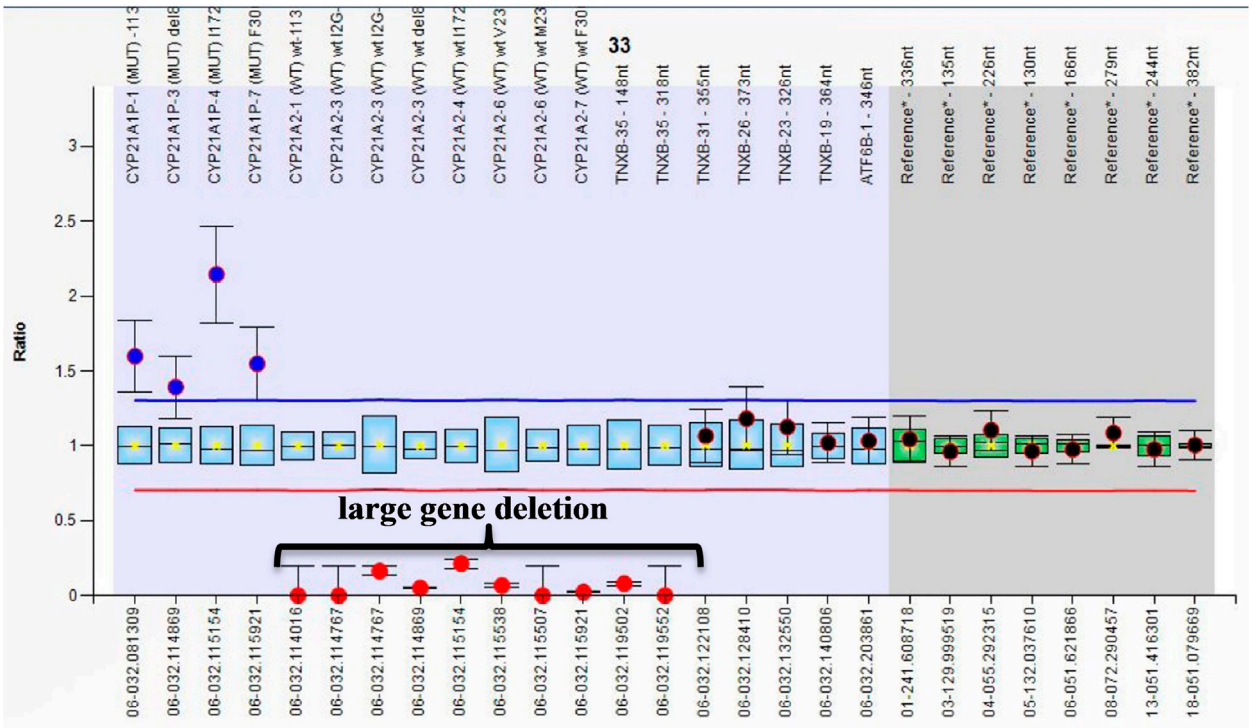



**FIGURE 2 |** Ratio chart showing homozygous large gene deletion from CYP21A2 extending to exon 35 of TNXB which produced chimeric gene TNXA/TNXB (CAH-X CH-1).

females). Ultrasound can detect the presence of Mullerian structures as a guide for the female internal genitalia. However, it is operator dependent, and sometimes, it can be very difficult to visualize the tiny uterus and ovaries in neonates.

Regarding genotyping, the most frequent variant in the present study was I2G, with an allelic frequency of 31%. The second common variant was large deletions (15.45%), and the third most frequent variant was LGC (7.3%) (**Table 2**). In 2002, Krone et al.'s study showed similar results (Krone et al., 2002). Copy number variants presented in 15% of cases as simple (11.6%) or combined with SNVs. These simple cases may be missed in genetic diagnosis when using commercial molecular techniques that only test for SNVs. Thus, copy number variations also need to be checked in the cohort due to the high frequency of this cohort (Gao et al., 2019).

New et al. highlighted the high frequency of homozygous I2G variants in the Middle Eastern population (New et al., 2013). The I2G mutation presented in 31% of cases alone, or combined with other point mutations (9%) or with copy number variations (LGC or deletions) (3.5%). The E6 cluster variants (p.V237E, p.M239K) showed an allele frequency of 0.86%. The homozygous I2G (as the common mutation) and compound heterozygous for I2G and E6 (the least frequent point mutation) mutations cause the same disease phenotype (SW). The genotype frequency among different phenotypes in the studied cases is summarized in **Supplementary Table S1**.

The lack of identified variants in the present study highlights the need for further genetic studies to detect such a discrepancy between the genotype and phenotype. This may be explained by the presence of common variants of the CYP21A2 gene that were not detected by MLPA probes of probemix P050-C1 CAH as a mutation, P.Q318X, in addition to rare variants that were not tested. Moreover, variants within the promoter region of the gene may be present.

Sequencing is essential to optimize molecular diagnosis in CAH patients. It is considered the only available method that allows for a 100% detection rate for all single nucleotide variants of CYP21A2, either common or rare. Direct sequencing in combination with MLPA offers the highest diagnostic information (Khattab et al., 2016). The gene dosage of the active gene and pseudogene reported in the current study using the MLPA technique showed the CYP21A2 copy number was zero in 24/112 (21.4%). Three copies were found in 7 of the 112 patients (6.3%). To our knowledge, no studies investigate the clinical importance of pseudogene copy number role in the CAH. Pseudogene dosage analysis is important in the differentiation between large deletions and LGC and to avoid misinterpretation of the wild-type sequence of a pseudogene as a duplication of CYP21A2. The present study results showed three copies of CYP21A1P in 21 out of 112 patients (18.8%) and four copies in 4 out of112 (3.6%).

An added diagnostic potential of using the MLPA technique is that the pseudogene analysis was performed simultaneously with CYP21A2 analysis, so this can help to detect the extent of deletion and different forms of chimeric genes. In the present study, a 30-KB Del-produced chimeric gene (CH-1) was detected in 6/112

cases and a chimeric gene (CH-7) in one case. Coeli-Lacchini et al. conducted a study using the MLPA method. The results revealed a 30-kB Del in 3/90 (3.3%) and a complete CYP21A2 deletion in one of the patients (Coeli-Lacchini et al., 2013).

Deletions typically extend approximately 30 kb and have their breakpoint somewhere between exons 3 and 8 of the CYP21A1P pseudogene and ending somewhere at the corresponding point in CYP21A2, thus yielding a single remaining gene in which the 5′ end corresponds to the CYP21A1P pseudogene and the 3′ end corresponds to the CYP21A2 gene. To date, nine types of CYP21A1P/CYP21A2 chimera (CH1–CH9) with different junction sites have been identified (New et al., 2013). Furthermore, loss of MLPA probes specific to the coding regions of the CYP21A2 gene, which constituted loss of exons 1–3, is Del 1–3, loss of exons 1–6 is Del 1–6, and loss of exons 1–8 is Del 1–8 variants in CYP21A2 (Balraj et al., 2013). The classic TNXA/TNXB chimeric gene is a 120 bp deletion at the boundary of exon 35 and intron 35 (CAH-X CH-1). The CYP21A2 gene is completely deleted and replaced by the CYP21A1P pseudogene (Merke et al., 2013). The 148 and 318 nucleotide (nt) probes of TNXB exon 35 of the MLPA probemix (P050-C1 CAH) are located within the 121 nt sequence that is absent in the TNXA pseudogene. Using the MLPA technique, the prevalence of large deletions and large gene conversions in the present study was found in 29/112 (25.9%) patients with the SW phenotype. The results were in agreement with the work of Ben Charfeddine et al., 2012, who reported gene deletion/conversion in 11/50 Tunisian patients (22%) (Kharrat et al., 2004; Ben Charfeddine et al., 2012; Chi et al., 2019; Umaña-Calderón et al., 2021).

The present study is the first to report large gene Del in Egyptian CAH patients, extending to exon 35 of the TNXB gene that produced a chimeric gene (CAH-X CH-1) in 3/112 patients. All cases with large gene deletions presented with the SW form (**Figure 2**). This molecular diagnosis coincides with Ehlers–Danlos syndrome (EDS). These results highlight the importance of a complete clinical evaluation for CAH children. Clinical evaluation for connective tissue dysplasia should be routinely performed in CAH patients, especially those harboring a CYP21A2 deletion (Merke et al., 2013,) as CAH-X recently considered a subtype of CAH (Lao and Merke, 2021; Miller, 2021). Merke et al. reported CAH-X syndrome in 7% of CAH patients. Patients with CAH-X were more likely to have joint hypermobility, chronic joint pain, multiple joint dislocations, and structural cardiac valves.

## CONCLUSION

Early genetic sex assignment, genetic diagnosis, and phenotype prediction are important in the early treatment of CAH patients. Due to the sophisticated mechanisms of genetic variations in the CYP21A2 gene, there is an urgent need to use MLPA as a multipotent and cost-effective laboratory routine tool for the molecular diagnosis of 21-OHD CAH in countries with a high frequency of CAH for proper diagnosis of cases. Recommendations for tailoring the panel of MLPA probe mix to include missed important variants will add to its potent diagnostic effect.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the clinical and chemical pathology ethical committee at Kasr al Ainy Hospital. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

## AUTHOR CONTRIBUTIONS

AT: substantial contributions to the conception or design of the work and acquisition and analysis of data. IM: drafting the work and revising it critically for important intellectual content. NM: diagnosing of patients and data collection. FE: substantial contributions to the conception or design of the work and acquisition of data. MH: substantial contributions to the conception or design of the work and acquisition of data. SA: drafting the work or revising it critically for important intellectual content. AI: data collection. HS: data collection. BL: provided approval for publication of the content. YE: drafting the work. TR: contribution in lab work. ME: substantial contributions to the conception and design of the work and acquisition, analysis, and interpretation of data for the work.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.785570/full#supplementary-material

## REFERENCES

Al-obaidi, R. G. Y., Al-musawi, B. M. S., Al-zubaidi, M. A. K., Oberkanins, C., Németh, S., and Al-obaidi, Y. G. Y. (2016). Molecular Analysis of CYP21A2 Gene Mutations Among Iraqi Patients with Congenital Adrenal Hyperplasia. *Enzyme Res.* 2016, 1–7. doi:10.1155/2016/9040616

Alshabab, L. I. S., Alebrahem, A., Kaddoura, A., and Al-Fahoum, S. (2015). Congenital Adrenal Hyperplasia Due to 21-hydroxylase Deficiency: A Five-Year Retrospective Study in the Children's Hospital of Damascus, Syria. *Qatar Med. J.* 2015, 11. doi:10.5339/qmj.2015.11

Balraj, P., Lim, P. G., Sidek, H., Wu, L. L., and Khoo, A. S. (2013). Mutational Characterization of Congenital Adrenal Hyperplasia Due to 21-hydroxylase Deficiency in Malaysia. *J. Endocrinol. Invest.* 36, 366–374. doi:10.3275/8648

Ben Charfeddine, I., Riepe, F. G., Clauser, E., Ayedi, A., Makni, S., Sfar, M. T., et al. (2012). Steroid 21-hydroxylase Gene Mutational Spectrum in 50 Tunisian Patients: Characterization of Three Novel Polymorphisms. *Gene* 507, 20–26. doi:10.1016/j.gene.2012.07.027

Blaskó, B., Bánlaki, Z., Gyapay, G., Pozsonyi, É., Sasvári-Székely, M., Rajczy, K., et al. (2009). Linkage Analysis of the C4A/C4B Copy Number Variation and Polymorphisms of the Adjacent Steroid 21-hydroxylase Gene in a Healthy Population. *Mol. Immunol.* 46, 2623–2629. doi:10.1016/j.molimm.2009.04.033

Chan, A. O. K., But, W. M., Ng, K. L., Wong, L. M., Lam, Y. Y., Tiu, S. C., et al. (2011). Molecular Analysis of Congenital Adrenal Hyperplasia Due to 21-hydroxylase Deficiency in Hong Kong Chinese Patients. *Steroids* 76, 1057–1062. doi:10.1016/j.steroids.2011.04.010

Chen, W., Xu, Z., Sullivan, A., Finkielstain, G. P., Van Ryzin, C., Merke, D. P., et al. (2012). Junction Site Analysis of Chimeric CYP21A1P/CYP21A2 Genes in 21-hydroxylase Deficiency. *Clin. Chem.* 58, 421–430. doi:10.1373/clinchem.2011.174037

Chi, D. V., Tran, T. H., Nguyen, D. H., Luong, L. H., Le, P. T., Ta, M. H., et al. (2019). Novel Variants of CYP21A2 in Vietnamese Patients with Congenital Adrenal Hyperplasia. *Mol. Genet. Genomic Med.* 7, e623. doi:10.1002/mgg3.623

Choi, J.-H., Kim, G.-H., and Yoo, H.-W. (2016). Recent Advances in Biochemical and Molecular Analysis of Congenital Adrenal Hyperplasia Due to 21-hydroxylase Deficiency. *Ann. Pediatr. Endocrinol. Metab.* 21, 1–6. doi:10.6065/apem.2016.21.1.1

Coeli-Lacchini, F. B., Turatti, W., Elias, P. C. L., Elias, L. L. K., Martinelli, C. E., Moreira, A. C., et al. (2013). A Rational, Non-radioactive Strategy for the Molecular Diagnosis of Congenital Adrenal Hyperplasia Due to 21-hydroxylase Deficiency. *Gene* 526, 239–245. doi:10.1016/j.gene.2013.03.082

Concolino, P., Mello, E., Toscano, V., Ameglio, F., Zuppi, C., and Capoluongo, E. (2009). Multiplex Ligation-dependent Probe Amplification (MLPA) Assay for the Detection of CYP21A2 Gene Deletions/duplications in Congenital Adrenal Hyperplasia: First Technical Report. *Clinica Chim. Acta* 402, 164–170. doi:10.1016/j.cca.2009.01.008

Daae, E., Feragen, K. B., Nermoen, I., and Falhammar, H. (2018). Psychological Adjustment, Quality of Life, and Self-Perceptions of Reproductive Health in Males with Congenital Adrenal Hyperplasia: a Systematic Review. *Endocrine* 62, 3–13. doi:10.1007/s12020-018-1723-0

den Dunnen, J. T., Dalgleish, R., Maglott, D. R., Hart, R. K., Greenblatt, M. S., McGowan-Jordan, J., et al. HGVS Recommendations for the Description of Sequence Variants: 2016 Update. *Hum. Mutat.* 2016;37:564–569. doi:10.1002/humu.22981

Erlandson, A., Appelqvist, F., and Enerbäck, C. (2008). Epigenetic Mutations in CDKN2A in Western Swedish Families with Hereditary Malignant Melanoma. *Mol. Med. Rep.* 1, 89–91.

Freire, A. V., Grinspon, R. P., and Rey, R. A. (2018). Importance of Serum Testicular Protein Hormone Measurement in the Assessment of Disorders of Sex Development. *Sex. Dev.* 12, 30–40. doi:10.1159/000479572

Gao, Y. J., Yu, B. Q., Lu, L., Wu, X. Y., Mao, J. F., Wang, X., et al. (2019). Analysis of Copy Number Variation of CYP21A2 Gene and the Type of CYP21A1P/CYP21A2 Fused Gene in Patients with 21-hydroxylase Deficiency. *Zhonghua Yi Xue Za Zhi* 99, 3765–3769. doi:10.3760/cma.j.issn.0376-2491.2019.48.002

Guerrero-Fernández, J., Azcona San Julián, C., Barreiro Conde, J., Bermúdez de la Vega, J. A., Carcavilla Urquí, A., Castaño González, L. A., et al. (2018). Guía de actuación en las anomalías de la diferenciación sexual (ADS)/desarrollo sexual diferente (DSD). *Anales de Pediatría* 89, e1–315. doi:10.1016/j.anpedi.2018.06.009

Huynh, T., McGown, I., Cowley, D., Nyunt, O., Leong, G. M., Harris, M., et al. (2009). The Clinical and Biochemical Spectrum of Congenital Adrenal Hyperplasia Secondary to 21-hydroxylase Deficiency. *Clin. Biochem. Rev.* 30, 75–86.

Kharrat, M., Tardy, V., M'Rad, R., Maazoul, F., Jemaa, L. B., Refaï, M., et al. (2004). Molecular Genetic Analysis of Tunisian Patients with a Classic Form of 21-hydroxylase Deficiency: Identification of Four Novel Mutations and High Prevalence of Q318X Mutation. *J. Clin. Endocrinol. Metab.* 89, 368–374. doi:10.1210/jc.2003-031056

Khattab, A., Yuen, T., Al-Malki, S., Yau, M., Kazmi, D., Sun, L., et al. (2016). A rareCYP21A2mutation in a Congenital Adrenal Hyperplasia kindred Displaying Genotype-Phenotype Nonconcordance. *Ann. N.Y. Acad. Sci.* 1364, 5–10. doi:10.1111/nyas.12864

Krone, N., Braun, A., Weinert, S., Peter, M., Roscher, A. A., Partsch, C.-J., et al. (2002). Multiplex Minisequencing of the 21-hydroxylase Gene as a Rapid Strategy to Confirm Congenital Adrenal Hyperplasia. *Clin. Chem.* 48, 818–825. doi:10.1093/clinchem/48.6.818

Kutney, K., Konczal, L., Kaminski, B., and Uli, N. Challenges in the Diagnosis and Management of Disorders of Sex Development. *Birth Defects Res. C Embryo Today* 2016;108:293–308. doi:10.1002/bdrc.21147

Lao, Q., and Merke, D. P. (2021). Molecular Genetic Testing of Congenital Adrenal Hyperplasia Due to 21-hydroxylase Deficiency Should Include CAH-X Chimeras. *Eur. J. Hum. Genet.* 29, 1047–1048. doi:10.1038/s41431-021-00870-5

Lee, H. J., Choi, J., Hwang, T. S., Shong, Y. K., Hong, S. J., and Gong, G. (2010). Detection ofBRAFMutations in Thyroid Nodules by Allele-specific PCR Using a Dual Priming Oligonucleotide System. *Am. J. Clin. Pathol.* 133, 802–808. doi:10.1309/AJCPO3F2ENKMDTUS

Merke, D. P., Chen, W., Morissette, R., Xu, Z., Van Ryzin, C., Sachdev, V., et al. (2013). Tenascin-X Haploinsufficiency Associated with Ehlers-Danlos Syndrome in Patients with Congenital Adrenal Hyperplasia. *J. Clin. Endocrinol. Metab.* 98, E379–E387. doi:10.1210/jc.2012-3148

Miller, W. L. (2021). Tenascin-X-Discovery and Early Research. *Front. Immunol.* 11, 612497. doi:10.3389/fimmu.2020.612497

Musa, N., Asem, N., Basyony, S., and Fawaz, L. (2020). Assessment of Health-Related Quality of Life in Egyptian Children and Adolescents with Congenital Adrenal Hyperplasia. *J. Pediatr. Endocrinol. Metab.* 33, 295–304. doi:10.1515/jpem-2019-0345

New, M. I., Abraham, M., Gonzalez, B., Dumic, M., Razzaghy-Azar, M., Chitayat, D., et al. (2013). Genotype-phenotype Correlation in 1,507 Families with Congenital Adrenal Hyperplasia Owing to 21-hydroxylase Deficiency. *Proc. Natl. Acad. Sci.* 110, 2611–2616. doi:10.1073/pnas.1300057110

Nimkarn, S., Gangishetti, P. K., Yau, M., and New, M. I. (1993). in *21-Hydroxylase-Deficient Congenital Adrenal Hyperplasia*. M. P. Adam, H. H. Ardinger, R. A. Pagon, S. E. Wallace, L. J. H. Bean, and K. Stephens. Editors (Seattle (WA.

Rabbani, B., Mahdieh, N., Haghi Ashtiani, M. T., Akbari, M. T., and Rabbani, A. (2011). Molecular Diagnosis of Congenital Adrenal Hyperplasia in Iran: Focusing on CYP21A2 Gene. *Iran J. Pediatr.* 21, 139–150.

Savanelli, A., Alicchio, F., Esposito, C., De Marco, M., and Settimi, A. (2008). A Modified Approach for Feminizing Genitoplasty. *World J. Urol.* 26, 517–520. doi:10.1007/s00345-008-0298-4

Speiser, P. W., Arlt, W., Auchus, R. J., Baskin, L. S., Conway, G. S., Merke, D. P., et al. (2018). Congenital Adrenal Hyperplasia Due to Steroid 21-Hydroxylase Deficiency: An Endocrine Society* Clinical Practice Guideline. *J. Clin. Endocrinol. Metab.* 103, 4043–4088. doi:10.1210/jc.2018-01865

Tayel, S. M., Ismael, H., Kandil, H., Abd Rabuh, A. R., and Sallam, H. (2011). Congenital Adrenal Hyperplasia in Alexandria, Egypt: A High Prevalence Justifying the Need for a Community-Based Newborn Screening Program. *J. Trop. Pediatr.* 57, 232–234. doi:10.1093/tropej/fmq064

Temtamy, S., and Aglan, M. (2012). Consanguinity and Genetic Disorders in Egypt. *Middle East. J. Med. Genet.* 1. doi:10.1097/01.mxe.0000407744.14663.d8

Toraman, B., Ökten, A., Kalay, E., Karagüzel, G., Dinçer, T., Açıkgöz, E. G., et al. (2013). Investigation of CYP21A2 Mutations in Turkish Patients with 21-hydroxylase Deficiency and a Novel Founder Mutation. *Gene* 513, 202–208. doi:10.1016/j.gene.2012.10.059

Tsai, M.-C., Chou, Y.-Y., Lin, S.-J., and Tsai, L.-P. (2012). A Novel SRD5A2 Mutation in a Taiwanese Newborn with Ambiguous Genitalia. *Kaohsiung J. Med. Sci.* 28, 231–235. doi:10.1016/j.kjms.2011.10.011

Umaña-Calderón, A., Acuña-Navas, M. J., Alvarado, D., Jiménez, M., and Cavallo-Aita, F. (2021). CYP21A2 Mutations in Pediatric Patients with Congenital Adrenal Hyperplasia in Costa Rica. *Mol. Genet. Metab. Rep.* 27, 100728. doi:10.1016/j.ymgmr.2021.100728

Veldhuisen, B., van der Schoot, C. E., and de Haas, M. (2015). Multiplex Ligation-dependent Probe Amplification (MLPA) Assay for Blood Group Genotyping, Copy Number Quantification, and Analysis of RH Variants. *Immunohematology* 31, 58–61.

Vrzalová, Z., Hrubá, Z., Hrabincová, E. S., Vrábelová, S., Votava, F., Kolouškova, S., et al. (2011). Chimeric CYP21A1P/CYP21A2 Genes Identified in Czech Patients with Congenital Adrenal Hyperplasia. *Eur. J. Med. Genet.* 54, 112–117. doi:10.1016/j.ejmg.2010.10.005

Wisniewski, A. B. (2017). Psychosocial Implications of Disorders of Sex Development Treatment for Parents. *Curr. Opin. Urol.* 27, 11–13. doi:10.1097/MOU.0000000000000344

Xu, Z., Chen, W., Merke, D. P., and McDonnell, N. B. (2013). Comprehensive Mutation Analysis of the CYP21A2 Gene. *J. Mol. Diagn.* 15, 745–753. doi:10.1016/j.jmoldx.2013.06.001

**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.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

*Copyright © 2022 Tolba, Mandour, Musa, Elmougy, Hafez, Abdelatty, Ibrahim, Soliman, Labib, Elshiwy, Ramzy and Elsharkawy. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*

frontiers | Frontiers in Genetics

ORIGINAL RESEARCH
published: 11 August 2022
doi: 10.3389/fgene.2022.875182

# The Usefulness of a Targeted Next Generation Sequencing Gene Panel in Providing Molecular Diagnosis to Patients With a Broad Spectrum of Neurodevelopmental Disorders

*Simona Mellone[1], Chiara Puricelli[1,2], Denise Vurchio[2], Sara Ronzani[1], Simone Favini[1], Arianna Maruzzi[2], Cinzia Peruzzi[3], Amanda Papa[4], Alice Spano[1], Fabio Sirchia[5], Giorgia Mandrile[6], Alessandra Pelle[6], Paolo Rasmini[7], Fabiana Vercellino[8], Andrea Zonta[6], Ivana Rabbone[2,9], Umberto Dianzani[1,2], Maurizio Viri[4] and Mara Giordano[1,2]**

[1]Laboratory of Genetics, Clinical Biochemistry, University Hospital Maggiore della Carità, Novara, Italy, [2]Department of Health Sciences, Università del Piemonte Orientale, Novara, Italy, [3]Child Neuropsychiatry Unit, Ospedale San Gerardo Monza-Università degli Studi di Milano Bicocca, Monza, Italy, [4]Department of Child Neuropsychiatry, Hospital Maggiore della Carità, Novara, Italy, [5]Department of Molecular Medicine, University of Pavia, Pavia, Italy, [6]Medical Genetics Unit, Città della Salute e della Scienza University Hospital, Torino, Italy, [7]Child Neuropsychiatry ASL, Vercelli, Italy, [8]Child Neuropsychiatry Unit, SS. Antonio e Biagio e Cesare Arrigo Hospital, Alessandria, Italy, [9]Division of Pediatrics, University Hospital Maggiore della Carità, Novara, Italy

OPEN ACCESS

**Edited by:**
Kathleen M. Gorman,
Temple Street Children's University Hospital, Ireland

**Reviewed by:**
Ivan Iossifov,
Cold Spring Harbor Laboratory, United States
Olivia J. Veatch,
University of Kansas Medical Center, United States
M. Chiara Manzini,
The State University of New Jersey, United States

***Correspondence:**
Mara Giordano
mara.giordano@med.uniupo.it

**Specialty section:**
This article was submitted to Genetics of Common and Rare Diseases, a section of the journal Frontiers in Genetics

**Received:** 13 February 2022
**Accepted:** 26 May 2022
**Published:** 11 August 2022

**Citation:**
Mellone S, Puricelli C, Vurchio D, Ronzani S, Favini S, Maruzzi A, Peruzzi C, Papa A, Spano A, Sirchia F, Mandrile G, Pelle A, Rasmini P, Vercellino F, Zonta A, Rabbone I, Dianzani U, Viri M and Giordano M (2022) The Usefulness of a Targeted Next Generation Sequencing Gene Panel in Providing Molecular Diagnosis to Patients With a Broad Spectrum of Neurodevelopmental Disorders. Front. Genet. 13:875182. doi: 10.3389/fgene.2022.875182

**Background:** Neurodevelopmental disorders comprise a clinically and genetically heterogeneous group of conditions that affect 2%–5% of children and represents a public health challenge due to complexity of the etiology. Only few patients with unexplained syndromic and non-syndromic NDDs receive a diagnosis through first-tier genetic tests as array-CGH and the search for *FMR1* CGG expansion. The aim of this study was to evaluate the clinical performance of a targeted next-generation sequencing (NGS) gene panel as a second-tier test in a group of undiagnosed patients with NDDs.

**Method:** A 221-gene next-generation sequencing custom panel was designed and used to analyze a cohort of 338 patients with a broad spectrum of NDDs (202 males and 136 females) including Intellectual Disability (ID), Autism Spectrum Disorders (ASD), Epilepsy, language and motor disorders.

**Results:** A molecular diagnosis was established in 71 patients (21%) and a *de novo* origin was present in 38 (64.4%) of the available trios. The diagnostic yield was significantly higher in females than in males (29.4% vs. 15.3%; $p = 0.0019$) in particular in ASD (36.8% vs. 7.6%; $p = 0.0026$) and Epilepsy (38.9% vs. 14.4% $p = 0.001$). The most involved genes were *SLC2A1*, *SCN1A, ANKRD11*, *ATP1A2*, *CACNA1A*, *FOXP1*, and *GNAS* altered in more than two patients and accounting for the 19.7% of the diagnosis.

**Conclusion:** Our findings showed that this NGS panel represents a powerful and affordable clinical tool, significantly increasing the diagnostic yield in patients with different form of NDDs in a cost- and time-effective manner without the need of large investments in data storage and bioinformatic analysis.

**Keywords: neurodevelopmental disorders, NGS gene panel, autism, intellectual disability, epilepsy**

## INTRODUCTION

Neurodevelopmental disorders (NNDs) represent a group of heterogeneous early-onset conditions including autism spectrum disorder (ASD), intellectual disability (ID), language disorders, developmental delay (DD), and epilepsy, that globally affect 2%–5% of children (Boyle et al., 2011; American Psychiatric Association, 2013).

Different phenotypes frequently co-exist in the same patient, thus blurring the lines in the classification of individuals with these disorders. Intellectual disability (ID) is the most common NDD with a prevalence varying between 0.5 and 3% in the general population, characterized by deficits in both intellectual and adaptive functioning that first manifest during early childhood (Leonard et al., 2011). Children with intellectual disability (ID) exhibit an increased risk to present potential co-occurring developmental disorders. Epilepsy, which affects 0.7%–1.0% of the general population, represents a common comorbidity in the ID populations. Estimates of the prevalence of epilepsy in ID patients show that it increases with ID severity, varying from 16.1% to more than 50% (Robertson et al., 2015). ASD represents a broad clinical condition including deficits in social communication and interactions, repetitive behaviors, and restrictive interests (Totsika et al., 2011). As well as epilepsy ASD frequently coexists with ID: it is estimated that 40%–70% of ASD individuals suffers also from ID (Bertrand et al., 2001; Chakrabarti and Fombonne, 2001; Charman et al., 2011; Almuhtaseb et al., 2014; Jensen and Girirajan, 2017; Kazeminasab et al., 2018).

Multiple causes have been correlated to NDDs, including genetic, traumatic, and environmental factors that largely interact with each other (Van Bokhoven, 2011), and although the precise etiology remains unknown in most cases, the genetic component has been deciphered in a broad range of clinical phenotypes also thanks to ever increasingly sophisticated diagnostic tools.

The usefulness of a molecular diagnosis mainly relies on the possibility of giving information on the clinical outcome, preventing further superfluous testing, and proceeding with active monitoring of the patient. Furthermore, knowing the genetic cause of a disorder may disclose the involved biological pathway and, in some cases, suggests tailored pharmacological and behavioral treatments (Henderson et al., 2014). The current gold-standard tests for the molecular diagnosis of this heterogeneous group of conditions are mostly based on unbiased genomic approaches such as array comparative genomic hybridization (array-CGH), recommended as the first-tier genetic test (Di Gregorio et al., 2017), followed by multiple gene panels and whole-exome sequencing (WES), that are currently considered as the second-tier tests. In particular, large-scale genomic studies (WES and Whole genome sequencing, WGS) have established the pathogenic role of *de novo* and recessive variations in different forms of NDDs (Doan et al., 2019).

Despite its high performance, genetic testing is still not implemented routinely in clinical practice for patients with NDDs, mainly because of a scarcity of instrumental resources and trained specialized staff. The WES approach offers a comprehensive view of the entire mutational landscape, but it generates a huge amount of data making bioinformatics challenging especially in the diagnostic workflow. Although genome sequencing is expected to revolutionize diagnostics in the near future, gene panel sequencing remain crucial for an efficient and precise identification of pathogenic variants representing a more affordable second-tier test that, although limited to specific sets of genes, allows accurate identification of variants with greater sensitivity and lower overall cost compared to WES and WGS. In a recent meta-analysis, Stefanski et al. (2021) examined several cohorts with different forms of NDDs that had been screened either through a targeted gene panel ($N$ = 73) or WES ($N$ = 36). The diagnostic yield of WES was higher but the difference was not statistically significant (27.2% vs. 22.6%, $p$ = 0.071) (Stefanski et al., 2021). Therefore, to date, targeted sequencing might represent an efficient approach achievable in clinical contexts, provided that well-conceived gene panels are used.

The aim of the present study was to evaluate the clinical usefulness of a 221-gene custom panel in a cohort of 338 individuals with a broad spectrum of NDDs. The presence of clinically relevant variants was detected overall in 21% of the individuals (with a significantly higher yield in females) suggesting that it might represent a powerful tool in the routine diagnostic workflow of these conditions.

## PATIENTS AND METHODS

### Patients

The study retrospectively included all the 338 patients with NNDs that were referred to the Laboratory of Genetics of the University Hospital "Maggiore della Carità" in Novara (Italy) from January 2017 to May 2021 for genetic testing. All patients had been already screened by aCGH, resulting either negative or carrying a variant of uncertain significance (VUS), and were tested for *FMR1* CGG expansion to exclude Fragile X syndrome. Clinical data were obtained from the patients' clinical charts provided by the unit or the physician that originally requested genetic testing. Information about the clinical presentation, cognitive performance scores, family history, and, when available, past diagnostic tests was used as the main source to characterize and stratify the subjects. All requests were accompanied by written informed consent obtained from the patients' parents or legal representatives and written consent was requested for variant segregation analysis. Most patients were referred to the Genetics Unit from the Child Neuropsychiatry Unit ($N$ = 264, 78.1%) and the others from the Pediatric Unit ($N$ = 7, 2.1%) and the Neonatal Intensive Care Unit ($N$ = 2, 0.6%) of the Hospital "Maggiore della Carità" of Novara. Moreover, since this Hospital is a tertiary reference center for specialized genetic tests and the hub center of the laboratory network in the North Eastern Piedmont quadrant for non-urgent tests, requests were also received from other hospitals in Piedmont ($N$ = 37, 10.9%) and other Italian regions ($N$ = 20, 5.9%).

## Gene Panel Design

The panel included 221 genes (**Supplementary Table S1**) found to be altered in NDDs, selected from OMIM (https://www.omim.org), medical literature (https://pubmed.ncbi.nlm.nih.gov/), the public resources SFARI (https://sfari.org/resources/sfari-gene) and Genomic England PanelApp (https://panelapp.genomicsengland.co.uk/). Both well-known disease-causing genes and a sub-group of a few candidate NDD genes were included; 106 genes are associated with autosomal dominant (AD) disorders, 30 to autosomal recessive (AR) disorders; 11 genes have both an AD and an AR inheritance, and 66 genes have been associated with X-linked (XL) diseases. For eight genes, the inheritance pattern is still controversial (**Supplementary Table S1**).

The genes were stratified into nine functional categories according to their main functional properties, with many involved in several functions: 1. Cell cycle regulation; 2. Cell structure and polarity; 3. Genomic stability/DNA repair; 4. Membrane polarity/electrochemical gradient; 5. Metabolic pathways; 6. Brain function/development and neuronal signaling; 7. Transcriptional/translational regulation and cell differentiation; 8. Intra- and intercellular signal transduction; 9. Vesicular trafficking. Information about their function was obtained from the scientific literature (https://pubmed.ncbi.nlm.nih.gov/), online databases including OMIM (https://www.omim.org) and GeneCards (https://www.genecards.org) (**Supplementary Table S1**).

## Targeted Next-Generation Sequencing

DNA was isolated from 200 µl of peripheral blood using ReliaPrepTM Blood gDNA Miniprep System (Promega, Fitchburg, WI, United States), according to the manufacturer's protocols. The amount of DNA obtained from each sample was quantified using NanoDrop One (Thermo Fisher Scientific, Wilmington, DE, United States). To create the panel, the SureDesign software (Agilent Technologies, Santa Clara, CA, United States) was utilized and the size corresponded to 1.342 Mb. DNA libraries were prepared by SureSelectQXT Target Enrichment Kit according to the protocol for Illumina Multiplex Sequencing (Agilent Technologies, Santa Clara, CA, United States). Sequencing probes were designed to cover all coding exons ±20 bp flanking sequences from the intron-exon boundary. Sequencing baits were designed with 2X density so that each desired region was covered by at least four overlapping probes. DNA libraries were diluted to 12 pmol/l pools with 11 samples analyzed in parallel per one MiSeq sequencing run using a MiSeq sequencing reagent kit v3 150 cycles (Illumina, Inc., San Diego, CA, United States) to obtain an estimated coverage of 150X. In each sample, the estimated coverage exceeded 50 reads for over 96% of the target gene sequence.

Read alignment to the human genome reference (hg19/GRCh37), variant calling, and annotations of genetic variants were performed with the SureCall v3.5 software (Agilent Technologies, Santa Clara, CA, United States) and wANNOVAR (https://wannovar.wglab.org/) was used to annotate them. Variants showing low read depth (<20x) or poor base quality score (Phred quality score <20) were excluded.

Variant prioritization was performed using the following bioinformatic criteria (**Figure 1**): 1) non-synonymous SNV or indels located in exonic or splicing regions; 2) all synonymous variants not predicted to alter the splicing mechanism were excluded; 3) allele frequency <0.01 and no homozygotes or hemizygotes if recessive in public sequence in the following databases: 1,000 Genome Project (http://browser.1000genomes.org), Exome Aggregation Consortium (ExAC) (http://exac.broadinstitute.org), Genome Aggregation Database (gnomAD) (https://gnomad.broadinstitute.org/), dbSNP (http://www.ncbi.nlm.nih.gov/projects/SNP/); 4) variant predicted to be damaging by at least four of the *in silico* predictive tools (SIFT, Polyphen-2, Mutation Taster, CADD, PROVEAN, MutationAssessor, FATHMM). Then, a literature search in PubMed (https://pubmed.ncbi.nlm.nih.gov/) was performed to investigate the role of the involved genes and to compare the phenotypes of reported patients carrying variants in the same genes (**Figure 1**).

Evaluation of the clinical significance of the variants was performed following the consensus guidelines published by the American College of Medical Genetics and Genomics (ACMG; https://www.acmg.net), by using the Varsome tool [https://varsome.com/, (Kopanos et al., 2019),] and by consulting the ClinVar database (https://www.ncbi.nlm.nih.gov/clinvar/). Accordingly variants were initially classified into five classes: "pathogenic," "likely pathogenic," of "uncertain significance," "likely benign," and "benign," based on multiple lines of evidence (conservation, allele frequency in population databases, computational inferences of variant effect). In most cases, pathogenic (P) and likely pathogenic (LP) variants were definitively considered as causative after segregation analysis when parental DNA was available and after evaluating the consistency with the clinical phenotype. In essence, P/LP variants that either arose *de novo* or were inherited from an affected parent or were located within genes causing disease consistent with the patient's phenotype were considered as the major contributors to the patient's clinical phenotype.

## Sanger Sequencing

P and LP variants and VUS were subsequently confirmed by Sanger sequencing, which was also used to evaluate familial segregation when the parents were available. Sanger sequencing was performed using dye-terminator chemistry with the kit BigDyeTM Terminator v.1.1 Cycle Sequencing RR-100 (Applied Biosystems, Unuted States) and run on automated sequencer "SeqStudio Genetic Analyzer" (Applied Biosystems, United States). Primers were specifically designed by using the software "Primer3web" v 4.1.0 (https://primer3.ut.ee/).

## Analysis of VPS13B Through Multiplex Ligation Dependent Probe Amplification (MLPA)

Search for deletions/duplications in the *VPS13B* gene was performed by an MLPA assay using the commercial Kits SALSA MLPA Probemix P321-B3 VPS13B mix 1 and SALSA MLPA Probemix P322-C2 VPS13B mix 2 (MRC-Holland, Amsterdam, Netherlands) following the manufacturer's instructions.

## *In-Vitro* SCN1A Minigene Splicing Assay

Fragments carrying the wild-type or mutant *SCN1A* exon 9 flanked by 338 bp of the 3′region intron 8 and 168 bp of the 5′region of

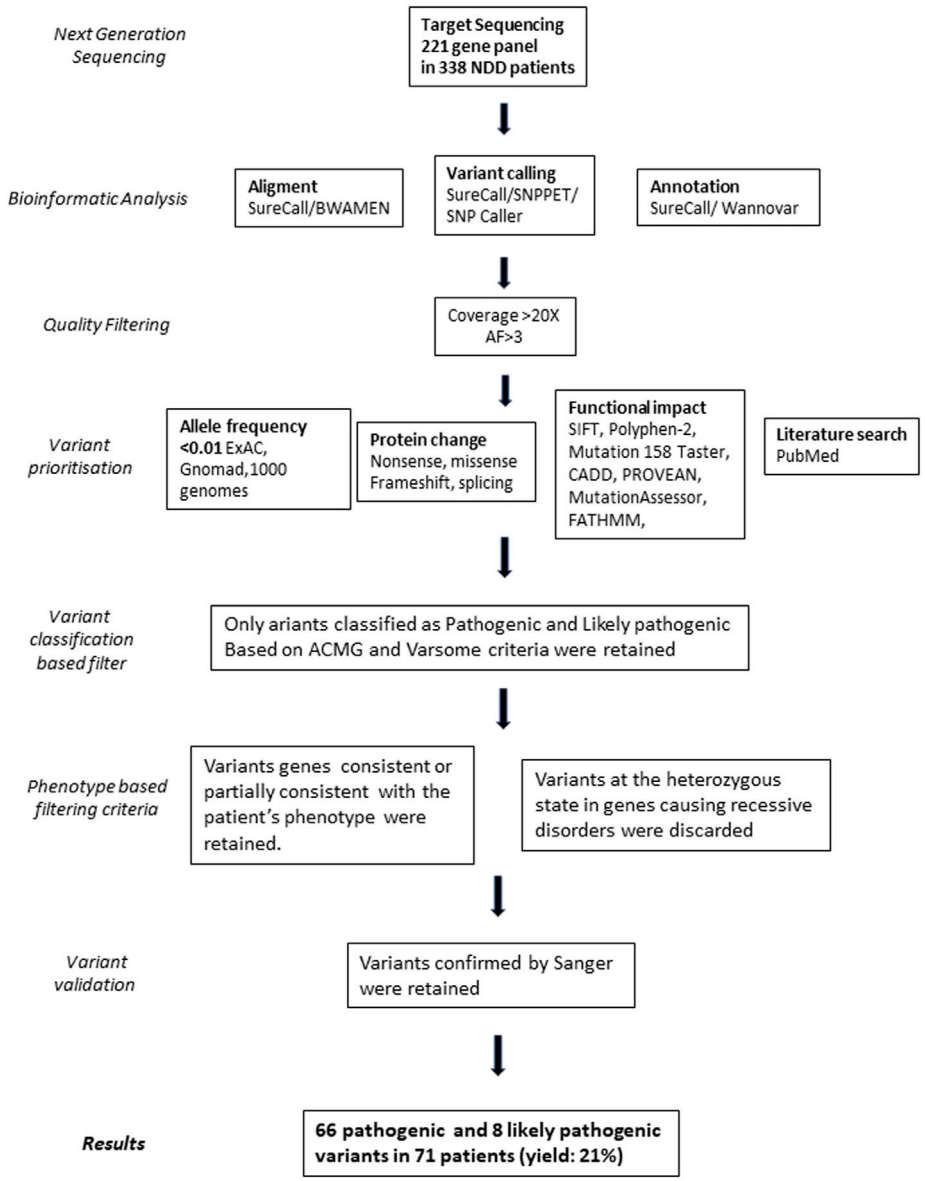



**FIGURE 1 |** Workflow describing the main step in the identification of the selected variants. AF, Allele frequency of the variant base relative to the read number.

intron 9 were amplified and cloned into the pSPL3 vector between the exons SD (Splice Donor) and SA (Splice acceptor) using SacI and BamHI restriction sites as previously described (Babu et al., 2021). U2O-S cells ($3 \times 10^5$) were seeded in a 6-well culture plate and incubated at 37°C in a 5% $CO_2$ atmosphere in Dulbecco's modified Eagle medium supplemented with 10% fetal bovine serum (Gibco-BRL, Carlsbad, CA, United States). The following day, 2 μg of wild-type or mutant vectors were transfected using Lipofectamine 2000 transfection reagent (Life Technologies, United States). cDNA was synthesized from 1 μg of RNA by the High Capacity cDNA Reverse Transcription kit (Applied Biosystems, United States), according to the manufacturer's instructions. Using vector exon-specific primers, cDNAs produced from the mini-gene constructs were specifically PCR amplified and Sanger sequenced.

# RESULTS

## Cohort Description

Three-hundreds-thirty-eight individuals (202 males and 136 females) with a broad spectrum of NDDs (ID, ASD, motor and/or language disorders, and dysmorphic features potentially due to a syndromic condition) were included in the study (**Table 1**). Epilepsy and ASD, as well as other clinical features, were present both as a comorbidity in subjects with ID and as conditions in subjects with normal cognitive function.

The majority of subjects were children (69.2% in the 1–12 age range) and the median age at the time of first clinical presentation was 7.0. The cohort also included 18 newborns (<12 months) and 25 young adults, so the final age ranged

**TABLE 1 |** Clinical characteristics of the 338 subjects and diagnostic yield.

| Characteristic | Patients<br>Absolute count<br>(percentage over total patients) | Patients carrying variants<br>Absolute count<br>(percentage over total patients in each category) |
|---|---|---|
| **Total count** | 338 | 71 (21.0) |
| **Gender** | | |
| Males | 202 (59.8) | 31 (15.3) |
| Females | 136 (40.2) | 40 (29.4) |
| **Age at diagnosis**(years) | | |
| Median | 7.0 | 10.0 |
| Range | 0-40 | 0-40 |
| < 1 years old | 18 (5.3) | 4(22.2) |
| 1 –12 years old | 234 (69.2) | 46 (19.7) |
| 13 –19 years old | 61 (18.0) | 12 (19.7) |
| ≥ 20 years old | 25 (7.4) | 9 (36.0) |
| **ID** | | |
| Total | 222 (65.7) | 49 (22.1) |
| Pure ID[a] | 44 (13.0) | 7 (15.9) |
| **Comorbidities associated with ID** | | |
| ASD | 28 (8.3) | 5 (17.9) |
| Motor disorder | 123 (36.4) | 26 (21.1) |
| Language disorder | 109 (32.2) | 23 (21.1) |
| Psychiatric disorder | 13 (3.8) | 8 (61.5) |
| Attention/Learning/Behavioral disorder | 59 (17.5) | 14 (23.7) |
| Epilepsy | 73 (21.6) | 19 (26.0) |
| Dysmorphism/Congenital defects | 74 (21.9) | 23(31.1) |
| **Other diagnoses without ID**[2] | | |
| ASD | 43 (12.7) | 6 (14.0) |
| Motor disorder | 24 (7.1) | 8 (33.3) |
| Language disorder | 27 (8.0) | 5 (18.5) |
| Psychiatric disorder | 2 (0.6) | 0 (0.0) |
| Attention/Learning/<br>Behavioral disorder | 22 (6.5) | 4 (18.2) |
| Pure epilepsy | 35(10.4) | 7 (20.0) |
| Dysmorphism/Congenital defects | 20 (5.9) | 1 (5.0) |
| **Copy number variants** | 36 (10.7) | 10 (27.8) |

*[a]Pure ID, refers to a diagnosis of intellectual disability in the absence of distinct comorbidities such as ASD, attention/learning/behavioral disorders, epilepsy, psychiatric disorders, and dysmorphic features. However, considering the high prevalence of motor and language impairment in subjects with ID, these two clinical features were included in the definition.*
*[b]Each additional diagnosis other than intellectual disability is considered including a potential overlap among the clinical manifestations, so that some patients may have, for instance, ASD, and epilepsy or ASD, and a learning disorder at the same time. Conversely, pure epilepsy is intended here as an isolated clinical entity, i.e. it refers to the presence of symptomatic chronic seizures in the absence of ID, motor or language disorders; ASD, attention/learning/behavioral disorders, psychiatric disorders, or dysmorphic traits.*
*Abbreviations: ASD, autism spectrum disorder; ID, intellectual disability; LP, likely pathogenic; P, pathogenic.*

from 0 to 40. ID was diagnosed in 222 individuals (65.7%), based on clinical assessment and objective performance scales, such as Wechsler Preschool and Primary Scale of Intelligence-III (WPPSI-III), Griffiths Mental Development Scales (GMDS), Leiter-R scale, Vineland-II scale, or Brunet-Lézine scale. A total of 116 individuals had no ID but were affected by one or more of other conditions, in particular, ASD was present in 43 subjects, motor disorders in 24, language disorders in 27; attention deficit-hyperactivity disorder (ADHD), learning difficulties, and behavioral alterations, included in the same category were observed in 22 subjects; 35 individuals presented epilepsy without any other comorbidity; dysmorphism/congenital defects were present in 20 and psychiatric disorders in two patients (**Table 1**). Positive family history was referred in 32 cases and in 12 of them the heredity was compatible with an autosomal dominant/recessive disorder. It has to be considered that this might be an underestimation since it is based on the retrospective records obtained by clinicians regarding the parents and the information reported during counseling where in some cases only one accompanying family member was present and in other cases, the presence of NND signs in the relatives might not have been evaluated.

An exceeding number of males was present in the whole NND cohort, with a male-to-female ratio of 1.5. This male excess was particularly evident in ASD (with or without ID), with a prevalence in males of 25.7% compared to 13.9% of females ($p = 0.01$).

In all patients, P and LP copy number variants (CNV) were ruled out, as well as the expansion at the Fragile X locus. In 36 (10.7%) patients, aCGH had revealed a VUS that was not considered the main cause of the disorder.

## Identification of P/LP Variants

All patients were analyzed with the 221 gene NGS panel (**Supplementary Table S1**). A manual prioritization procedure based on expert knowledge related to the disease phenotype and gene functions detected P ($N = 66$) and LP variants ($N = 8$) in 71 patients corresponding to 21% of the whole cohort (**Supplementary Table S2**; **Figure 1**). The clinical characteristics available for patients carrying P/LP are detailed in **Supplementary Table S3**. Fifty-five patients were heterozygous for a variant causing AD in a single gene and one patient for two in different genes; 3 patients were compound heterozygous/homozygous for biallelic variants causing autosomal recessive disorders (#8836, #559-19, #10456); 12 patients carried variants in X-linked genes (2 males and 10 females).

To assign pathogenicity, the variants were analyzed through Varsome (https://varsome.com/), a tool for human genetic variation analysis that follows the ACMG guidelines (Kopanos et al., 2019), and the verdict for each variant was adjusted according to the variant origin and to the correlation with the clinical phenotype (PS2 and PP4 Varsome criteria, respectively). Only variants identified in patients with a clinical phenotype consistent or partially consistent with that reported in subjects with alterations in the same genes were considered. By contrast, 315 variants classified as VUS (data not shown) based on the current knowledge are not considered in the present study although they are constantly revisited.

Despite the predominance of males in our whole cohort, a significantly higher number of disease contributing variants were identified in females (29.4%) than in males (15.3%; $p = 0.0019$; **Table 2**).

Segregation analysis was performed in 60 out of the 71 cases, for whom the DNA of both parents was available, and a *de novo* origin could be established in 38 (64.4%) of the analyzed trios. Of the 13 autosomal variants involved in dominant disorders and resulting transmitted, 10 were of maternal origin and 3 of paternal origin, with a gender transmission ratio of 3.3:1 (F: M). Both males carrying X-linked P/LP variations (#10851 and #1002-19) inherited the variant from the heterozygous unaffected mother. Among the 10 females showing X-linked variants, seven carried a variant that arose *de novo* and in one case the origin could not be assessed (**Table 2**).

In the whole cohort, a total of 22 P/LP variants causing well-known recessive disorders were identified at the heterozygous state (data not shown). Sixteen of them were not considered correlated to the patient's disorder owing to the lack of expected clinical features related to the involved gene (data not shown). Five patients carried monoallelic P/LP variants in *VPS13B*, mapping within the Cohen syndrome critical region on 8q22 (Kolehmainen et al., 2003), that manifest as a recessive disorder. Only one of them, #10456 (**Supplementary Tables S2, 3**), exhibited clinical features typical of this syndrome, including microcephaly, peculiar facies, hypotonia, non-progressive cognitive deficit, myopia, retinal dystrophy, neutropenia, and truncal obesity. Moreover, the analysis performed through the Surecall Software (Agilent) suggested that this patient might carry an intragenic *VPS13B* deletion. The specifically designed MLPA confirmed a deletion involving exons 4-7 with a minimum length of 19.46 kb. Thus the two *VPS13B* variants, that were inherited from the two heterozygous parents (**Supplementary Table S3**), were considered as causative of the Cohen Syndrome in this patient. The MLPA analysis was also extended to the other four patients, although they did not show any sign of Cohen syndrome, but no deletions/duplications were detected in *VPS13B*. As a consequence, their variants were not counted among the causative ones.

Among the clinically relevant variants 12.2% were variants predicted to cause a splicing defect falling within the highly conserved dinucleotide at the acceptor or donor splice site. Only one putative splicing variant, namely *SCN1A* c.695-6_695-3delTTTC in patient #10457, located at position -3 from exon 9, did not fall within the canonical dinucleotide but just upstream the acceptor splice site. Alteration of *SCN1A* (sodium voltage-gated channel alpha subunit 1) are associated with Dravet Syndrome (OMIM #607208, **Supplementary Table S1**), a Mendelian disorder characterized by drug-resistant epilepsy with early-onset seizures (usually during the first year of life) after apparently normal development, and usually induced by fever, at least initially (Dravet, 1978). The variant pathogenicity was evaluated following the ACMG criteria that classified it as pathogenic considering

**TABLE 2 |** Diagnostic yield in males and females.

| Characteristic | Males<br>Absolute count (percentage over total males) | Males carrying variants<br>Absolute count (percentage over total males in each category) | Females<br>Absolute count (percentage over total females) | Females carrying variants<br>Absolute count (percentage over total females in each category) | Significance of the difference in diagnostic yield between males and females[a] |
|---|---|---|---|---|---|
| **Total count** | 202 | 31 (15.3) | 136 | 40 (29.4) | P=0.0019 |
| **Average age (years)** | 8.2 | 8.6 | 9.8 | 12.0 | |
| **ID** | | | | | |
| Total | 125 (61.9) | 23 (18.4) | 97 (71.3) | 26 (26.8) | P=0.037 |
| Pure ID[b] | 21 (10.4) | 2 (9.5) | 23 (16.9) | 5 (21.7) | |
| **Comorbidities associated with ID** | | | | | |
| ASD | 19 (9.4) | 2 (10.5) | 9 (6.6) | 3 (33.3) | |
| Motor disorders | 69 (34.2) | 11 (15.9) | 54 (39.7) | 15 (27.8) | |
| Language disorder | 62 (30.7) | 8 (12.9) | 47 (34.6) | 15 (31.9) | P=0.01 |
| Psychiatric disorder | 5 (2.5) | 2 (40.0) | 8 (5.9) | 6 (75.0) | |
| Attention/Learning/ Behavioral disorder | 35 (17.3) | 7 (20.0) | 24 (17.6) | 7 (29.2) | |
| Epilepsy | 38 (18.8) | 7 (18.4) | 35 (25.7) | 12 (34.3) | |
| Dysmorphism/Congenital defects | 44 (21.8) | 11 (25.0) | 30 (22.1) | 12 (40.0) | |
| **Other major diagnoses without ID[c]** | | | | | |
| ASD | 33 (16.3) | 2 (6.1) | 10 (7.4) | 4 (40.0) | P=0.0067 |
| Motor disorder | 17 (8.4) | 4 (23.5) | 7 (5.1) | 4 (57.1) | |
| Language disorder | 18 (8.9) | 2 (11.1) | 9 (6.6) | 3 (33.3) | |
| Psychiatric disorder | 2 (1.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | |
| Attention/Learning/ Behavioral disorder | 16 (7.9) | 2 (12.5) | 6 (4.4) | 2 (33.3) | |
| Pure epilepsy | 20 (9.9) | 1 (5.0) | 15 (11.0) | 6 (40.0) | P=0.01 |
| Dysmorphism/ Congenital defects | 14 (6.9) | 1 (7.1) | 6 (4.4) | 0 (0.0) | |

*[a]Significance has been calculated by 2 × 2 contingency table. Only p value < 0.05 are reported.*
*[b]Pure ID, refers to a diagnosis of intellectual disability in the absence of distinct comorbidities such as ASD, attention/learning/behavioral disorders, epilepsy, psychiatric disorders, and dysmorphic features. However, considering the high prevalence of motor and language impairment in subjects with ID, these two clinical features were included in the definition.*
*[c]Each additional diagnosis other than intellectual disability is considered including a potential overlap among the clinical manifestations, so that some patients may have, for instance, ASD, and epilepsy or ASD, and a learning disorder at the same time. Conversely, pure epilepsy is intended here as an isolated clinical entity, i.e., it refers to the presence of symptomatic chronic seizures in the absence of ID, motor or language disorders; ASD, attention/learning/behavioral disorders, psychiatric disorders, or dysmorphic traits.*
*Abbreviations: ASD, autism spectrum disorder; ID, intellectual disability; LP, likely pathogenic; P, pathogenic.*

that it arose *de novo* in a patient with typical clinical features of Dravet syndrome. An *in vitro* splicing assay was set up to demonstrate the role of this novel variation. Sequencing of the PCR products from mini-genes bearing the mutant and wild type alleles revealed that c.695-6_695-3delTTTC causes aberrant splicing. The transcript originating from the wild-type construct resulted in a product of 532 bp whereas the variant construct produced a smaller transcript of 262 bp, corresponding to the size of the empty plasmid (**Supplementary Figure S1**). Sequence analysis of these transcripts revealed that the wild-type product results from the correct splicing of *SCNA1* exon 9 whereas the variant product results from complete skipping of exon 9 and junction of the two plasmid exons. The predicted product is a protein lacking the 90 amino acids encoded by exon 9 corresponding to a portion of the first transmembrane domain.

In the cohort there were a total of 12 individuals for whom at least one first/second degree relative with a clinical phenotype attributable to NDDs was reported, suggestive of autosomal heredity, although the phenotype was not necessarily the same observed in the index case as different diagnosis might represent variable expressivity of the same underlying genetic cause. A pathogenic variant segregating in accordance with a autosomal dominant disorder was identified in 3 of these cases (#10617,#211-20, #482-12) whereas in one case (#422-20) the variant arose *de novo*. In this latter case the patient had a sister and a first cousin (from maternal lineage) with NNDs and the mother reported four miscarriages, thus suggesting a second independent not investigated genetic alteration in this family (i.e., chromosomal balanced translocation in the mother).

## Genes Bearing Clinically Relevant Variants

Variants were identified in a total of 48 genes from the 221 genes of the panel. The most represented functional categories were those correlated to brain function and development (71%), regulation of gene expression and differentiation (58%), and cell signaling pathways (33%). Fourteen genes were mutated in more than one patient accounting for 19.7% of the cases receiving a molecular diagnosis. The most involved genes were *SLC2A1* altered in five patients, *SCNA1* in 4 patients, and *ANKRD11*, *ATP1A2*, *CACNA1A*, *FOXP1*, and *GNAS*, each altered in three patients (**Supplementary Table S3**).

## Complex Phenotype in a Patient With Two Pathogenic Variants

Patient # 298-19, a 12-year-old boy, carried a splicing variant in *TBX1* and a frameshift in *UBE3A* (**Supplementary Tables S2, 3**). Both variations were of unknown origin as they were not present in the mother, and the father's DNA was not available. The patient was referred for suspicion of Angelman syndrome (AS), characterized by severe intellectual disability behavioral disorder and unexplained falls. He showed developmental delay (Brunet-Lézine DQ = 48), predominantly in the performance, visuo-perceptive coordination, and communication. The EEG revealed occipital paroxysms in the occipital area and MRI showed a thinning of the *corpus callosum* and reduced white matter signaling. *UBE3A* fits the AS characteristics of this patient, and it might be hypothesized that the variant arose *de novo* in the maternal germline cells, since this gene is subjected to paternal imprinting (OMIM # 105830). The other altered gene, *TBX1*, has been recently reported to be involved in agenesis of the *corpus callosum* and reduced white matter in haploinsufficient mice (Hiramoto et al., 2021), which were both features displayed by the patient that are not typical of AS. The complex phenotype might thus be the result of two different causative variants in different genes.

## Variants Identified in ID Patients

ID was present in 222 patients (125 males and 97 females), and in 44 of them, it was not associated with either ASD or epilepsy, behavioral disorders, or dysmorphic traits. However, considering the high prevalence of motor and/or language impairment in ID patients, these two clinical features were still considered in the definition of "pure ID." Clinically relevant variants were found in 49 (22.1%) patients with ID, and seven of them did not present comorbidities other than motor or language difficulties. A higher proportion of females with ID (26.8%) carried causative variants in comparison to males (18.4%) although with a borderline significance ($p$ = 0.037, **Table 2**).

## Variants Identified in Patients With Epilepsy

A total of 135 patients (76 males and 59 females) manifested epilepsy and in 73 of them it was associated with ID. Among the total, 34 (25.6%) patients carried causative variants in 23 different genes (**Supplementary Table S2**). Thirty-five patients manifested epilepsy without any other comorbidity and seven of them (20.0%) carried variants in six genes, namely *KCNQ2*, *NALCN*, *PAK3*, *PRRT2*, *SCN1A*, and *SLC2A1* that are among the well-known genes causing epilepsies or syndromes including epilepsy (Wang et al., 2017).

A significantly higher proportion of females with epilepsy carried pathogenic variants in comparison to males (38.9% vs. 14.4% $p$ = 0.001) and the difference was also evident in isolated epilepsy (40.0% females *versus* 5.0% males; $p$ = 0.01, **Table 2**).

## Variants Identified in Patients With ASD

ASD was present in 71 patients (52 males and 19 females) and 11 of them (15.5%) carried pathogenic variants in 11 different genes (**Supplementary Table S3**). All these genes were listed in SFARI (https://gene.sfari.org/), a comprehensive database including genes associated with autism, and six of them (*KANSL1, SLC9A6, CDKL5, SYNGAP, SCN1A,* and *FOXP1*) are currently classified in the first category that includes those that have been clearly implicated in ASD with high confidence.

As for epilepsy, a significantly higher proportion of females with ASD carried P/LP compared with males either with other co-morbidity (36.8% *vs.* 7.6%; $p$ = 0.0026) or isolated (40.0% *vs.* 6.1%; $p$ = 0.0067).

## DISCUSSION

The clinical heterogeneity of patients suffering from NNDs and the wide spectrum of comorbidities render large genomic unbiased NGS analyses the best approach to identify the underlying genetic defects. On the one hand, WES is the gold standard method to identify pathogenic variants affecting the primary protein structure including those that are present in less frequently mutated and still undiscovered genes. On the other hand, WGS offers, at higher costs, additional benefits since it can uncover regulatory and deep intronic alterations, as well as CNV by using appropriate bioinformatics tools. However, in most clinical diagnostic centers, systematic use of WES and WGS poses the problem of the high costs due to the need of storing and analyzing huge amounts of data and co-sequencing both patients' parents. Therefore, the creation of affordable NGS panels containing a careful selection of genes involved in the diseases of interest may be a good compromise to reach a molecular diagnosis in a cost- and time-effective manner, with the possibility of limiting the NGS analysis to the index case to detect the causal variant and then investigate the parental origin by conventional sequencing targeting of the variant. Targeted gene panels depending on the platform used, require from one third to half of the expenses needed to analyze and store the WES data of the trios in terms of reagents, bioinformatics supplies and operator costs. Thus these tests might be easily implemented in the Hospitals that already represent the tertiary reference center for specialized genetic tests.

Aiming to avoid large investments but, at the same time, to obtain adequate coverage of the selected genes (mean 150 X) with the possibility of identifying causative variants in a heterogeneous population of NND patients, we designed a panel of 221 genes that currently represent the most promising ones for our diagnostic purposes. This panel allowed us to identify a genetic defect contributing to the phenotype in 21% of the patients analyzed, which is in line with the results described in a recent meta-analysis by Stefanski et al. (2021) showing an overall yield of 22.6% for targeted gene panels sequencing, compared to 27.2% for WES. Therefore, this result can be considered a success as our cohort includes individuals with heterogeneous clinical outcomes, which limits the number of genes associated with each disorder as opposed to highly specific panels designed for homogeneous groups of patients used in several studies evaluated by the meta-analysis. Grouping patients on the basis of two main clinical phenotypes, the yield was 26.0% in epilepsy and 17.9% in ASD, in line with the data detected by the meta-analysis (24% and 17.1%, respectively).

Increased male prevalence has been repeatedly reported in several studies as well as in our NDD cohort (**Table1**). Epidemiologic studies report a 30%–50% excess of males in ID (Polyak et al., 2015) and, in ASD, a male-to-female ratio of 4: 1 that increases to 7:1 in high-functioning autism (McLaren and Bryson, 1987). Several mechanisms might cause this bias, including sex hormones, as suggested for testosterone exposure in fetal life in ASD (Baron-Cohen et al., 2015). Possible genetic mechanisms underlying this gender seem not to involve X-linked variants, since X-linked ID is too rare to account for the 30% excess of males with this disorder and a recent burden analysis on a large group of patients with NDDs reported that X-linked causal variants were carried in similar proportions of males (6%) and females (6.9%) (Martin et al., 2021). Our panel allows good coverage of X-linked genes, since it includes 67 of the 130 X-linked genes that have been implicated in NNDs, and detected X-linked variants in two males and 10 females, with seven originated *de novo*. This enriched X-linked *de novo* variant in females is expected and might be in part explained by a higher lethality for mutations in the hemizygous male and milder phenotypes in females, in particular for those genes that escapers of X-inactivation (Tukiainen et al., 2017). In fact, it has been observed that 40% of the genes that are enriched in *de novo* mutations in females with NDDs escape X-inactivation (Turner et al., 2019), as DDX3X and *NAA10*, mutated in two patients of our cohort.

Despite the prevailing number of males present in our cohort, a significantly higher number of variants were identified in females. Global data showed relevant variants in 29.4% of females *versus* 15.3% of males ($p = 0.0046$), with even a higher bias in epilepsy and ASD (**Table 2**). This higher yield in females has emerged previously (Turner et al., 2019). For instance, an excess of pathogenic CNV in females has been shown in a large cohort of NDD patients with an O. R of 1.46 ($p = 8 \times 10^{-8}$) and an excess of deleterious SNV in females (O.R = 1.34) has been found in ASD (Jacquemont et al., 2014). These data are consistent with the "female protective model" that suggests that the clinical manifestation of NNDs in females requires a higher "mutational burden" to reach the threshold for a diagnosis.

Although family history for NNDs was referred in the 9.5% of the cases only in few of them the inheritance pattern suggested an autosomal disorder (even with reduced penetrance). The low percentage of familial cases in the whole cohort is attributable to the reduced fitness associated with NDDs and to an increased rate of *de novo* variants as recently demonstrated by a large WES study (McRae et al., 2017) reporting that these alterations account for approximately half of the genetic architecture of severe developmental disorders; in the present study *de novo* variations were identified in the 64.4% of the analyzed trios. The detection of pathogenic variants in only few familial cases might be due the limitations of the utilized gene panel that don't cover intronic and regulatory regions and includes only most promising genes as well as to the presence of more complex pattern of inheritance in many patients. In fact the wide phenotypic variability of the NDDs likely reflects the interaction of multiple genes within an individual's genetic background and different genetic combinations might exist among patients and affected member of the same family. The 96% of the detected VUS were missense (data not shown) and although in some families these segregated consistently with the phenotype, they remained classified as VUS. Protein disrupting and *de novo* variations are less ambiguous and in most cases classified as pathogenic by bioinformatics tools. However some of the inherited missense VUS (maybe with milder effect) might actually contribute to patients' phenotype. Although challenging, VUS should be more extensively studied, periodically reanalyzed and much more efforts should be dedicated to decipher their functional role.

In our subjects with ID, most variants were found in genes related to transcriptional regulation, brain development, and neuronal and intracellular signaling while genes mutated in ASD were associated with the same functional categories but also with the control of cell structure and polarity. ASD is a pervasive NDD with multifactorial etiology (Walsh et al., 2008) and previous attempts to identify causative candidate genes supported the implication of synapse pathology and abnormal neural network formation as the two main pathogenic mechanisms (Rylaarsdam and Guemez-Gamboa, 2019). In line with these findings, most gene variants detected in our ASD subjects are related to the synaptic structure or signaling pathways (*CDKL5*, *DOCK8, GNAS*, *SCN1A*, *SLC9A6*, *SYNGAP1*), or transcriptional regulation with a putative role in the patterning of neural circuits (*FOXP1*) (Walsh et al., 2008; Glessner et al., 2017) (**Supplementary Tables S1, 2**).

In the seven patients with isolated epilepsy carrying causative variants, the genes involved play roles in synaptic signaling, cell and energy metabolism, and intracellular signal transduction. In particular, *KCNQ2, NALCN*, and *SCN1A* are responsible for ion conductance and membrane polarity, *PRRT2* codes for a protein interacting with key components of the presynaptic density, *PAK3* is involved in intracellular signal transduction pathways, and *SLC2A1* encodes GLUT1, mainly expressed in the blood-brain barrier (BBB) and working as a key transporter of glucose, the first-choice energy source in the central nervous system (CNS).

It must be underlined that *SLC2A1* was the most frequently altered gene in both comorbid and isolated epilepsy, and in the whole NDDs cohort irrespectively of the disease considered. GLUT1 deficiency syndrome has been characterized as a specific clinical entity encompassing neurodevelopmental impairment, movement disorders, and seizures, and it has been predominately associated with pathogenic *SLC2A1* variants (Graham, 2012). Early detection of this genetic alteration is clinically important, since affected subjects may benefit from ketogenic diets, which may mitigate symptoms and even prevent their progression provided that treatment is started soon when the CNS development is still not complete.

## CONCLUSION

This work shows the utility of a customized 221 gene panel that allowed to reach a molecular diagnosis in 21% of patients with NDDs and supports the opportunity of using a common gene panel as a second-tier test for different forms of NNDs. This approach allows to update the gene panel according to changes in the diagnostic criteria of this clinically heterogeneous group of disorders which would consequently increase the diagnostic power. Given the diagnostic yield and the potential clinical benefits, this testing should be offered to all patients with global developmental delay, intellectual disability, and/or ASD. In this scenario, WES and WGS might be reserved for the definition of complex phenotypes or familial unexplained cases, allowing a more efficient allocation of the resources.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

## ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

## AUTHOR CONTRIBUTIONS

SM and MG coordinated the study and performed the genetic and bioinformatics analysis. DV, SR, SF, AM, and MV conducted the genetic test in laborator. CP, AmP, AS, FS, GM, AlP, PR, FV, AZ, IR, and MV recruited and collected the anamnestic data for the patients. CP and SM performed the literature search, reviewed and extracted data from the papers. MG, SM, CiP, and UD elaborated the clinical data and coordinated the manuscript writing. All authors discussed the results and contributed to the final manuscript.

## FUNDING

Fondi Vari del Dipartimento di Scienze della Salute, Università del Piemonte Orientale.

## ACKNOWLEDGMENTS

We are grateful to all the patients and their families.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.875182/full#supplementary-material

## REFERENCES


Almuhtaseb, S., Oppewal, A., and Hilgenkamp, T. I. M. (2014). Gait Characteristics in Individuals with Intellectual Disabilities: A Literature Review. *Res. Dev. Disabil.* 35, 2858–2883. doi:10.1016/J.RIDD.2014.07.017

American Psychiatric Association (2013). "Neurodevelopmental Disorders," in *Diagnostic and Statistical Manual of Mental Disorders* (Arlington, VA: American Psychiatric Publishing).

Babu, D., Vannelli, S., Fanelli, A., Mellone, S., Baffico, A. M., Corrado, L., et al. (2021). Variants in the 5′UTR Reduce SHOX Expression and Contribute to SHOX Haploinsufficiency. *Eur. J. Hum. Genet.* 29, 110–121. doi:10.1038/S41431-020-0676-Y

Baron-Cohen, S., Auyeung, B., Nørgaard-Pedersen, B., Hougaard, D. M., Abdallah, M. W., Melgaard, L., et al. (2015). Elevated Fetal Steroidogenic Activity in Autism. *Mol. Psychiatry* 20, 369–376. doi:10.1038/MP.2014.48

Bertrand, J., Mars, A., Boyle, C., Bove, F., Yeargin-Allsopp, M., and Decoufle, P. (2001). Prevalence of Autism in a United States Population: The Brick Township, New Jersey, Investigation. *Pediatrics* 108, 1155–1161. doi:10.1542/PEDS.108.5.1155

Boyle, C. A., Boulet, S., Schieve, L. A., Cohen, R. A., Blumberg, S. J., Yeargin-Allsopp, M., et al. (2011). Trends in the Prevalence of Developmental Disabilities in US Children, 1997-2008. *Pediatrics* 127, 1034–1042. doi:10.1542/PEDS.2010-2989

Chakrabarti, S., and Fombonne, E. (2001). Pervasive Developmental Disorders in Preschool Children. *JAMA* 285, 3093–3099. doi:10.1001/JAMA.285.24.3093

Charman, T., Pickles, A., Simonoff, E., Chandler, S., Loucas, T., and Baird, G. (2011). IQ in Children with Autism Spectrum Disorders: Data from the Special Needs and Autism Project (SNAP). *Psychol. Med.* 41, 619–627. doi:10.1017/S0033291710000991

Di Gregorio, E., Riberi, E., Belligni, E. F., Biamino, E., Spielmann, M., Ala, U., et al. (2017). Copy Number Variants Analysis in a Cohort of Isolated and Syndromic Developmental Delay/Intellectual Disability Reveals Novel Genomic Disorders, Position Effects and Candidate Disease Genes. *Clin. Genet.* 92, 415–422. doi:10.1111/CGE.13009

Doan, R. N., Lim, E. T., Lim, E. T., De Rubeis, S., Betancur, C., Cutler, D. J., et al. (2019). Recessive Gene Disruptions in Autism Spectrum Disorder. *Nat. Genet.* 51, 1092–1098. doi:10.1038/S41588-019-0433-8

Dravet, C. (1978). Les epilepsies graves de l'enfant. *Vie médicale* 8, 543–548.

Glessner, J. T., Li, J., Li, J., Wang, D., March, M., Lima, L., et al. (2017). Copy Number Variation Meta-Analysis Reveals a Novel Duplication at 9p24 Associated with Multiple Neurodevelopmental Disorders. *Genome Med.* 9, 106. doi:10.1186/S13073-017-0494-1

Graham, J. M. (2012). GLUT1 Deficiency Syndrome as a Cause of Encephalopathy that Includes Cognitive Disability, Treatment-Resistant Infantile Epilepsy and a Complex Movement Disorder. *Eur. J. Med. Genet.* 55, 332–334. doi:10.1016/J.EJMG.2011.11.009

Henderson, L. B., Applegate, C. D., Wohler, E., Sheridan, M. B., Hoover-Fong, J., and Batista, D. A. S. (2014). The Impact of Chromosomal Microarray on Clinical Management: A Retrospective Analysis. *Genet. Med.* 16, 657–664. doi:10.1038/GIM.2014.18

Hiramoto, T., Sumiyoshi, A., Yamauchi, T., Tanigaki, K., Shi, Q., Kang, G., et al. (2021). Tbx1, a Gene Encoded in 22q11.2 Copy Number Variant, Is a Link Between Alterations in Fimbria Myelination and Cognitive Speed in Mice. *Mol. Psychiatry* 27, 929–938. doi:10.1038/S41380-021-01318-4

Jacquemont, S., Coe, B. P., Hersch, M., Duyzend, M. H., Krumm, N., Bergmann, S., et al. (2014). A Higher Mutational Burden in Females Supports a "Female Protective Model" in Neurodevelopmental Disorders. *Am. J. Hum. Genet.* 94, 415–425. doi:10.1016/J.AJHG.2014.02.001

Jensen, M., and Girirajan, S. (2017). Mapping a Shared Genetic Basis for Neurodevelopmental Disorders. *Genome Med.* 9, 109. doi:10.1186/S13073-017-0503-4

Kazeminasab, S., Taskiran, I. I., Fattahi, Z., Bazazzadegan, N., Hosseini, M., Rahimi, M., et al. (2018). CNKSR1 Gene Defect Can Cause Syndromic Autosomal Recessive Intellectual Disability. *Am. J. Med. Genet.* 177, 691–699. doi:10.1002/AJMG.B.32648

Kolehmainen, J., Black, G. C. M., Saarinen, A., Chandler, K., Clayton-Smith, J., Träskelin, A.-L., et al. (2003). Cohen Syndrome Is Caused by Mutations in a Novel Gene, COH1, Encoding a Transmembrane Protein with a Presumed Role in Vesicle-Mediated Sorting and Intracellular Protein Transport. *Am. J. Hum. Genet.* 72, 1359–1369. doi:10.1086/375454

Kopanos, C., Tsiolkas, V., Kouris, A., Chapple, C. E., Albarca Aguilera, M., Meyer, R., et al. (2019). VarSome: The Human Genomic Variant Search Engine. *Bioinformatics* 35, 1978–1980. doi:10.1093/BIOINFORMATICS/BTY897

Leonard, H., Glasson, E., Nassar, N., Whitehouse, A., Bebbington, A., Bourke, J., et al. (2011). Autism and Intellectual Disability Are Differentially Related to Sociodemographic Background at Birth. *PLoS One* 6, e17875. doi:10.1371/JOURNAL.PONE.0017875

Martin, H. C., Gardner, E. J., Gardner, E. J., Samocha, K. E., Kaplanis, J., Akawi, N., et al. (2021). The Contribution of X-Linked Coding Variation to Severe Developmental Disorders. *Nat. Commun.* 12, 1–13. doi:10.1038/s41467-020-20852-3

McLaren, J., and Bryson, S. E. (1987). Review of Recent Epidemiological Studies of Mental Retardation: Prevalence, Associated Disorders, and Etiology. *Am. J. Ment. Retard.* 92, 243–254.

McRae, J. F., Clayton, S., Fitzgerald, T. W., Kaplanis, J., Prigmore, E., Rajan, D., et al. (2017). Prevalence and Architecture of De Novo Mutations in Developmental Disorders. *Nature* 542, 433–438. doi:10.1038/NATURE21062

Polyak, A., Rosenfeld, J. A., and Girirajan, S. (2015). An Assessment of Sex Bias in Neurodevelopmental Disorders. *Genome Med.* 7, 94–104. doi:10.1186/S13073-015-0216-5

Robertson, J., Hatton, C., Emerson, E., and Baines, S. (2015). Prevalence of Epilepsy Among People with Intellectual Disabilities: A Systematic Review. *Seizure* 29, 46–62. doi:10.1016/J.SEIZURE.2015.03.016

Rylaarsdam, L., and Guemez-Gamboa, A. (2019). Genetic Causes and Modifiers of Autism Spectrum Disorder. *Front. Cell. Neurosci.* 13, 385. doi:10.3389/FNCEL.2019.00385/BIBTEX

Stefanski, A., Calle-López, Y., Leu, C., Pérez-Palma, E., Pestana-Knight, E., and Lal, D. (2021). Clinical Sequencing Yield in Epilepsy, Autism Spectrum Disorder, and Intellectual Disability: A Systematic Review and Meta-analysis. *Epilepsia* 62, 143–151. doi:10.1111/EPI.16755

Totsika, V., Hastings, R. P., Emerson, E., Lancaster, G. A., and Berridge, D. M. (2011). A Population-Based Investigation of Behavioural and Emotional Problems and Maternal Mental Health: Associations with Autism Spectrum Disorder and Intellectual Disability. *J. Child. Psychol. Psychiatry* 52, 91–99. doi:10.1111/J.1469-7610.2010.02295.X

Tukiainen, T., Villani, A. C., Villani, A.-C., Yen, A., Rivas, M. A., Marshall, J. L., et al. (2017). Landscape of X Chromosome Inactivation Across Human Tissues. *Nature* 550, 244–248. doi:10.1038/NATURE24265

Turner, T. N., Wilfert, A. B., Bakken, T. E., Bernier, R. A., Pepper, M. R., Zhang, Z., et al. (2019). Sex-Based Analysis of De Novo Variants in Neurodevelopmental Disorders. *Am. J. Hum. Genet.* 105, 1274–1285./ATTACHMENT/9F8DDCA7-5D69-4E9A-85EA-1E2DC42F1D42/MMC2.XLSX. doi:10.1016/J.AJHG.2019.11.003

Van Bokhoven, H. (2011). Genetic and Epigenetic Networks in Intellectual Disabilities. *Annu. Rev. Genet.* 45, 81–104. doi:10.1146/ANNUREV-GENET-110410-132512

Walsh, C. A., Morrow, E. M., and Rubenstein, J. L. R. (2008). Autism and Brain Development. *Cell* 135, 396–400. doi:10.1016/J.CELL.2008.10.015

Wang, J., Lin, Z.-J., Liu, L., Xu, H.-Q., Shi, Y.-W., Yi, Y.-H., et al. (2017). Epilepsy-Associated Genes. *Seizure* 44, 11–20. doi:10.1016/J.SEIZURE.2016.11.030


**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.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

*Copyright © 2022 Mellone, Puricelli, Vurchio, Ronzani, Favini, Maruzzi, Peruzzi, Papa, Spano, Sirchia, Mandrile, Pelle, Rasmini, Vercellino, Zonta, Rabbone, Dianzani, Viri and Giordano. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*

# 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:** frontiersin.org/about/contact

**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